Brownian Motion

Draft version of May 25, 2008

Peter Mörters and Yuval Peres

Contents

Foreword 7

List of frequently used notation 9

Chapter 0. Motivation 13

Chapter 1. Definition and first properties of Brownian motion 21
1. Paul Lévy's construction of Brownian motion 21
2. Continuity properties of Brownian motion 27
3. Nondifferentiability of Brownian motion 31
4. The Cameron-Martin theorem 37
Exercises 38
Notes and Comments 41

Chapter 2. Brownian motion as a strong Markov process 43
1. The Markov property and Blumenthal's 0-1 Law 43
2. The strong Markov property and the reflection principle 46
3. Markov processes derived from Brownian motion 53
4. The martingale property of Brownian motion 57
Exercises 64
Notes and Comments 68

Chapter 3. Harmonic functions, transience and recurrence 69
1. Harmonic functions and the Dirichlet problem 69
2. Recurrence and transience of Brownian motion 75
3. Occupation measures and Green's functions 80
4. The harmonic measure 87
Exercises 94
Notes and Comments 96

Chapter 4. Hausdorff dimension: Techniques and applications 97
1. Minkowski and Hausdorff dimension 97
2. The mass distribution principle 106
3. The energy method 108
4. Frostman's lemma and capacity 111
Exercises 117
Notes and Comments 119

Chapter 5. Brownian motion and random walk 121
1. The law of the iterated logarithm 121
2. Points of increase for random walk and Brownian motion 126
3. The Skorokhod embedding problem 129
4. The Donsker invariance principle 134
5. The arcsine laws 137
Exercises 142
Notes and Comments 144

Chapter 6. Brownian local time 147
1. The local time at zero 147
2. A random walk approach to the local time process 158
3. The Ray-Knight theorem 163
4. Brownian local time as a Hausdorff measure 171
Exercises 179
Notes and Comments 181

Chapter 7. Stochastic integrals and applications 183
1. Stochastic integrals with respect to Brownian motion 183
2. Conformal invariance and winding numbers 194
3. Tanaka's formula and Brownian local time 202
4. Feynman-Kac formulas and applications 206
Exercises 213
Notes and Comments 215

Chapter 8. Potential theory of Brownian motion 217
1. The Dirichlet problem revisited 217
2. The equilibrium measure 220
3. Polar sets and capacities 226
4. Wiener's test of regularity 233
Exercises 236
Notes and Comments 238

Chapter 9. Intersections and self-intersections of Brownian paths 239
1. Intersection of paths: existence and Hausdorff dimension 239
2. Intersection equivalence of Brownian motion and percolation limit sets 247
3. Multiple points of Brownian paths 256
4. Kaufman's dimension doubling theorem 264
Exercises 270
Notes and Comments 272

Chapter 10. Exceptional sets for Brownian motion 275
1. The fast times of Brownian motion 275
2. Packing dimension and limsup fractals 283
3. Slow times of Brownian motion 292
4. Cone points of planar Brownian motion 296
Exercises 306
Notes and Comments 309

Appendix I: Hints and solutions for selected exercises 311

Appendix II: Background and prerequisites 331
1. Convergence of distributions on metric spaces 331
2. Gaussian random variables 334
3. Martingales in discrete time 337
4. The max-flow min-cut theorem 342

Index 345

Bibliography 349

Foreword

The aim of this book is to introduce Brownian motion as the central object of probability and discuss its properties, putting particular emphasis on the sample path properties. Our hope is to capture as much as possible the spirit of Paul Lévy's investigations on Brownian motion, by moving quickly to the fascinating features of the Brownian motion process, and filling in more and more details into the picture as we move along.

Inevitably, while exploring the nature of Brownian paths one encounters a great variety of other subjects: Hausdorff dimension serves from early on in the book as a tool to quantify subtle features of Brownian paths, stochastic integrals helps us to get to the core of the invariance properties of Brownian motion, and potential theory is developed to enable us to control the probability the Brownian motion hits a given set.

An important idea of this book is to make it as *interactive* as possible and therefore we have included more than 100 exercises collected at the end of each of the ten chapters. Exercises marked with a diamond have either a hint, a reference to a solution, or a full solution given in the appendix. Exercises marked with a star are more challenging. We have also marked some theorems with a star to indicate that the results will not be used in the remainder of the book and may be skipped on first reading.

This book grew out of lectures given by Yuval Peres at the Statistics Department, University of California, Berkeley in Spring 1998. We are grateful to the students who attended the course and wrote the first draft of the notes: Diego Garcia, Yoram Gat, Diogo A. Gomes, Charles Holton, Frédéric Latrémolière, Wei Li, Ben Morris, Jason Schweinsberg, Bálint Virág, Ye Xia and Xiaowen Zhou. The first draft of these notes was edited by Bálint Virág and Elchanan Mossel and at this stage corrections were made by Serban Nacu and Yimin Xiao. The notes were distributed via the internet and turned out to be very popular — this demand motivated us to expand these notes to a full book hopefully retaining the character.

Several people read drafts of the book, supplied us with helpful lists of corrections or tested exercises. We thank Jian Ding, Richard Kiefer, Achim Klenke, Nathan Levy, Arjun Malhotra, Marcel Ortgiese, Jeff Steif, and Kamil Szczegot.

Peter Mörters lectured on the topics of this book in the Graduate School in Mathematical Sciences at the University of Bath in Autumn 2003, thanks are due to the audience, and in particular to Alex Cox and Pascal Vogt, for their contributions.

Yuval Peres thanks Pertti Mattila for the invitation to lecture on this material at the joint summer school in Jyväskyla, August 1999, and Peter Mörters thanks Michael Scheutzow for the invitation to lecture at the winter school of the Berlin graduate school in probability in Stralsund, April 2003.

Peter Mörters
Yuval Peres

List of frequently used notation

Numbers:

$\lceil x \rceil$	the smallest integer bigger or equal to x
$\lfloor x \rfloor$	the largest integer smaller or equal to x

Topology of Euclidean space $\mathbb{R}^d$:

$\mathbb{R}^d$	Euclidean space consisting of all column vectors $x = (x_1, \dots, x_d)^{\mathrm{T}}$
$\lvert \cdot \rvert$	Euclidean norm $\lvert x \rvert = \sqrt{\sum_{i=1}^{d} x_i^2}$
$\mathcal{B}(x, r)$	the open ball of radius $r > 0$ centred in $x \in \mathbb{R}^d$, i.e. $\mathcal{B}(x, r) = \{y \in \mathbb{R}^d \colon \lvert x - y \rvert < r\}$
$\overline{U}$	closure of the set $U \subset \mathbb{R}^d$
∂U	boundary of the set $U \subset \mathbb{R}^d$
$\mathfrak{B}(A)$	the collection of all Borel subsets of $A \subset \mathbb{R}^d$.

Binary relations:

$a \wedge b$	the minimum of a and b
$a \vee b$	the maximum of a and b
$X \overset{\mathrm{d}}{=} Y$	the random variables X and Y have the same distribution
$a(n) \asymp b(n)$	the ratio of the two sides is bounded from above and below by positive constants that do not depend on n
$a(n) \sim b(n)$	the ratio of the two sides converges to one.

Vectors, functions, and measures:

I_d	$d \times d$ identity matrix
$\mathbb{1}_A$	indicator function with $\mathbb{1}_A(x) = 1$ if $x \in A$ and 0 otherwise
δ_x	Dirac measure with mass concentrated on x, i.e. $\delta_x(A) = 1$ if $x \in A$ and 0 otherwise
f^+	the positive part of the function f, i.e. $f^+(x) = f(x) \vee 0$
f^-	the negative part of the function f, i.e. $f^-(x) = -(f(x) \wedge 0)$
$\mathcal{L}_d$ or $\mathcal{L}$	Lebesgue measure on $\mathbb{R}^d$
$\sigma_{x,r}$	$(d-1)$-dimensional surface measure on $\partial\mathcal{B}(x,r) \subset \mathbb{R}^d$ if $x = 0$, $r = 1$ we also write $\sigma = \sigma_{0,1}$
$\varpi_{x,r}$	uniform distribution on $\partial\mathcal{B}(x,r)$, $\varpi_{x,r} = \frac{\sigma_{x,r}}{\sigma_{x,r}(\partial\mathcal{B}(x,r))}$, if $x = 0$, $r = 1$ we also write $\varpi = \varpi_{0,1}$.

Function spaces:

$\mathcal{C}(K)$	topological space of all continuous functions on the compact $K \subset \mathbb{R}^d$, equipped with the supremum norm $\|f\| = \sup_{x\in K} |f(x)|$

Probability measures and σ-algebras:

$\mathbb{P}_x$	a probability measure on a measure space $(\Omega, \mathcal{A})$ such that the process $\{B(t) \colon t \geq 0\}$ is a Brownian motion started in x
$\mathbb{E}_x$	the expectation associated with $\mathbb{P}_x$
$\mathfrak{p}(t,x,y)$	the transition density of Brownian motion $\mathbb{P}_x\{B(t) \in A\} = \int_A \mathfrak{p}(t,x,y)\, dy$
$\mathcal{F}^0(t)$	the smallest σ-algebra that makes $\{B(s) \colon 0 \leq s \leq t\}$ measurable
$\mathcal{F}^+(t)$	the right-continuous augmentation $\mathcal{F}^+(t) = \bigcap_{s>t} \mathcal{F}^0(s)$.

Stopping times:

For any Borel sets $A_1, A_2, \ldots \subset \mathbb{R}^d$ and a Brownian motion $B\colon [0,\infty) \to \mathbb{R}^d$,

$$\tau(A_1) := \inf\{t \geq 0\colon\ B(t) \in A_1\}, \qquad \text{the entry time into } A_1,$$

$$\tau(A_1,\ldots,A_n) := \begin{cases} \inf\{t \geq \tau(A_1,\ldots,A_{n-1}) \,:\, B(t) \in A_n\}, & \text{if } \tau(A_1,\ldots,A_{n-1}) < \infty, \\ \infty, & \text{otherwise.} \end{cases}$$

the time to enter A_1 and then A_2 and so on until A_n.

Systems of subsets in $\mathbb{R}^d$:

For any fixed d-dimensional unit cube $\mathsf{Cube} = x + [0,1]^d$ we denote:

$\mathfrak{D}_k$	family of all half-open dyadic subcubes $D = x + \prod_{i=1}^d \big[k_i 2^{-k}, (k_i+1)2^{-k}\big) \subset \mathbb{R}^d$, $k_i \in \{0,\ldots,2^k-1\}$, of sidelength 2^{-k}
$\mathfrak{D}$	all half-open dyadic cubes $\mathfrak{D} = \bigcup_{k=0}^{\infty} \mathfrak{D}_k$ in Cube
$\mathfrak{C}_k$	family of all compact dyadic subcubes $D = x + \prod_{i=1}^d \big[k_i 2^{-k}, (k_i+1)2^{-k}\big] \subset \mathbb{R}^d$, $k_i \in \{0,\ldots,2^k-1\}$, of sidelength 2^{-k}
$\mathfrak{C}$	all compact dyadic cubes $\mathfrak{C} = \bigcup_{k=0}^{\infty} \mathfrak{C}_k$ in Cube.

Potential theory:

For a metric space (E,ρ) and mass distribution μ on E:

$\phi_\alpha(x)$	the α-potential of a point $x \in E$ defined as $\phi_\alpha(x) = \int \frac{d\mu(y)}{\rho(x,y)^\alpha}$,
$I_\alpha(\mu)$	the α-energy of the measure μ defined as $I_\alpha(\mu) = \iint \frac{d\mu(x)\,d\mu(y)}{\rho(x,y)^\alpha}$,
$\mathrm{Cap}_\alpha(E)$	the (Riesz) α-capacity of E defined as $\mathrm{Cap}_\alpha(E) = \sup\{I_\alpha(\mu)^{-1} \,:\, \mu(E)=1\}$.

For a general kernel $K\colon E \times E \to [0,\infty]$:

$U_\mu(x)$	the potential of μ at x defined as $U_\mu(x) = \int K(x,y)\,d\mu(y)$,
$I_K(\mu)$	K-energy of μ defined as $I_K(\mu) = \iint K(x,y)\,d\mu(x)\,d\mu(y)$,
$\mathrm{Cap}_K(E)$	K-capacity of E defined as $\mathrm{Cap}_K(E) = \sup\{I_K(\mu)^{-1} \,:\, \mu(E)=1\}$.

If $K(x,y) = f(\rho(x,y))$ we also write:

$I_f(\mu)$	instead of $I_K(\mu)$,
$\mathrm{Cap}_f(E)$	instead of $\mathrm{Cap}_K(E)$.

Sets and processes associated with Brownian motion:

For a linear Brownian motion $\{B(t) \colon t \geq 0\}$:

$\{M(t) \colon t \geq 0\}$	the maximum process defined by $M(t) = \sup_{s \leq t} B(s)$,
Rec	the set of record points $\{t \geq 0 \colon B(t) = M(t)\}$,
Zero	the set of zeros $\{t \geq 0 \colon B(t) = 0\}$.

For a Brownian motion $\{B(t) \colon t \geq 0\}$ in $\mathbb{R}^d$ for $d \geq 1$:

Graph	the graph $\{(t, B(t)) \colon t \geq 0\} \subset \mathbb{R}^{d+1}$,
Range	the range $\{B(t) \colon t \geq 0\} \subset \mathbb{R}^d$.

CHAPTER 0

Motivation

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. 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 $\{S_n : n \geq 0\}$ is a random walk, the displacements represent the microscopic inputs. When we think about the macroscopic picture, what we mean is questions such as:

- Does S_n drift to infinity?
- Does S_n return to the neighbourhood of the origin infinitely often?
- What is the speed of growth of $\max\{|S_1|, \ldots, |S_n|\}$ as $n \to \infty$?
- What is the asymptotic number of windings of $\{S_n : n \geq 0\}$ around the origin?

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 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.

Any continuous time stochastic process $\{B(t) : t \geq 0\}$ describing the macroscopic features of a random walk should have the following properties:

(1) for all times $0 \leq t_1 \leq t_2 \leq \ldots \leq t_n$ 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 $\{B(t) : t \geq 0\}$ 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 $t \geq 0$ and $h \geq 0$ 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.

We call the process $\{B(t) : t \geq 0\}$ a *Brownian motion* if the drift vector is zero, and the diffusion matrix is the identity. 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 $\{B(t) : t \geq 0\}$. 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 $\{\tilde{B}(t) : t \geq 0\}$ given by

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

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 can be fully described by a standard Brownian motion.

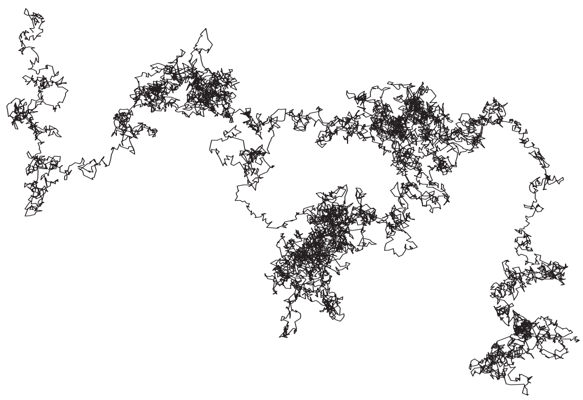

FIGURE 1. The range $\{B(t) : 0 \leq t \leq 1\}$ of a planar Brownian motion

In *Chapter 1* we start exploring Brownian motion by looking at dimension $d = 1$. Here Brownian motion is a random continuous function and we ask about its *regularity*, for example:

- For which parameters α is the random function $B\colon [0, 1] \to \mathbb{R}$ α-Hölder continuous?
- Is the random function $B\colon [0, 1] \to \mathbb{R}$ differentiable?

The surprising answer to the second question was given by Paley, Wiener and Zygmund in 1933: Almost surely, the random function $B\colon [0, 1] \to \mathbb{R}$ is *nowhere* differentiable! This is particularly interesting, as it is not easy to construct a continuous, nowhere differentiable function without the help of randomness. We will give a modern proof of the Paley, Wiener and Zygmund theorem, see Theorem 1.30.

In *Chapter 2* we move to general dimension d. We shall explore the strong Markov property, which roughly says that at suitable random times Brownian motion starts afresh. Among the facts we derive are: Almost surely,

- the set of all points visited by Brownian motion in $d = 2$ has area zero,
- the set of times when Brownian motion in $d = 1$ revisits the origin is uncountable.

Besides these sample path properties, the strong Markov property is also the key to some fascinating distributional identities. It enables us to understand, for example,

- the process $\{M(t) \colon t \geq 0\}$ of the running maxima $M(t) = \max_{0 \leq s \leq t} B(s)$ of a one-dimensional Brownian motion,
- the process $\{T_a \colon a \geq 0\}$ of the first hitting times $T_a = \inf\{t \geq 0 \colon B(t) = a\}$ of level a of a one-dimensional Brownian motion,
- the process of the vertical first hitting positions by a two-dimensional Brownian motion of the lines $\{(x, y) \in \mathbb{R}^2 \colon x = a\}$, as a function of a.

In *Chapter 3* we start exploring the rich relations of Brownian motion to harmonic analysis. To motivate this relation by a discrete analogue, suppose that $\{S_n : n \in \mathbb{N}\}$ is a simple, symmetric random walk in $\mathbb{Z}^2$ started at some $x \in \mathbb{Z}^2$. Here simple and symmetric means that the increments take each of the values $(0, 1), (1, 0), (0, -1), (-1, 0)$ with probability $\frac{1}{4}$. Suppose that $A \subset \mathbb{Z}^2$ is a bounded subset of the two-dimensional lattice and let ∂A be the set of all vertices in $\mathbb{Z}^2 \setminus A$ which are adjacent to a vertex in A. Let $T = \inf\big\{n \geq 0 \colon S_n \notin A\big\}$ be the first exit time from A. Suppose moreover that $\varphi : \partial A \to \mathbb{R}$ is given and define

$$f \colon A \cup \partial A \to \mathbb{R}, \quad f(x) = \mathbb{E}\big[\varphi(S_T) \,\big|\, S_0 = x\big].$$

Then it easy to see that,

$$f(x) = \frac{1}{4} \sum_{y \sim x} f(y), \qquad \text{for all } x \in A,$$

where we write $x \sim y$ if x and y are adjacent on the lattice $\mathbb{Z}^2$. This means that the value $f(x)$ is the mean over all the values at the adjacent vertices. A function with this property is called *discrete harmonic*, and we have solved the (easy) problem of finding the discrete harmonic function on $A \cup \partial A$ with given boundary values φ on ∂A. A more challenging problem, which we solve in Chapter 3, is the corresponding continuous problem, called the *Dirichlet problem.*

For its formulation, fix a connected open set $U \subset \mathbb{R}^2$ with nice boundary, and let $\varphi \colon \partial U \to \mathbb{R}$ be continuous. The harmonic functions $f \colon U \to \mathbb{R}$ on the domain U are characterised by the differential equation

$$\Delta f(x) := \sum_{j=1}^{2} \frac{\partial^2 f}{\partial x_j^2}(x) = 0, \quad \text{for all } x \in U.$$

The Dirichlet problem is to find, for a given domain U and boundary data φ, a continuous function $f \colon U \cup \partial U \to \mathbb{R}$, which is harmonic on U and agrees with φ on ∂U. In Theorem 3.12 we show that the unique solution of this problem is given as

$$f(x) = \mathbb{E}\big[\varphi(B(T)) \,\big|\, B(0) = x\big], \text{ for } x \in \overline{U},$$

FIGURE 2. Brownian motion and the Dirichlet problem

where $\{B(t) : t \geq 0\}$ is a Brownian motion and $T = \inf\{t \geq 0 : B(t) \notin U\}$ is the first exit time from U. We shall exploit this result, for example, to show exactly in which dimensions a particle following a Brownian motion drifts to infinity, see Theorem 3.19.

In *Chapter 4* we provide one of the major tools in our study of Brownian motion, the concept of Hausdorff dimension, and show how it can be applied in the context of Brownian motion. Indeed, when describing the sample paths of a Brownian motion one frequently encounters questions of the size of a given set: How big is the set of all points visited by a Brownian motion in the plane? How big is the set of double-points of a planar Brownian motion? How big is the set of times where Brownian motion visits a given set, say a point?

For an example, let $\{B(t) : t \geq 0\}$ be Brownian motion on the real line and look at

$$\mathsf{Zero} = \big\{t \geq 0 : B(t) = 0\big\},$$

the set of its zeros. Although $t \mapsto B(t)$ is a continuous function, Zero is an infinite set. This set is *big*, as it is an uncountable set without isolated points. However, it is also *small* in the sense that its Lebesgue measure, denoted $\mathcal{L}$, is zero. Indeed, we have by Fubini's theorem:

$$\mathbb{E}\big[\mathcal{L}(\mathsf{Zero})\big] = \mathbb{E}\int_0^\infty 1_{\mathsf{Zero}}(s)\,ds = \int_0^\infty \mathbb{P}\{s \in \mathsf{Zero}\}\,ds = \int_0^\infty \mathbb{P}\{B(s) = 0\}\,ds = 0.$$

Zero is a fractal set and we show in Theorem 4.24 that its Hausdorff dimension is $1/2$.

In *Chapter 5* we explore the relationship of random walk and Brownian motion. There are two natural ways to relate random walks directly to Brownian motion: Assume $d = 1$ for the moment and let X be an arbitrary random variable, for simplicity with $\mathbb{E}X = 0$ and $\mathrm{Var}\, X = 1$.

- *Random walks can be embedded into Brownian motion.*
  The idea is that, given any centred distribution with finite variance, one can define a sequence $T_1 < T_2 < \cdots$ of (stopping) times for Brownian motion of controllable size, such that $\{S_n : n \geq 1\}$ given by $S_n = B(T_n)$ is a random walk with increments distributed like X. We say that the random walk is *embedded* in Brownian motion.

- *Random walk paths converge in distribution to Brownian motion paths.*
  The main result is Donsker's invariance principle, which states that, for the random walk $\{S_k : k \in \mathbb{N}\}$ with increments distributed like X, the law of the random curve obtained by connecting the points $S_k/\sqrt{n}$ in order $k = 1, 2, \ldots, n$ linearly in $1/n$ time units converges in law to Brownian motion.

These two principles allow us to answer a lot of questions about random walks by looking at Brownian motion instead. Why can this be advantageous? First, in many cases the fact that Brownian motion is a continuous time process is an advantage over discrete time random walks. For example, as we discuss in the next paragraph, Brownian motion has scaling invariance properties, which can be a powerful tool in the study of its path properties. Second, even in cases where the discrete, combinatorial structures of a simple random walk model are the right tool in the proof of a statement, the translation into a Brownian motion setting frequently helps extending the result from a specific random walk, e.g. the simple random walk on the integers, where X_i takes values ± 1, to a wider range of random walks. In Chapter 5 we give several examples how results about Brownian motion can be exploited for random walks.

In *Chapter 6* we look again at Brownian motion in dimension $d = 1$. In this case the occupation measure μ_t of the Brownian motion, defined by

$$\mu_t(A) = \int_0^t \mathbb{1}\{B(s) \in A\}\, ds \qquad \text{for } A \subset \mathbb{R} \text{ Borel},$$

has a density. To see this use first Fatou's lemma and then Fubini's theorem,

$$\begin{aligned}\mathbb{E} \int \liminf_{r \downarrow 0} \frac{\mu_t(\mathcal{B}(x,r))}{\mathcal{L}(\mathcal{B}(x,r))}\, d\mu_t(x) &\leq \liminf_{r \downarrow 0} \frac{1}{2r}\, \mathbb{E} \int \mu_t(\mathcal{B}(x,r))\, d\mu_t(x) \\ &= \liminf_{r \downarrow 0} \frac{1}{2r} \int_0^t \int_0^t \mathbb{P}\big\{|B(s_1) - B(s_2)| \leq r\big\}\, ds_1\, ds_2\,.\end{aligned}$$

Using that the density of a standard normal random variable X is bounded by one, we get

$$\mathbb{P}\big\{|B(s_1) - B(s_2)| \leq r\big\} = \mathbb{P}\big\{|X| \leq \tfrac{r}{\sqrt{|s_1 - s_2|}}\big\} \leq \tfrac{r}{\sqrt{|s_1 - s_2|}},$$

and this implies that

$$\liminf_{r \downarrow 0} \frac{1}{2r} \int_0^t \int_0^t \mathbb{P}\big\{|B(s_1) - B(s_2)| \leq r\big\}\, ds_1\, ds_2 \leq \frac{1}{2} \int_0^t \int_0^t \frac{ds_1\, ds_2}{\sqrt{|s_1 - s_2|}} < \infty.$$

Hence

$$\liminf_{r \downarrow 0} \frac{\mu_t(\mathcal{B}(x,r))}{\mathcal{L}(\mathcal{B}(x,r))} < \infty \qquad \text{for } \mu_t\text{-almost every } x.$$

By the Radon-Nikodym this implies that a density $\{L^a(t) \colon a \in \mathbb{R}\}$ exists, which we call the *Brownian local time*. We shall construct this density by probabilistic means, show that it is jointly continuous in a and t, and characterise it as a stochastic process.

One of the most important invariance properties of Brownian motion is *conformal invariance*, which we discuss in *Chapter 7*. To make this plausible think of an angle-preserving linear mapping $L\colon \mathbb{R}^d \to \mathbb{R}^d$, like a rotation followed by multiplication by a. Take a random walk started in zero with increments of mean zero and covariance matrix the identity, and look at its image under L. This image is again a random walk and its increments are distributed like LX. Appropriately rescaled as in Donsker's invariance principle, both random walks converge, the first to a Brownian motion, the second to a process satisfying our conditions (1)–(4), but with a slightly different covariance matrix. This process can be identified as a time-changed Brownian motion

$$\{B(a^2t) : t \geq 0\}.$$

This easy observation has a deeper, local counterpart for planar Brownian motion: Suppose that $\phi\colon U \to V$ is a conformal mapping of a simply connected domain $U \subset \mathbb{R}^2$ onto a domain $V \subset \mathbb{R}^2$. Conformal mappings are locally angle-preserving and the Riemann mapping theorem of complex analysis tells us that a lot of these mappings exist.

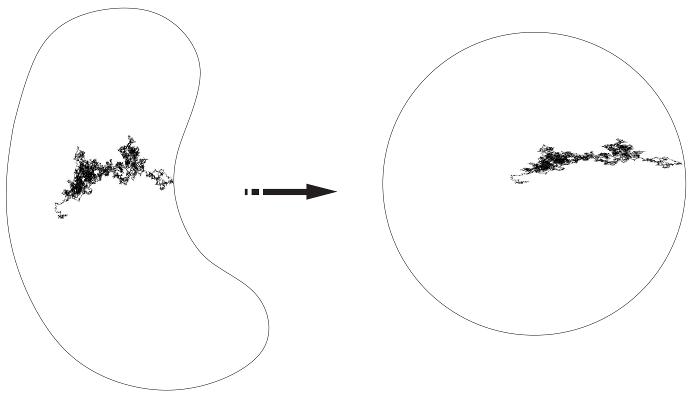

FIGURE 3. A conformal mapping of Brownian paths

Suppose that $\{B(t) : t \geq 0\}$ is a standard Brownian motion started in some point $x \in U$ and $\tau = \inf\{t > 0\colon B(t) \notin U\}$ is the first exit time of the path from the domain U. Then it turns out that the image process $\{\phi(B(t)) : 0 \leq t \leq \tau\}$ is a *time-changed* Brownian motion in the domain V, stopped when it leaves V. In order to prove this we have to develop a little bit of the theory of stochastic integration with respect to a Brownian motion, and we shall give a lot of further applications of this tool in Chapter 7.

In *Chapter 8* we develop the potential theory of Brownian motion. The problem which is the motivation behind this is, given a compact set $A \subset \mathbb{R}^d$, to find the probability that a Brownian motion $\{B(t)\colon t \geq 0\}$ hits the set A, i.e. that there exists $t > 0$ with $B(t) \in A$. This problem will be answered in the best possible way by Theorem 8.23, which is modern extension of a classical result of Kakutani: The hitting probability can be approximated by the capacity of A with respect to the Martin kernel up to a factor of two.

With a wide range of tools at our hand, in *Chapter 9* we study the self-intersections of Brownian motion: For example, a point $x \in \mathbb{R}^d$ is called a double point of $\{B(t)\colon t \geq 0\}$ if there exist times $0 < t_1 < t_2$ such that $B(t_1) = B(t_2) = x$. In which dimensions does Brownian motion have double points? How big is the set of double points? We show that, almost surely,

- in dimensions $d \geq 4$ no double points exist,
- in dimension $d = 3$ (and $d = 1$) double points exist and the set of double points has Hausdorff dimension one,
- in dimension $d = 2$ double points exists and the set of double points has Hausdorff dimension two.

In dimension $d = 2$ we find a surprisingly complex situation: While every point $x \in \mathbb{R}^2$ is almost surely not visited by a Brownian motion, there exist (random) points in the plane, which are visited infinitely often, even uncountably often. This result, Theorem 9.24, is one of the highlights of this book.

Chapter 10 deals with exceptional points for Brownian motion and Hausdorff dimension spectra of families of exceptional points. To explain an example, we look at a Brownian motion in the plane run for one time unit, which is a continuous curve $\{B(t) : t \in [0, 1]\}$. If $x = B(t) \in \mathbb{R}^2$, for some $0 < t < 1$, is a point on the curve one can use polar coordinates centred in x to define for every time interval $(t + \varepsilon, 1)$ the number of windings the curve performs around x in this time interval, with counterclockwise windings having a positive and clockwise windings having a negative sign. Denoting this number by $\theta(\varepsilon)$, we obtain in Chapter 7 that, almost surely,

$$\limsup_{\varepsilon \downarrow 0} \theta(\varepsilon) = \infty \text{ and } \liminf_{\varepsilon \downarrow 0} \theta(\varepsilon) = -\infty.$$

In other words, for any point on the curve, almost surely, the Brownian motion performs an infinite number of full windings in both directions.

Still, there exist random points on the curve, which are exceptional in the sense that Brownian motion performs no windings around them at all. This follows from an easy geometric argument: Take a point in $\mathbb{R}^2$ with coordinates (x_1, x_2) such that

$$x_1 = \min\{x \colon (x, x_2) \in B[0, 1]\},$$

i.e. a point which is the leftmost on the intersection of the Brownian curve and the line $\{(z, y) \colon z \in \mathbb{R}\}$, for some $x_2 \in \mathbb{R}$. Then Brownian motion does not perform any full windings around (x_1, x_2), as this would necessarily imply that it crosses the half-line $\{(x, x_2) \colon x < x_2\}$, contradicting the minimality of x_1.

One can ask for a more extreme deviation from typical behaviour: A point $x = B(t)$ is an α-cone point if the Brownian curve is contained in an open cone with tip in $x = (x_1, x_2)$, central axis $\{(x_1, x) \colon x > x_2\}$ and opening angle α. Note that the points described in the previous paragraph are 2π-cone points in this sense. We show that α-cone points exist exactly if $\alpha \in [\pi, 2\pi]$, and prove that for every such α, almost surely,

$$\dim \big\{x \in \mathbb{R}^2 \colon x \text{ is an } \alpha\text{-cone point } \big\} = 2 - \frac{2\pi}{\alpha}.$$

This is an example of a Hausdorff dimension spectrum, a topic which has been at the centre of some research activity at the beginning of the current millennium.

CHAPTER 1

Definition and first properties of Brownian motion

In this chapter we focus on one-dimensional, or linear, Brownian motion. We start with Paul Lévy's construction of Brownian motion and discuss two fundamental sample path properties, continuity and differentiability.

1. Paul Lévy's construction of Brownian motion

1.1. Definition of Brownian motion. Brownian motion is closely linked to the normal distribution. Recall that a random variable X is normally distributed with mean μ and variance σ^2 if

$$\mathbb{P}\{X > x\} = \frac{1}{\sqrt{2\pi\sigma^2}} \int_x^\infty e^{-\frac{(u-\mu)^2}{2\sigma^2}}\, du, \qquad \text{for all } x \in \mathbb{R}.$$

DEFINITION 1.1. *A real-valued stochastic process $\{B(t) : t \geq 0\}$ is called a* **(linear) Brownian motion** *with start in $x \in \mathbb{R}$ if the following holds:*

- $B(0) = x$,
- *the process has* **independent increments**, *i.e. for all times $0 \leq t_1 \leq t_2 \leq \ldots \leq t_n$ the increments $B(t_n) - B(t_{n-1})$, $B(t_{n-1}) - B(t_{n-2})$, ..., $B(t_2) - B(t_1)$ are independent random variables,*
- *for all $t \geq 0$ and $h > 0$, the increments $B(t+h) - B(t)$ are normally distributed with expectation zero and variance h,*
- *almost surely, the function $t \mapsto B(t)$ is continuous.*

We say that $\{B(t) : t \geq 0\}$ is a **standard Brownian motion** *if $x = 0$.* ◇

Let us step back and look at some technical points. We have defined Brownian motion as a *stochastic process* $\{B(t) : t \geq 0\}$ which is just a family of (uncountably many) random variables $\omega \mapsto B(t, \omega)$ defined on a single probability space $(\Omega, \mathcal{A}, \mathbb{P})$. At the same time, a stochastic process can also be interpreted as a *random function* with the sample functions defined by $t \mapsto B(t, \omega)$. The *sample path properties* of a stochastic process are the properties of these random functions, and it is these properties we will be most interested in in this book.

REMARK 1.2. When considering a stochastic process as a random function it is sometimes useful to assume that the mapping $(t, \omega) \mapsto B(t, \omega)$ is measurable on the product space $[0, \infty) \times \Omega$. We shall not need any such assumption before Chapter 7. ◇

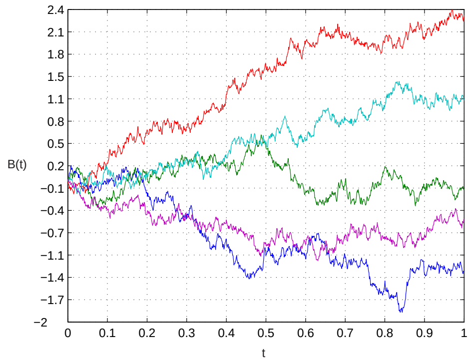

FIGURE 1. Graphs of five sampled Brownian motions

By the **marginal distributions** of a stochastic process $\{B(t) : t \geq 0\}$ we mean the laws of all the finite dimensional random vectors

$$\big(B(t_1), B(t_2), \ldots, B(t_n)\big), \text{ for all } 0 \leq t_1 \leq t_2 \leq \ldots \leq t_n.$$

To describe these joint laws it suffices to describe the joint law of $B(0)$ and all the increments

$$\big(B(t_1) - B(0), B(t_2) - B(t_1), \ldots, B(t_n) - B(t_{n-1})\big), \text{ for all } 0 \leq t_1 \leq t_2 \leq \ldots \leq t_n.$$

This is what we have done in the first three items of the definition, which specify the marginal distributions of Brownian motion. However, the last item, almost sure continuity, is also crucial, and this is information which goes beyond the marginal distributions of the process in the sense above, technically because the set $\{\omega \in \Omega : t \mapsto B(t, \omega) \text{ continuous}\}$ is in general not in the σ-algebra generated by the random vectors $(B(t_1), B(t_2), \ldots, B(t_n))$, $n \in \mathbb{N}$.

EXAMPLE 1.3. Suppose that B is a Brownian motion and U is an independent random variable, which is uniformly distributed on $[0, 1]$. Then the process $\{\tilde{B}(t) : t \geq 0\}$ defined by

$$\tilde{B}(t) = \begin{cases} B(t) & \text{if } t \neq U, \\ 0 & \text{if } t = U, \end{cases}$$

has the same marginal distributions as a Brownian motion, but is discontinuous if $B(U) \neq 0$, i.e. with probability one, and hence this process is not a Brownian motion. ⋄

We see that, if we are interested in the sample path properties of a stochastic process, we may need to specify more than just its marginal distributions. Suppose $\mathfrak{X}$ is a property a function might or might not have, like continuity, differentiability, etc. We say that a process $\{X(t) : t \geq 0\}$ **has property $\mathfrak{X}$ almost surely** if there exists $A \in \mathcal{A}$ such that $\mathbb{P}(A) = 1$ and $A \subset \big\{\omega \in \Omega : t \mapsto X(t, \omega) \text{ has property } \mathfrak{X}\big\}$. Note that the set on the right need not lie in $\mathcal{A}$.

1.2. Paul Lévy's construction of Brownian motion. It is a nontrivial issue whether the conditions imposed on the marginal distributions in the definition of Brownian motion allow the process to have continuous sample paths, or whether there is a contradiction. In this section we show that there is no contradiction and, fortunately, Brownian motion exists.

THEOREM 1.4 (Wiener 1923). *Standard Brownian motion exists.*

We construct Brownian motion as a uniform limit of continuous functions, to ensure that it automatically has continuous paths. Recall that we need only construct a *standard* Brownian motion $\{B(t) : t \geq 0\}$, as $X(t) = x + B(t)$ is a Brownian motion with starting point x. The proof exploits properties of Gaussian random vectors, which are the higher-dimensional analogue of the normal distribution.

DEFINITION 1.5. *A random vector $(X_1, \ldots, X_n)$ is called a* **Gaussian random vector** *if there exists an $n \times m$ matrix A, and an n-dimensional vector b such that $X^{\mathrm{T}} = AY + b$, where Y is an m-dimensional vector with independent standard normal entries.* ◇

Basic facts about Gaussian random variables are collected in Appendix II.2. The following exercise is an easy warm-up to the proof of Wiener's theorem.

Proof of Wiener's theorem. We first construct Brownian motion on the interval $[0, 1]$ as a random element on the space $\mathcal{C}[0, 1]$ of continuous functions on $[0, 1]$. The idea is to construct the right joint distribution of Brownian motion step by step on the finite sets

$$\mathcal{D}_n = \left\{ \tfrac{k}{2^n} \, : \, 0 \leq k \leq 2^n \right\}$$

of dyadic points. We then interpolate the values on $\mathcal{D}_n$ linearly and check that the uniform limit of these continuous functions exists and is a Brownian motion.

To do this let $\mathcal{D} = \bigcup_{n=0}^{\infty} \mathcal{D}_n$ and let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space on which a collection $\{Z_t \, : \, t \in \mathcal{D}\}$ of independent, standard normally distributed random variables can be defined. Let $B(0) := 0$ and $B(1) := Z_1$. For each $n \in \mathbb{N}$ we define the random variables $B(d)$, $d \in \mathcal{D}_n$ such that

(1) for all $r < s < t$ in $\mathcal{D}_n$ the random variable $B(t) - B(s)$ is normally distributed with mean zero and variance $t - s$, and is independent of $B(s) - B(r)$,
(2) the vectors $(B(d) \, : \, d \in \mathcal{D}_n)$ and $(Z_t \, : \, t \in \mathcal{D} \setminus \mathcal{D}_n)$ are independent.

Note that we have already done this for $\mathcal{D}_0 = \{0, 1\}$. Proceeding inductively we may assume that we have succeeded in doing it for some $n - 1$. We then define $B(d)$ for $d \in \mathcal{D}_n \setminus \mathcal{D}_{n-1}$ by

$$B(d) = \frac{B(d - 2^{-n}) + B(d + 2^{-n})}{2} + \frac{Z_d}{2^{(n+1)/2}} \, .$$

Note that the first summand is the linear interpolation of the values of B at the neighbouring points of d in $\mathcal{D}_{n-1}$. Therefore $B(d)$ is independent of $(Z_t \, : \, t \in \mathcal{D} \setminus \mathcal{D}_n)$ and the second property is fulfilled.

Moreover, as $\frac{1}{2}[B(d + 2^{-n}) - B(d - 2^{-n})]$ depends only on $(Z_t \, : \, t \in \mathcal{D}_{n-1})$, it is independent of $Z_d / 2^{(n+1)/2}$. Both terms are normally distributed with mean zero and variance $2^{-(n+1)}$. Hence their sum $B(d) - B(d - 2^{-n})$ and their difference $B(d + 2^{-n}) - B(d)$ are independent and normally distributed with mean zero and variance 2^{-n} by Corollary II.3.4.

Indeed, all increments $B(d) - B(d - 2^{-n})$, for $d \in \mathcal{D}_n \setminus \{0\}$, are independent. To see this it suffices to show that they are pairwise independent, as the vector of these increments is Gaussian. We have seen in the previous paragraph that pairs $B(d) - B(d - 2^{-n})$, $B(d + 2^{-n}) - B(d)$ with

$d \in \mathcal{D}_n \setminus \mathcal{D}_{n-1}$ are independent. The other possibility is that the increments are over intervals separated by some $d \in \mathcal{D}_{n-1}$. Choose $d \in \mathcal{D}_j$ with this property and minimal j, so that the two intervals are contained in $[d-2^{-j}, d]$, respectively $[d, d+2^{-j}]$. By induction the increments over these two intervals of length 2^{-j} are independent, and the increments over the intervals of length 2^{-n} are constructed from the independent increments $B(d) - B(d-2^{-j})$, respectively $B(d+2^{-j}) - B(d)$, using a disjoint set of variables $(Z_t : t \in \mathcal{D}_n)$. Hence they are independent and this implies the first property, and completes the induction step.

Having thus chosen the values of the process on all dyadic points, we interpolate between them. Formally, define

$$F_0(t) = \begin{cases} Z_1 & \text{for } t = 1 \\ 0 & \text{for } t = 0 \\ \text{linear} & \text{in between.} \end{cases}$$

and, for each $n \geq 0$,

$$F_n(t) = \begin{cases} 2^{-(n+1)/2} Z_t & \text{for } t \in \mathcal{D}_n \setminus \mathcal{D}_{n-1} \\ 0 & \text{for } t \in \mathcal{D}_{n-1} \\ \text{linear between} & \text{consecutive points in } \mathcal{D}_n. \end{cases}$$

These functions are continuous on $[0, 1]$ and for all n and $d \in \mathcal{D}_n$

$$B(d) = \sum_{i=0}^{n} F_i(d) = \sum_{i=0}^{\infty} F_i(d). \tag{1.1}$$

This can be seen by induction. It holds for $n = 0$. Suppose that it holds for $n - 1$. Let $d \in \mathcal{D}_n \setminus \mathcal{D}_{n-1}$. Since for $0 \leq i \leq n-1$ the function F_i is linear on $[d - 2^{-n}, d + 2^{-n}]$, we get

$$\sum_{i=0}^{n-1} F_i(d) = \sum_{i=1}^{n-1} \frac{F_i(d-2^{-n}) + F_i(d+2^{-n})}{2} = \frac{B(d-2^{-n}) + B(d+2^{-n})}{2}.$$

Since $F_n(d) = 2^{-(n+1)/2} Z_d$, this gives (1.1).

On the other hand, we have, by definition of Z_d and by Lemma II.3.1, for $c > 0$ and large n,

$$\mathbb{P}\{|Z_d| \geq c\sqrt{n}\} \leq \exp\left(\frac{-c^2 n}{2}\right),$$

so that the series

$$\begin{aligned} \sum_{n=0}^{\infty} \mathbb{P}\{\text{ there exists } d \in \mathcal{D}_n \text{ with } |Z_d| \geq c\sqrt{n}\} &\leq \sum_{n=0}^{\infty} \sum_{d \in \mathcal{D}_n} \mathbb{P}\{|Z_d| \geq c\sqrt{n}\} \\ &\leq \sum_{n=0}^{\infty} (2^n + 1) \exp\left(\frac{-c^2 n}{2}\right), \end{aligned}$$

converges as soon as $c > \sqrt{2 \log 2}$. Fix such a c. By the Borel-Cantelli lemma there exists a random (but almost surely finite) N such that for all $n \geq N$ and $d \in \mathcal{D}_n$ we have $|Z_d| < c\sqrt{n}$. Hence, for all $n \geq N$,

$$\|F_n\|_\infty < c\sqrt{n} 2^{-n/2}. \tag{1.2}$$

This upper bound implies that, almost surely, the series

$$B(t) = \sum_{n=0}^{\infty} F_n(t)$$

is uniformly convergent on $[0,1]$. We denote the continuous limit by $\{B(t) : t \in [0,1]\}$.

It remains to check that the increments of this process have the right marginal distributions. This follows directly from the properties of B on the dense set $\mathcal{D} \subset [0,1]$ and the continuity of the paths. Indeed, suppose that $t_1 < t_2 < \cdots < t_n$ are in $[0,1]$. We find $t_{1,k} \le t_{2,k} \le \cdots \le t_{n,k}$ in $\mathcal{D}$ with $\lim_{k\uparrow\infty} t_{i,k} = t_i$ and infer from the continuity of B that, for $1 \le i \le n-1$,

$$B(t_{i+1}) - B(t_i) = \lim_{k\uparrow\infty} B(t_{i+1,k}) - B(t_{i,k})\,.$$

As $\lim_{k\uparrow\infty} \mathbb{E}[B(t_{i+1,k}) - B(t_{i,k})] = 0$ and

$$\begin{aligned}\lim_{k\uparrow\infty} \mathrm{Cov}\,\big(B(t_{i+1,k}) - B(t_{i,k}), B(t_{j+1,k}) - B(t_{j,k})\big)\big)& \\ = \lim_{k\uparrow\infty} \mathbb{1}_{\{i=j\}}\,\big(t_{i+1,k} - t_{i,k}\big) &= \mathbb{1}_{\{i=j\}}\,\big(t_{i+1} - t_i\big)\,,\end{aligned}$$

the increments $B(t_{i+1}) - B(t_i)$ are, by Proposition II.3.7, independent Gaussian random variables with mean 0 and variance $t_{i+1} - t_i$, as required.

We have thus constructed a continuous process $B\colon [0,1] \to \mathbb{R}$ with the same marginal distributions as Brownian motion. Take a sequence $B_1, B_2, \ldots$ of independent $\mathcal{C}[0,1]$-valued random variables with the distribution of this process, and define $\{B(t) : t \ge 0\}$ by gluing together the parts, more precisely by

$$B(t) = B_{\lfloor t\rfloor}(t - \lfloor t\rfloor) + \sum_{i=0}^{\lfloor t\rfloor - 1} B_i(1)\,,\ \text{for all } t \ge 0\,.$$

This defines a continuous random function $B\colon [0,\infty) \to \mathbb{R}$ and one can see easily from what we have shown so far that the requirements of a standard Brownian motion are fulfilled. ■

Remark 1.6. A stochastic process $\{Y(t) : t \ge 0\}$ is called a **Gaussian process**, if for all $t_1 < t_2 < \ldots < t_n$ the vector $(Y(t_1), \ldots, Y(t_n))$ is a Gaussian random vector. It is shown in Exercise 1.2 that Brownian motion with start in $x \in \mathbb{R}$ is a Gaussian process. ◇

1.3. Simple invariance properties of Brownian motion. One of the themes of this book is that many natural sets that can be derived from the sample paths of Brownian motion are in some sense *random fractals*. An intuitive approach to fractals is that they are sets which have a nontrivial geometric structure at all scales.

A key role in this behaviour is played by the very simple *scaling invariance* property of Brownian motion, which we now formulate. It identifies a transformation on the space of functions, which changes the individual Brownian random functions but leaves their distribution unchanged.

LEMMA 1.7 (Scaling invariance). *Suppose* $\{B(t) : t \geq 0\}$ *is a standard Brownian motion and let* $a > 0$. *Then the process* $\{X(t) : t \geq 0\}$ *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 rescaling. It remains to observe that $X(t) - X(s) = \frac{1}{a}\big(B(a^2t) - B(a^2s)\big)$ is normally distributed with expectation 0 and variance $(1/a^2)(a^2t - a^2s) = t - s$. ■

REMARK 1.8. Scaling invariance has many useful consequences. As an example, let $a < 0 < b$, and look at $T(a,b) = \inf\{t \geq 0 : B(t) = a \text{ or } B(t) = b\}$, the first exit time of a one-dimensional standard Brownian motion from the interval $[a,b]$. Then, with $X(t) = \frac{1}{a}B(a^2t)$ we have

$$\mathbb{E}T(a,b) = a^2\,\mathbb{E}\inf\big\{t \geq 0 : X(t) = 1 \text{ or } X(t) = b/a\big\} = a^2\,\mathbb{E}T(b/a,1)\,,$$

which implies that $\mathbb{E}T(-b,b)$ is a constant multiple of b^2. Also

$$\mathbb{P}\big\{\{B(t)\colon t \geq 0\} \text{ exits } [a,b] \text{ at } a\big\} = \mathbb{P}\big\{\{X(t)\colon t \geq 0\} \text{ exits } [1,b/a] \text{ at } 1\big\}$$

is only a function of the ratio b/a. The scaling invariance property will be used extensively in all the following chapters, and we shall often use the phrase that a fact holds 'by Brownian scaling' to indicate this. ◇

We shall discuss a very powerful extension of the scaling invariance property, the *conformal invariance property*, in Chapter 7 of the book. A further useful invariance property of Brownian motion, invariance under time-inversion, can be identified easily. As above, the transformation on the space of functions changes the individual Brownian random functions without changing the distribution.

THEOREM 1.9 (Time inversion). *Suppose* $\{B(t) : t \geq 0\}$ *is a standard Brownian motion. Then the process* $\{X(t) : t \geq 0\}$ *defined by*

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

is also a standard Brownian motion.

Proof. Recall that the finite dimensional marginals $(B(t_1), \dots, B(t_n))$ of Brownian motion are Gaussian random vectors and are therefore characterised by $\mathbb{E}[B(t_i)] = 0$ and $\operatorname{Cov}(B(t_i), B(t_j)) = t_i$ for $0 \leq t_i \leq t_j$.

Obviously, $\{X(t) : t \geq 0\}$ is also a Gaussian process and the Gaussian random vectors $(X(t_1), \dots, X(t_n))$ have expectation 0. The covariances, for $t > 0$, $h \geq 0$, are given by

$$\operatorname{Cov}(X(t+h), X(t)) = (t+h)t\operatorname{Cov}(B(1/(t+h)), B(1/t)) \quad = t(t+h)\,\frac{1}{t+h} = t\,.$$

Hence the law of all the finite dimensional marginals

$$\big(X(t_1), X(t_2), \dots, X(t_n)\big), \text{ for } 0 \leq t_1 \leq \dots \leq t_n,$$

are the same as for Brownian motion. The paths of $t \mapsto X(t)$ are clearly continuous for all $t > 0$ and in $t = 0$ we use the following two facts: First, the distribution of X on the rationals $\mathbb{Q}$ is the same as for a Brownian motion, hence

$$\lim_{\substack{t \downarrow 0 \\ t \in \mathbb{Q}}} X(t) = 0 \text{ almost surely.}$$

And second, X is almost surely continuous on $(0, \infty)$, so that

$$0 = \lim_{\substack{t \downarrow 0 \\ t \in \mathbb{Q}}} X(t) = \lim_{t \downarrow 0} X(t) \text{ almost surely.}$$

Hence $\{X(t) : t \geq 0\}$ has almost surely continuous paths, and is a Brownian motion. ■

Remark 1.10. The symmetry inherent in the time inversion property becomes more apparent if one considers the *Ornstein-Uhlenbeck diffusion* $\{X(t) : t \in \mathbb{R}\}$, which is given by

$$X(t) = e^{-t} B(e^{2t}) \text{ for all } t \in \mathbb{R}.$$

This is a Markov process (this will be explained properly in Chapter 2.3), such that $X(t)$ is standard normally distributed for all t. It is a diffusion with a drift towards the origin proportional to the distance from the origin. Unlike Brownian motion, the Ornstein-Uhlenbeck diffusion is time reversible: The time inversion formula gives that $\{X(t) : t \geq 0\}$ and $\{X(-t) : t \geq 0\}$ have the same law. For t near $-\infty$, $X(t)$ relates to the Brownian motion near time 0, and for t near ∞, $X(t)$ relates to the Brownian motion near ∞. ◇

Time inversion is a useful tool to relate the properties of Brownian motion in a neighbourhood of time $t = 0$ to properties at infinity. To illustrate the use of time-inversion we exploit Theorem 1.9 to get an interesting statement about the long-term behaviour from a trivial statement at the origin.

Corollary 1.11 (Law of large numbers). *Almost surely,* $\displaystyle\lim_{t \to \infty} \frac{B(t)}{t} = 0.$

Proof. Let $\{X(t) : t \geq 0\}$ be as defined in Theorem 1.9. Using this theorem, we see that $\lim_{t \to \infty} B(t)/t = \lim_{t \to \infty} X(1/t) = X(0) = 0$. ■

In the next two chapters we discuss the two basic analytic properties of Brownian motion as a random function, its *continuity* and *differentiability* properties.

2. Continuity properties of Brownian motion

The definition of Brownian motion already requires that the sample functions are continuous almost surely. This implies that on the interval $[0, 1]$ (or any other compact interval) the sample functions are uniformly continuous, i.e. there exists some (random) function φ with $\lim_{h \downarrow 0} \varphi(h) = 0$ called a **modulus of continuity** of the function $B : [0, 1] \to \mathbb{R}$, such that

$$\limsup_{h \downarrow 0} \sup_{0 \leq t \leq 1-h} \frac{|B(t+h) - B(t)|}{\varphi(h)} \leq 1. \tag{2.1}$$

Can we achieve such a bound with a deterministic function φ, i.e. is there a nonrandom modulus of continuity for the Brownian motion? The answer is yes, as the following theorem shows.

THEOREM 1.12. *There exists a constant $C > 0$ such that, almost surely, for every sufficiently small $h > 0$ and all $0 \le t \le 1 - h$,*

$$\big|B(t+h) - B(t)\big| \le C\sqrt{h\log(1/h)}.$$

Proof. This follows quite elegantly from Lévy's construction of Brownian motion. Recall the notation introduced there and that we have represented Brownian motion as a series

$$B(t) = \sum_{n=0}^{\infty} F_n(t)\,,$$

where each F_n is a piecewise linear function. The derivative of F_n exists almost everywhere, and by definition and (1.2), for any $c > \sqrt{2\log 2}$ there exists a (random) $N \in \mathbb{N}$ such that, for all $n > N$,

$$\|F_n'\|_\infty \le \frac{2\|F_n\|_\infty}{2^{-n}} \le 2c\sqrt{n}2^{n/2}\,.$$

Now for each $t, t+h \in [0,1]$, using the mean-value theorem,

$$|B(t+h) - B(t)| \le \sum_{n=0}^{\infty} |F_n(t+h) - F_n(t)| \le \sum_{n=0}^{\ell} h\|F_n'\|_\infty + \sum_{n=\ell+1}^{\infty} 2\|F_n\|_\infty\,.$$

Hence, using (1.2) again, we get for all $\ell > N$, that this is bounded by

$$h\sum_{n=0}^{N} \|F_n'\|_\infty + 2ch\sum_{n=N}^{\ell} \sqrt{n}2^{n/2} + 2c\sum_{n=\ell+1}^{\infty} \sqrt{n}2^{-n/2}.$$

We now suppose that h is (again random and) small enough that the first summand is smaller than $\sqrt{h\log(1/h)}$ and that ℓ defined by $2^{-\ell} < h \le 2^{-\ell+1}$ exceeds N. For this choice of ℓ the second and third summands are also bounded by constant multiples of $\sqrt{h\log(1/h)}$ as both sums are dominated by their largest element. Hence we get (2.1) with a deterministic function $\varphi(h) = C\sqrt{h\log(1/h)}$. ■

This upper bound is pretty close to the optimal result. The following lower bound confirms that the only missing bit is the precise value of the constant.

THEOREM 1.13. *For every constant $c < \sqrt{2}$, almost surely, for every $\varepsilon > 0$ there exist $0 < h < \varepsilon$ and $t \in [0, 1-h]$ with*

$$\big|B(t+h) - B(t)\big| \ge c\sqrt{h\log(1/h)}.$$

Proof. Let $c < \sqrt{2}$ and define, for integers $k, n \ge 0$, the events

$$A_{k,n} = \Big\{B((k+1)e^{-n}) - B(ke^{-n}) > c\sqrt{n}e^{-n/2}\Big\}.$$

Then, using Lemma II.3.1, for any $k \ge 0$,

$$\mathbb{P}(A_{k,n}) = \mathbb{P}\{B(e^{-n}) > c\sqrt{n}e^{-n/2}\} = \mathbb{P}\{B(1) > c\sqrt{n}\} \ge \frac{c\sqrt{n}}{c^2n+1}e^{-c^2n/2}\,.$$

By our assumption on c, we have $e^n \mathbb{P}(A_{k,n}) \to \infty$ as $n \uparrow \infty$. Therefore, using $1 - x \le e^{-x}$ for all x,

$$\mathbb{P}\Big(\bigcap_{k=0}^{\lfloor e^n - 1 \rfloor} A_{k,n}^c \Big) = (1 - \mathbb{P}(A_{0,n}))^{e^n} \le \exp(-e^n \mathbb{P}(A_{0,n})) \to 0 .$$

By considering $h = e^{-n}$ one can now see that, for any $\varepsilon > 0$,

$$\mathbb{P}\Big\{ \text{ for all } h \in (0, \varepsilon) \text{ and } t \in [0, 1-h] \text{ we have } |B(t+h) - B(t)| \le c\sqrt{h \log(1/h)} \Big\} = 0 . \quad \blacksquare$$

One can determine the constant c in the best possible modulus of continuity $\varphi(h) = c\sqrt{h \log(1/h)}$ precisely. Indeed, our proof of the lower bound yields a value of $c = \sqrt{2}$, which turns out to be optimal. This striking result is due to Paul Lévy.

THEOREM* 1.14 (Lévy's modulus of continuity (1937)). *Almost surely,*

$$\limsup_{h \downarrow 0} \sup_{0 \le t \le 1-h} \frac{|B(t+h) - B(t)|}{\sqrt{2h \log(1/h)}} = 1 .$$

REMARK 1.15. We come back to the modulus of continuity of Brownian motion in Chapter 10, where we prove a substantial extension, the spectrum of fast points of Brownian motion. We will not use Theorem 1.14 in the sequel as Theorem 1.12 is sufficient to discuss all problems where an upper bound on the increase of a Brownian motion is needed. Hence the proof of Lévy's modulus of continuity may be skipped on first reading. $\diamond$

In the light of Theorem 1.13, we only need to prove the upper bound. We first look at increments over a class of intervals, which is chosen to be sparse, but big enough to approximate arbitrary intervals. More precisely, given natural numbers n, m, we let $\Lambda_n(m)$ be the collection of all intervals of the form

$$\big[(k-1+b)2^{-n+a}, (k+b)2^{-n+a}\big],$$

for $k \in \{1, \ldots, 2^n\}$, $a, b \in \{0, \frac{1}{m}, \ldots, \frac{m-1}{m}\}$. We further define $\Lambda(m) := \bigcup_n \Lambda_n(m)$.

LEMMA 1.16. *For any fixed m and $c > \sqrt{2}$, almost surely, there exists $n_0 \in \mathbb{N}$ such that, for any $n \ge n_0$,*

$$\big|B(t) - B(s)\big| \le c\sqrt{(t-s) \log \tfrac{1}{(t-s)}} \qquad \textit{for all } [s,t] \in \Lambda_m(n).$$

Proof. From the tail estimate for a standard normal random variable X, see Lemma II.3.1, we obtain

$$\begin{aligned}
\mathbb{P}\Big\{ \sup_{k \in \{1,\ldots,2^n\}} \sup_{a,b \in \{0, \frac{1}{m}, \ldots, \frac{m-1}{m}\}} & \Big| B\big((k-1+b)2^{-n+a}\big) - B\big((k+b)2^{-n+a}\big) \Big| > c\sqrt{2^{-n+a} \log(2^{n+a})} \Big\} \\
& \le 2^n (2m)\, \mathbb{P}\big\{ X > c\sqrt{\log(2^n)} \big\} \\
& \le \frac{2m}{c\sqrt{\log(2^n)}} \frac{1}{\sqrt{2\pi}} 2^{n(1-c^2/2)},
\end{aligned}$$

and as the right hand side is summable, the result follows from the Borel-Cantelli lemma. ∎

LEMMA 1.17. *Given* $\varepsilon > 0$ *there exists* $m \in \mathbb{N}$ *such that for every interval* $[s,t] \subset [0,1]$ *there exists an interval* $[s',t'] \in \Lambda(m)$ *with* $|t-t'| < \varepsilon\,(t-s)$ *and* $|s-s'| < \varepsilon\,(t-s)$.

Proof. Choose m large enough to ensure that $\frac{1}{m} < \varepsilon/4$ and $2^{1/m} < 1+\varepsilon/2$. Given an interval $[s,t] \subset [0,1]$, we first pick n such that $2^{-n} \leq t-s < 2^{-n+1}$, then $a \in \{0, \frac{1}{m}, \dots, \frac{m-1}{m}\}$ such that

$$2^{-n+a} \leq t-s < 2^{-n+a+\frac{1}{m}}.$$

Next, pick $k \in \{1, \dots, 2^n\}$ such that $(k-1)2^{-n+a} < s \leq k2^{-n+a}$, and $b \in \{0, \frac{1}{m}, \dots, \frac{m-1}{m}\}$ such that

$$(k-1+b)2^{-n+a} \leq s \leq (k-1+b+\tfrac{1}{m})2^{-n+a}.$$

Let $s' = (k-1+b)2^{-n+a}$, then

$$|s-s'| \leq \tfrac{1}{m}2^{-n+a} \leq \tfrac{\varepsilon}{4}\,2^{-n+1} \leq \tfrac{\varepsilon}{2}\,(t-s).$$

Choosing $t' = (k+b)2^{-n+a}$ ensures that $[s',t'] \in \Lambda_n(m)$ and, moreover,

$$\begin{aligned} |t-t'| &\leq |s-s'| + |(t-s)-(t'-s')| \leq \tfrac{\varepsilon}{2}\,(t-s) + \big(2^{-n+a+1/m} - 2^{-n+a}\big) \\ &\leq \tfrac{\varepsilon}{2}\,(t-s) + \tfrac{\varepsilon}{2}\,2^{-n+a} \leq \varepsilon\,(t-s), \end{aligned}$$

as required. ∎

Proof of Theorem 1.14. Given $c > \sqrt{2}$, pick $0 < \varepsilon < 1$ small enough to ensure that $\tilde{c} := c - \varepsilon > \sqrt{2}$ and $m \in \mathbb{N}$ as in Lemma 1.17. Using Lemma 1.16 we choose $n_0 \in \mathbb{N}$ large enough that, for all $n \geq n_0$ and all intervals $[s',t'] \in \Lambda_n(m)$, almost surely,

$$|B(t') - B(s')| \leq \tilde{c}\,\sqrt{(t'-s')\log\tfrac{1}{(t'-s')}}\,.$$

Now let $[s,t] \subset [0,1]$ be arbitrary, with $t-s < 2^{-n_0} \wedge \varepsilon$, and pick $[s',t'] \in \Lambda(m)$ with $|t-t'| < \varepsilon\,(t-s)$ and $|s-s'| < \varepsilon\,(t-s)$. Then, recalling Theorem 1.12, we obtain

$$\begin{aligned} \big|B(t)-B(s)\big| &\leq \big|B(t)-B(t')\big| + \big|B(t')-B(s')\big| + \big|B(s')-B(s)\big| \\ &\leq C\,\sqrt{|t-t'|\log\tfrac{1}{|t-t'|}} + \tilde{c}\,\sqrt{(t'-s')\log\tfrac{1}{t'-s'}} + C\,\sqrt{|s-s'|\log\tfrac{1}{|s-s'|}} \\ &\leq \big(4C\sqrt{\varepsilon} + \tilde{c}\,\sqrt{(1+2\varepsilon)(1+\log(1+2\varepsilon))}\big)\sqrt{(t-s)\log\tfrac{1}{t-s}}. \end{aligned}$$

By making $\varepsilon > 0$ small, the first factor on the right can be chosen arbitrarily close to c. This completes the proof of the upper bound, and hence of the theorem. ∎

REMARK 1.18. The limsup in Theorem 1.14 may be replaced by a limit, see Exercise 1.6. ◇

Definition 1.19. *A function* $f\colon [0,\infty) \to \mathbb{R}$ *is said to be* **locally α-Hölder continuous at** $x \geq 0$*, if there exists* $\varepsilon > 0$ *and* $c > 0$ *such that*

$$|f(x) - f(y)| \leq c\,|x - y|^\alpha, \qquad \text{for all } y \geq 0 \text{ with } |y - x| < \varepsilon.$$

We refer to $\alpha > 0$ *as the* **Hölder exponent** *and to* $c > 0$ *as the* **Hölder constant** . ◇

Clearly, α-Hölder continuity gets stronger, as the exponent α gets larger. The results of this chapter so far indicate that, for Brownian motion, the transition between paths which are α-Hölder continuous and paths which are not happens at $\alpha = 1/2$.

Corollary 1.20. *If* $\alpha < 1/2$*, then, almost surely, Brownian motion is everywhere locally* α*-Hölder continuous.*

Proof. Let $C > 0$ be as in Theorem 1.12. Applying this theorem to the Brownian motions $\{B(t) - B(k)\colon t \in [k, k+1]\}$, where k is a nonnegative integer, we see that, almost surely, for every k there exists $h(k) > 0$ such that for all $t \in [k, k+1)$ and $0 < h < (k + 1 - t) \wedge h(k)$,

$$\big|B(t+h) - B(t)\big| \leq C\sqrt{h\log(1/h)} \leq C\,h^\alpha\,.$$

Doing the same to the Brownian motions $\{\tilde{B}(t)\colon t \in [k, k{+}1]\}$ with $\tilde{B}(t) = B(k{+}1{-}t){-}B(k{+}1)$ gives the full result. ■

Remark 1.21. This result is optimal in the sense that, for $\alpha > 1/2$, almost surely, at every point, Brownian motion fails to be locally α-Hölder continuous, see Exercise 1.7. Points where Brownian motion is locally 1/2-Hölder continuous exist almost surely, but they are very rare. We come back to this issue when discussing 'slow points' of Brownian motion in Chapter 10. ◇

3. Nondifferentiability of Brownian motion

Having proved in the previous section that Brownian motion is somewhat *regular*, let us see why it is *erratic*. One manifestation is that the paths of Brownian motion have no intervals of monotonicity.

Proposition 1.22. *Almost surely, for all* $0 < a < b < \infty$*, Brownian motion is not monotone on the interval* $[a, b]$*.*

Proof. First fix an interval $[a, b]$. If $[a, b]$ is an interval of monotonicity, i.e. if $B(s) \leq B(t)$ for all $a \leq s \leq t \leq b$, then we pick numbers $a = a_1 \leq \ldots \leq a_{n+1} = b$ and divide $[a, b]$ into n sub-intervals $[a_i, a_{i+1}]$. Each increment $B(a_i) - B(a_{i+1})$ has to have the same sign. As the increments are independent, this has probability $2 \cdot 2^{-n}$, and taking $n \to \infty$ shows that the probability that $[a, b]$ is an interval of monotonicity must be 0. Taking a countable union gives that there is no interval of monotonicity with rational endpoints, but each monotone interval would have a nontrivial monotone rational sub-interval. ■

In order to discuss differentiability of Brownian motion we make use of the *time-inversion* trick, which allows us to relate differentiability at $t = 0$ to a long-term property. This property is a complementary result to the law of large numbers: Whereas Corollary 1.11 asserts that Brownian motion grows slower than linearly, the next proposition shows that the limsup growth of $B(t)$ is faster than $\sqrt{t}$.

PROPOSITION 1.23. *Almost surely,*

$$\limsup_{n\to\infty} \frac{B(n)}{\sqrt{n}} = +\infty, \quad \textit{and} \quad \liminf_{n\to\infty} \frac{B(n)}{\sqrt{n}} = -\infty. \tag{3.1}$$

For the proof of Proposition 1.23 we use the Hewitt-Savage 0-1 law for exchangeable events, which we briefly recall. Readers unfamiliar with the result are invited to give a proof as Exercise 1.8.

DEFINITION 1.24. *Let $X_1, X_2, \ldots$ be a sequence of random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and consider a set A of sequences such that*

$$\big\{X_1, X_2, \dots \in A\big\} \in \mathcal{F}.$$

The event $\{X_1, X_2, \dots \in A\}$ is called **exchangeable** *if*

$$\big\{X_1, X_2, \dots \in A\big\} \subset \big\{X_{\sigma_1}, X_{\sigma_2}, \dots \in A\big\}$$

for all finite permutations $\sigma \colon \mathbb{N} \to \mathbb{N}$. Here finite permutation *means that σ is a bijection with $\sigma_n = n$ for all sufficiently large n.* ◇

LEMMA 1.25 (Hewitt-Savage 0-1 law). *If A is an exchangeable event for an independent, identically distributed sequence, then $\mathbb{P}(A)$ is 0 or 1.*

Proof of Proposition 1.23. We clearly have, by Fatou's lemma,

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

By the scaling property, the expression in the lim sup equals $\mathbb{P}\{B(1) > c\}$, which is positive. Let $X_n = B(n) - B(n-1)$, and note that

$$\big\{B(n) > c\sqrt{n} \text{ infinitely often}\big\} = \Big\{\sum_{j=1}^{n} X_j > c\sqrt{n} \text{ infinitely often}\Big\}$$

is an exchangeable event. Hence the Hewitt-Savage 0-1 law gives that, with probability one, $B(n) > c\sqrt{n}$ infinitely often. Taking the intersection over all positive integers c gives the first part of the statement and the second part is proved analogously. ■

REMARK 1.26. It is natural to ask whether there exists a 'gauge' function $\varphi \colon [0,\infty) \to [0,\infty)$ such that $B(t)/\varphi(t)$ has a lim sup which is greater than 0 but less than ∞. An answer will be given by the law of the iterated logarithm in the first section of Chapter 5. ◇

For a function f, we define the **upper** and **lower right derivatives**

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

and

$$D_*f(t) \quad = \quad \liminf_{h\downarrow 0} \frac{f(t+h)-f(t)}{h}.$$

We now show that for any fixed time t, almost surely, Brownian motion is not differentiable at t. For this we use Proposition 1.23 and the invariance under time-inversion.

THEOREM 1.27. *Fix $t \geq 0$. Then, almost surely, Brownian motion is not differentiable at t. Moreover, $D^*B(t) = +\infty$ and $D_*B(t) = -\infty$.*

Proof. Given a standard Brownian motion B we construct a further Brownian motion X by time-inversion as in Theorem 1.9. Then

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

which is infinite by Proposition 1.23. Similarly, $D_*X(0) = -\infty$, showing that X is not differentiable at 0. Now let $t > 0$ be arbitrary and $\{B(t) : t \geq 0\}$ a Brownian motion. Then $X(s) = B(t+s) - B(t)$ defines a standard Brownian motion and differentiability of X at zero is equivalent to differentiability of B at t. ■

While the previous proof shows that every t is almost surely a point of nondifferentiability for the Brownian motion, this does *not* imply that almost surely *every* t is a point of nondifferentiability for the Brownian motion! The order of the quantifiers *for all* t and *almost surely* in results like Theorem 1.27 is of vital importance. Here the statement holds for all Brownian paths outside a set of probability zero, which may depend on t, and the union of all these sets of probability zero may not itself be a set of probability zero.

To illustrate this point, consider the following example: The argument in the proof of Theorem 1.27 also shows that the Brownian motion X crosses 0 for arbitrarily small values $s > 0$. Defining the level sets $Z(t) = \{s > 0 : X(s) = X(t)\}$, this shows that every t is almost surely an accumulation point from the right for $Z(t)$. But not every point $t \in [0,1]$ is an accumulation point from the right for $Z(t)$. For example the last zero of $\{X(t) : t \geq 0\}$ before time 1 is, by definition, never an accumulation point from the right for $Z(t) = Z(0)$. This example illustrates that there can be random *exceptional times* at which Brownian motion exhibits atypical behaviour. These times are so rare that any fixed (i.e. nonrandom) time is almost surely not of this kind.

REMARK 1.28. The behaviour of Brownian motion at a fixed time $t > 0$ reflects the behaviour at *typical times* in the following sense: Suppose $\mathfrak{X}$ is a measurable event (a set of paths) such that for all fixed $t \geq 0$,

$$\mathbb{P}\big\{s \mapsto B(t+s) - B(t) \text{ satisfies } \mathfrak{X}\big\} = 1.$$

Then, almost surely, the set of exceptional times

$$\big\{t : s \mapsto B(t+s) - B(t) \text{ does not satisfy } \mathfrak{X}\big\}$$

has Lebesgue measure 0. Indeed, write Θ_t for the operator that shifts paths by t, such that $(\Theta_t B)(s) = B(t+s) - B(t)$. Then, by Fubini's theorem,

$$\mathbb{E}\int_0^\infty \mathbb{1}\{t : s \mapsto (\Theta_t B)(s) \text{ does not satisfy } \mathfrak{X}\}\, dt = \int_0^\infty \mathbb{P}\{\Theta_t B \notin \mathfrak{X}\}\, dt = 0.$$

For example, the previous result shows that, almost surely, the path of a Brownian motion is not differentiable at Lebesgue-almost every time t. ◇

Remark 1.29. Exercise 1.9 shows that, almost surely, there exist times $t_*, t^* \in [0,1)$ with $D^*B(t^*) \le 0$ and $D_*B(t_*) \ge 0$. Hence the almost sure behaviour at a fixed point t, which is described in Theorem 1.27, does not hold at all points simultaneously. ◇

Theorem 1.30 (Paley, Wiener and Zygmund 1933). *Almost surely, Brownian motion is nowhere differentiable. Furthermore, almost surely, for all t,*

$$\textit{either} \quad D^*B(t) = +\infty \quad \textit{or} \quad D_*B(t) = -\infty \quad \textit{or both..}$$

Proof. Suppose that there is a $t_0 \in [0,1]$ such that $-\infty < D_*B(t_0) \le D^*B(t_0) < \infty$. Then

$$\limsup_{h \downarrow 0} \frac{|B(t_0+h) - B(t_0)|}{h} < \infty,$$

and, using the boundedness of Brownian motion on $[0,2]$, this implies that for some finite constant M there exists t_0 with

$$\sup_{h \in [0,1]} \frac{|B(t_0+h) - B(t_0)|}{h} \le M.$$

It suffices to show that this event has probability zero for any M. From now on fix M. If t_0 is contained in the binary interval $[(k-1)/2^n, k/2^n]$ for $n > 2$, then for all $1 \le j \le 2^n - k$ the triangle inequality gives

$$\begin{aligned}\big|B\left((k+j)/2^n\right) - &B\left((k+j-1)/2^n\right)\big| \\ &\le |B\left((k+j)/2^n\right) - B(t_0)| + |B(t_0) - B\left((k+j-1)/2^n\right)| \\ &\le M(2j+1)/2^n.\end{aligned}$$

Define events

$$\Omega_{n,k} := \Big\{\big|B\left((k+j)/2^n\right) - B\left((k+j-1)/2^n\right)\big| \le M(2j+1)/2^n \text{ for } j = 1,2,3\Big\}.$$

Then by independence of the increments and the scaling property, for $1 \le k \le 2^n - 3$,

$$\begin{aligned}\mathbb{P}(\Omega_{n,k}) &\le \prod_{j=1}^{3} \mathbb{P}\big\{\big|B\left((k+j)/2^n\right) - B\left((k+j-1)/2^n\right)\big| \le M(2j+1)/2^n\big\} \\ &\le \mathbb{P}\Big\{|B(1)| \le 7M/\sqrt{2^n}\Big\}^3,\end{aligned}$$

which is at most $(7M2^{-n/2})^3$, since the normal density is bounded by 1/2. Hence

$$\mathbb{P}\left(\bigcup_{k=1}^{2^n-3}\Omega_{n,k}\right) \le 2^n(7M2^{-n/2})^3 = (7M)^3 2^{-n/2},$$

which is summable over all n. Hence, by the Borel-Cantelli lemma,

$$\begin{aligned}\mathbb{P}\Big\{ \text{ there is } t_0 \in [0,1] \text{ with } \sup_{h\in[0,1]} \frac{|B(t_0+h)-B(t_0)|}{h} \le M\Big\} \\ \le \mathbb{P}\left(\bigcup_{k=1}^{2^n-3}\Omega_{n,k} \text{ for infinitely many } n\right) = 0.\end{aligned}$$ ■

REMARK 1.31. The proof of Theorem 1.30 can be tightened to prove that, for any $\alpha > \frac{1}{2}$, Brownian motion is, almost surely, nowhere locally α-Hölder continuous, see Exercise 1.7. ◇

REMARK 1.32. There is an abundance of interesting statements about the right derivatives of Brownian motion, which we state as exercises at the end of the chapter. As a taster we mention here that Lévy [**Le54**] asked whether, almost surely, $D^*B(t) \in \{-\infty, \infty\}$ for every $t \in [0, 1)$. Exercise 1.11 shows that this is not the case. ◇

Another important regularity property, which Brownian motion does *not* possess is to be of bounded variation. We first define what it means for a function to be of bounded variation.

DEFINITION 1.33. *A right-continuous function $f\colon [0,t] \to \mathbb{R}$ is a function of* **bounded variation** *if*

$$V_f^{(1)}(t) := \sup \sum_{j=1}^{k} \big|f(t_j) - f(t_{j-1})\big| < \infty,$$

where the supremum is over all $k \in \mathbb{N}$ and partitions $0 = t_0 \le t_1 \le \dots \le t_{k-1} \le t_k = t$. If the supremum is infinite f is said to be of **unbounded variation**. ◇

REMARK 1.34. It is not hard to show that f is of bounded variation if and only if it can be written as the difference of two increasing functions. ◇

THEOREM 1.35. *Suppose that the sequence of partitions*

$$0 = t_0^{(n)} \le t_1^{(n)} \le \dots \le t_{k(n)-1}^{(n)} \le t_{k(n)}^{(n)} = t$$

is nested, i.e. at each step one or more partition points are added, and the mesh

$$\Delta(n) := \sup_{1\le j\le k(n)} \big\{t_j^{(n)} - t_{j-1}^{(n)}\big\}$$

converges to zero. Then, almost surely,

$$\lim_{n\to\infty} \sum_{j=1}^{k(n)} \big(B(t_j^{(n)}) - B(t_{j-1}^{(n)})\big)^2 = t,$$

and therefore Brownian motion is of unbounded variation.

Remark 1.36. For a sequence of partitions as above, we define

$$V^{(2)}(t) := \lim_{n\to\infty} \sum_{j=1}^{k(n)} \big(B(t_j^{(n)}) - B(t_{j-1}^{(n)})\big)^2$$

to be the **quadratic variation** of Brownian motion. The fact that Brownian motion has finite quadratic variation will be of crucial importance in Chapter 7, however, the analogy to the notion of bounded variation of a function is not perfect: In Exercise 1.13 we find a sequence of partitions

$$0 = t_0^{(n)} \le t_1^{(n)} \le \cdots \le t_{k(n)-1}^{(n)} \le t_{k(n)}^{(n)} = t$$

with mesh converging to zero, such that almost surely

$$\limsup_{n\to\infty} \sum_{j=1}^{k(n)} \big(B(t_j^{(n)}) - B(t_{j-1}^{(n)})\big)^2 = \infty.$$

In particular, the condition that the partitions in Theorem 1.35 are nested cannot be dropped entirely, though it can be replaced by other conditions, see Exercise 1.14. $\diamond$

The proof of Theorem 1.35 is based on the following simple lemma.

Lemma 1.37. *If X, Z are independent, symmetric random variables in L^2, then*

$$\mathbb{E}\big[(X+Z)^2 \,\big|\, X^2+Z^2\big] = X^2 + Z^2.$$

Proof. By symmetry of Z we have

$$\mathbb{E}\big[(X+Z)^2 \,\big|\, X^2+Z^2\big] = \mathbb{E}\big[(X-Z)^2 \,\big|\, X^2+Z^2\big].$$

Both sides of the equation are finite, so that we can take the difference and obtain

$$\mathbb{E}\big[XZ \,\big|X^2+Z^2\big] = 0,$$

and the result follows immediately. ■

Proof of Theorem 1.35. By the Hölder property, we can find, for any $\alpha \in (0, 1/2)$, an n such that $|B(a) - B(b)| \le |a-b|^\alpha$ for all $a, b \in [0,t]$ with $|a-b| \le \Delta(n)$. Hence

$$\sum_{j=1}^{k(n)} \big|B\big(t_j^{(n)}\big) - B\big(t_{j-1}^{(n)}\big)\big| \ge \Delta(n)^{-\alpha} \sum_{j=1}^{k(n)} \big(B\big(t_j^{(n)}\big) - B\big(t_{j-1}^{(n)}\big)\big)^2.$$

Therefore, once we show that the random variables

$$X_n := \sum_{j=1}^{k(n)} \big(B\big(t_j^{(n)}\big) - B\big(t_{j-1}^{(n)}\big)\big)^2$$

converge almost surely to a positive random variable it follows immediately that Brownian motion is almost surely of unbounded variation. By inserting elements in the sequence, if necessary, we may assume that at each step exactly one point is added to the partition.

To see that $\{X_n : n \in \mathbb{N}\}$ converges we use the theory of martingales in discrete time, see Appendix II.3 for basic facts on martingales. We denote by $\mathcal{G}_n$ the σ-algebra generated by the random variables $X_n, X_{n+1}, \ldots$. Then

$$\mathcal{G}_\infty := \bigcap_{k=1}^{\infty} \mathcal{G}_k \subset \cdots \subset \mathcal{G}_{n+1} \subset \mathcal{G}_n \subset \cdots \subset \mathcal{G}_1 .$$

We show that $\{X_n : n \in \mathbb{N}\}$ is a reverse martingale, i.e. that almost surely,

$$X_n = \mathbb{E}\big[X_{n-1} \,\big|\, \mathcal{G}_n\big] \qquad \text{for all } n \geq 2\,.$$

This is easy with the help of Lemma 1.37. Indeed, if $s \in (t_1, t_2)$ is the inserted point we apply it to the symmetric, independent random variables $B(s) - B(t_1)$, $B(t_2) - B(s)$ and denote by $\mathcal{F}$ the σ-algebra generated by $(B(s) - B(t_1))^2 + (B(t_2) - B(s))^2$. Then

$$\mathbb{E}\big[\big(B(t_2) - B(t_1)\big)^2 \big| \,\mathcal{F}\big] = \big(B(s) - B(t_1)\big)^2 + \big(B(t_2) - B(s)\big)^2,$$

and hence

$$\mathbb{E}\big[\big(B(t_2) - B(t_1)\big)^2 - \big(B(s) - B(t_1)\big)^2 - \big(B(t_2) - B(s)\big)^2 \big| \,\mathcal{F}\big] = 0,$$

which implies that $\{X_n : n \in \mathbb{N}\}$ is a reverse martingale.

By the Lévy downward theorem, see Theorem II.4.10,

$$\lim_{n\uparrow\infty} X_n = \mathbb{E}[X_1 \,|\, \mathcal{G}_\infty] \qquad \text{almost surely.}$$

The limit has expectation $\mathbb{E}[X_1] = t$ and, by Fatou's lemma, its variance is bounded by

$$\liminf_{n\uparrow\infty} \mathbb{E}\big[(X_n - \mathbb{E}X_n)^2\big] = \liminf_{n\uparrow\infty} 3 \sum_{j=1}^{k(n)} \big(t_j^{(n)} - t_{j-1}^{(n)}\big)^2 \leq 3 \liminf_{n\uparrow\infty} \Delta(n) = 0.$$

Hence, $\mathbb{E}[X_1 \,|\, \mathcal{G}_\infty] = t$ almost surely, as required. ∎

4. The Cameron-Martin theorem

In this section we have a look at Brownian motion with drift and ask, when we can transfer results about driftless Brownian motion to Brownian motion with drift.

Exercises

Exercise 1.1. Let $\{B(t) : t \geq 0\}$ be a Brownian motion with arbitrary starting point. Show that, for all $s, t \geq 0$, we have $\mathrm{Cov}(B(s), B(t)) = s \wedge t$.

Exercise 1.2 (∗). Show that Brownian motion with start in $x \in \mathbb{R}$ is a Gaussian process.

Exercise 1.3. Show that, for every point $x \in \mathbb{R}$, there exists a *two-sided Brownian motion* $\{B(t) \colon t \in \mathbb{R}\}$ with $B(0) = x$, which has continuous paths, independent increments and the property that, for all $t \in \mathbb{R}$ and $h > 0$, the increments $B(t+h) - B(t)$ are normally distributed with expectation zero and variance h.

Exercise 1.4 (∗). Fix $x, y \in \mathbb{R}$. The *Brownian bridge* with start in x and end in y is the process $\{X(t) \colon 0 \leq t \leq 1\}$ defined by

$$X(t) = B(t) - t\big(B(1) - y\big), \qquad \text{for } 0 \leq t \leq 1\,,$$

where $\{B(t) \colon t \geq 0\}$ is a Brownian motion started in x. The Brownian bridge is an almost surely continuous process such that $X(0) = x$ and $X(1) = y$.

(a) Show that, for every bounded $f \colon \mathbb{R}^n \to \mathbb{R}$,

$$\mathbb{E}\big[f\big(X(t_1), \ldots, X(t_n)\big)\big] = \int f(x_1, \ldots, x_n)\, \frac{\mathfrak{p}(t_1, x, x_1)}{\mathfrak{p}(1, x, y)} \times \prod_{i=2}^{n} \mathfrak{p}(t_i - t_{i-1}, x_i, x_{i+1}) \mathfrak{p}(1 - t_n, x_n, y)\, dx_1 \ \ldots dx_n,$$

for all $0 < t_1 < \cdots < t_n < 1$ where

$$\mathfrak{p}(t, x, y) = \frac{1}{\sqrt{2\pi t}}\, e^{-\frac{(y-x)^2}{2t}}\,.$$

(b) Infer that, for any $t_0 < 1$, the processes $\{X(t) \colon 0 \leq t \leq t_0\}$ and $\{B(t) \colon 0 \leq t \leq t_0\}$ are mutually absolutely continuous.

Exercise 1.5 (∗). Prove the law of large numbers in Corollary 1.11 directly. Use the law of large numbers for sequences of independent identically distributed random variables to show that $\lim_{n\to\infty} B(n)/n = 0$. Then show that $B(t)$ does not oscillate too much between n and $n+1$.

Exercise 1.6 (∗). Show the following improvement to Theorem 1.14: Almost surely,

$$\lim_{h \downarrow 0} \sup_{0 \leq t \leq 1-h} \frac{|B(t+h) - B(t)|}{\sqrt{2h \log(1/h)}} = 1\,.$$

Exercise 1.7 (∗). Show that, if $\alpha > 1/2$, then, almost surely, at every point, Brownian motion fails to be locally α-Hölder continuous.

Exercise 1.8 (∗). Show that, if A is an exchangeable event for an independent, identically distributed sequence, then $\mathbb{P}(A)$ is 0 or 1.

Exercise 1.9. Show that, for a Brownian motion $\{B(t)\colon t \geq 0\}$,

(a) for all $t \geq 0$ we have $\mathbb{P}\{t \text{ is a local maximum}\} = 0$;
(b) almost surely local maxima exist;
(c) almost surely, there exist times $t_*, t^* \in [0, 1)$ with $D^*B(t^*) \leq 0$ and $D_*B(t_*) \geq 0$.

Exercise 1.10 (∗). Let $f \in \mathcal{C}[0, 1]$ any fixed continuous function. Show that, almost surely, the function $\{B(t) + f(t) : t \in [0, 1]\}$ is nowhere differentiable.

Exercise 1.11 (∗). Show that, almost surely, there exists a time t at which $D^*B(t) = 0$.

Exercise 1.12 (∗∗). Show that, almost surely,

$$D^*B(t_0) = -\infty,$$

where t_0 is uniquely determined by

$$B(t_0) = \max_{0 \leq t \leq 1} B(t).$$

Hint. Try this exercise *after* the discussion of the strong Markov property in Chapter 2.

Exercise 1.13 ($**$).

(a) Show that, almost surely, there exists a family
$$0 = t_0^{(n)} \le t_1^{(n)} \le \cdots \le t_{k(n)-1}^{(n)} \le t_{k(n)}^{(n)} = t$$
of (random) partitions such that
$$\lim_{n\uparrow\infty} \sum_{j=1}^{k(n)} \big(B(t_j^{(n)}) - B(t_{j-1}^{(n)})\big)^2 = \infty\,.$$
Hint. Use the construction of Brownian motion to pick a partition consisting of dyadic intervals, such that the increment of Brownian motion over any chosen interval is large relative to the square root of its length.

(b) Construct a (non-random) sequence of partitions
$$0 = t_0^{(n)} \le t_1^{(n)} \le \cdots \le t_{k(n)-1}^{(n)} \le t_{k(n)}^{(n)} = t$$
with mesh converging to zero, such that, almost surely,
$$\limsup_{n\to\infty} \sum_{j=1}^{k(n)} \big(B(t_j^{(n)}) - B(t_{j-1}^{(n)})\big)^2 = \infty.$$

Exercise 1.14 ($*$). Consider a (not necessarily nested) sequence of partitions
$$0 = t_0^{(n)} \le t_1^{(n)} \le \cdots \le t_{k(n)-1}^{(n)} \le t_{k(n)}^{(n)} = t$$
with mesh converging to zero.

(a) Show that, in the sense of L^2-convergence,
$$\lim_{n\to\infty} \sum_{j=1}^{k(n)} \big(B(t_j^{(n)}) - B(t_{j-1}^{(n)})\big)^2 = t.$$

(b) Show that, if additionally
$$\sum_{n=1}^{\infty} \sum_{j=1}^{k(n)} \big(t_j^{(n)} - t_{j-1}^{(n)}\big)^2 < \infty,$$
then the convergence in (a) also holds almost surely.

Notes and Comments

The first study of the mathematical process of Brownian motion is due to Bachelier in [**Ba00**] in the context of modeling stock market fluctuations, see [**DE06**] for a modern edition. Bachelier's work was long forgotten and has only recently been rediscovered, today an international society for mathematical finance is named after him. The physical phenomenon of Brownian motion is usually attributed to Brown [**Br28**] and was explained by Einstein in [**Ei05**]. The first rigorous construction of mathematical Brownian motion is due to Wiener [**Wi23**].

There is a variety of constructions of Brownian motion in the literature. The approach we have followed goes back to one of the great pioneers of Brownian motion, the French mathematician *Paul Lévy*, see [**Le48**]. Lévy's construction has the advantage that continuity properties of Brownian motion can be obtained from the construction. An alternative is to first show that a Markov process with the correct transition probabilities can be constructed, and then to use an abstract criterion, like Kolmogorov's criterion for the existence of a continuous version of the process. See, for example, [**RY94**], [**KS88**] and [**Ka85**] for further alternative constructions.

Gaussian processes, only briefly mentioned here, are one of the richest and best understood class of processes in probability theory. Some good references for this are [**Ad90**] and [**Li95**]. A lot of effort in current research is put into trying to extend our understanding of Brownian motion to more general Gaussian processes like the so-called *fractional Brownian motion*. The main difficulty is that these processes do not have the extremely useful Markov property — which we shall discuss in the next chapter, and which we will make heavy use of throughout the book.

The modulus of continuity, Theorem 1.14, goes back to Lévy [**Le37**]. Observe that this result describes continuity of Brownian motion near its *worst* time. By contrast, the law of the iterated logarithm in the form of Corollary 5.3 shows that at a *typical* time the continuity properties of Brownian motion are better: For every fixed time $t > 0$ and $c > \sqrt{2}$, almost surely, there exists $\varepsilon > 0$ with $|B(t) - B(t+h)| \le c\sqrt{h \log\log(1/h)}$ for all $|h| < \varepsilon$. In Chapter 10 we explore for how many times $t > 0$ we are close to the worst case scenario.

The existence of points where Brownian motion is locally 1/2-Hölder continuous is a very tricky question. Dvoretzky [**Dv63**] showed that, for a sufficiently small $c > 0$, almost surely no point satisfies 1/2-local Hölder continuity with Hölder constant c. Later, Davis [**Da83**] and, independently, Greenwood and Perkins [**GP83**] identified the maximal possible Hölder constant, we will discuss their work in Chapter 10.

There is a lot of discussion about nowhere differentiable, continuous functions in the analysis literature of the early twentieth century. Examples are Weierstrass' function, see e.g. [**MG84**], and van der Waerden's function, see e.g. [**Bi82**]. Nowhere differentiability of Brownian motion was first shown by Paley, Wiener and Zygmund in [**PWZ33**], but the proof we give is due to Dvoretzky, Erdős and Kakutani [**DEK61**].

Besides the discussion of special examples of such functions, the statement that in some sense 'most' or 'almost all' continuous functions are nowhere differentiable is particularly fascinating. A topological form of this statement is that nowhere differentiability is a generic property for the space $\mathcal{C}([0,1])$ in the sense of Baire category. A newer, measure theoretic approach based on an idea of Christensen [**Ch72**], which was later rediscovered by Hunt, Sauer, and Yorke [**HSY92**], is the notion of prevalence. A subset A of a separable Banach space X is called *prevalent* if there exists a Borel probability measure μ on X such that $\mu(x+A)=1$ for any $x \in X$. A strengthening of the proof of Theorem 1.30, see Exercise 1.10, shows that the set of nowhere differentiable functions is prevalent.

The time t where $D^*B(t)=0$ which we constructed in Exercise 1.11 is an exceptional time, i.e. a time where Brownian motion behaves differently from almost every other time. In Chapter 10 we enter a systematic discussion of such times, and in particular address the question how many exceptional points (in terms of Hausdorff dimension) of a certain type exist. The set of times where $D^*B(t)=0$ has Hausdorff dimension $1/4$, see Barlow and Perkins [**BP84**].

The interesting fact that the 'true' quadratic variation of Brownian motion, taken as a supremum over arbitrary partitions with mesh going to zero, is infinite is a result of Lévy, see [**Le40**]. Finer variation properties of Brownian motion have been studied by Taylor in [**Ta72**]. He shows, for example, that the ψ-variation

$$V^{\psi} = \sup \sum_{i=1}^{k} \psi\big(|B(t_i) - B(t_{i-1})|\big),$$

where the supremum is taken over all partitions $0 = t_0 < \cdots < t_k = 1$, $k \in \mathbb{N}$, is finite almost surely for $\psi_1(s) = s^2/(2\log\log(1/s))$, but is infinite for any ψ with $\psi(s)/\psi_1(s) \to \infty$ as $s \to 0$.

CHAPTER 2

Brownian motion as a strong Markov process

In this chapter we discuss the strong Markov property of Brownian motion. We also briefly discuss Markov processes in general and show that some processes, which can be derived from Brownian motion, are also Markov processes. We then exploit these facts to get finer properties of Brownian sample paths.

1. The Markov property and Blumenthal's 0-1 Law

For the discussion of the Markov property we include higher dimensional Brownian motion, which can be defined easily by requiring the characteristics of a linear Brownian motion in every component, and independence of the components.

DEFINITION 2.1. *If $B_1, \ldots, B_d$ are independent linear Brownian motions started in $x_1, \ldots, x_d$, then the stochastic process $\{B(t) : t \geq 0\}$ given by*

$$B(t) = (B_1(t), \ldots, B_d(t))^{\mathrm{T}}$$

is called a **d-dimensional Brownian motion** *started in $(x_1, \ldots, x_d)^{\mathrm{T}}$. The d-dimensional Brownian motion started in the origin is also called* **standard Brownian motion**. *Two-dimensional Brownian motion is also called* **planar Brownian motion**. ◇

NOTATION 2.2. Throughout this book we write $\mathbb{P}_x$ for the probability measure which makes the d-dimensional process $\{B(t) : t \geq 0\}$ a Brownian motion started in $x \in \mathbb{R}^d$, and $\mathbb{E}_x$ for the corresponding expectation. ◇

Suppose now that $\{X(t) : t \geq 0\}$ is a stochastic process. Intuitively, the **Markov property** says that if we know the process $\{X(t) : t \geq 0\}$ on the interval $[0, s]$, for the prediction of the future $\{X(t) : t \geq s\}$ this is as useful as just knowing the endpoint $X(s)$. Moreover, a process is called a **(time-homogeneous) Markov process** if it starts afresh at any fixed time s. Slightly more precisely this means that, supposing the process can be started in any point $X(0) = x \in \mathbb{R}^d$, the time-shifted process $\{X(s+t) : t \geq 0\}$ has the same distribution as the process started in $X(s) \in \mathbb{R}^d$. We shall formalise the notion of a Markov process later in this chapter, but start by giving a straight formulation of the facts for a Brownian motion.

Note that two stochastic processes $\{X(t) : t \geq 0\}$ and $\{Y(t) : t \geq 0\}$ are called **independent**, if for any sets $t_1, \ldots, t_n \geq 0$ and $s_1, \ldots, s_m \geq 0$ of times the vectors $(X(t_1), \ldots, X(t_n))$ and $(Y(s_1), \ldots, Y(s_m))$ are independent.

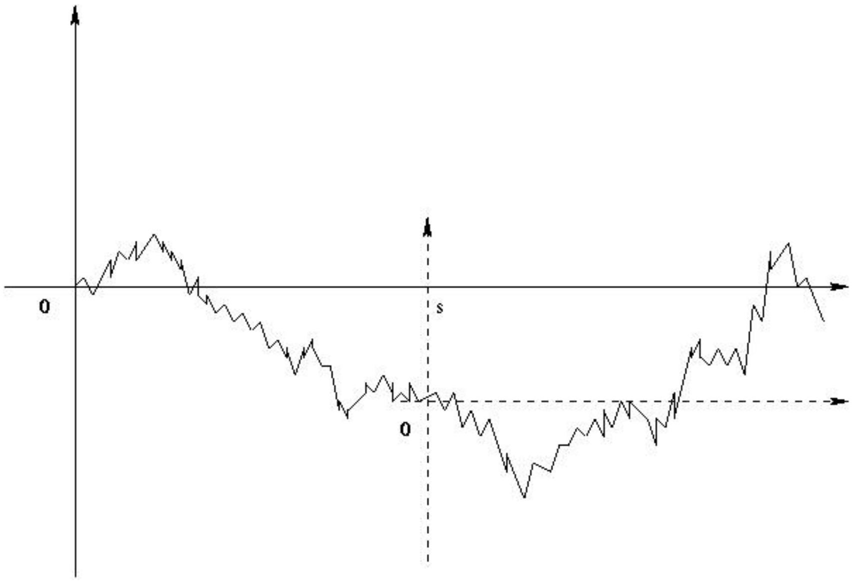

FIGURE 1. Brownian motion starts afresh at time s.

THEOREM 2.3 (Markov property). *Suppose that $\{B(t) : t \geq 0\}$ is a Brownian motion started in $x \in \mathbb{R}^d$. Let $s > 0$, then the process $\{B(t+s) - B(s) : t \geq 0\}$ is again a Brownian motion started in the origin and it is independent of the process $\{B(t) : 0 \leq t \leq s\}$.*

PROOF. It is trivial to check that $\{B(t+s) - B(s) : t \geq 0\}$ satisfies the definition of a d-dimensional Brownian motion. The independence statement follows directly from the independence of the increments of a Brownian motion. ■

We now improve this result slightly and introduce some useful terminology.

DEFINITION 2.4. *A* **filtration** *on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a family $(\mathcal{F}(t) : t \geq 0)$ of σ-algebras such that $\mathcal{F}(s) \subset \mathcal{F}(t) \subset \mathcal{F}$ for all $s < t$. A probability space together with a filtration is sometimes called a* **filtered probability space**. *A stochastic process $\{X(t) : t \geq 0\}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ is called* **adapted** *if $X(t)$ is $\mathcal{F}(t)$-measurable for any $t \geq 0$.* ◇

Suppose we have a Brownian motion $\{B(t) : t \geq 0\}$ defined on some probability space, then we can define a filtration $(\mathcal{F}^0(t) : t \geq 0)$ by letting $\mathcal{F}^0(t)$ be the σ-algebra generated by the random variables $\{B(s) : 0 \leq s \leq t\}$. With this definition, the Brownian motion is obviously adapted to the filtration. Intuitively, this σ-algebra contains all the information available from observing the process up to time t.

By Theorem 2.3, the process $\{B(t+s) - B(s) : t \geq 0\}$ is independent of $\mathcal{F}^0(s)$. In a first step, we improve this and allow a slightly larger (augmented) σ-algebra $\mathcal{F}^+(s)$ defined by

$$\mathcal{F}^+(s) := \bigcap_{t>s} \mathcal{F}^0(t) .$$

Clearly, the family $(\mathcal{F}^+(t) : t \geq 0)$ is again a filtration and $\mathcal{F}^+(s) \supset \mathcal{F}^0(s)$, but intuitively $\mathcal{F}^+(s)$ is a bit larger than $\mathcal{F}^0(s)$, allowing an additional infinitesimal glance into the future.

Theorem 2.5. *For every $s \geq 0$ the process $\{B(t+s) - B(s) : t \geq 0\}$ is independent of $\mathcal{F}^+(s)$.*

Proof. By continuity $B(t+s) - B(s) = \lim_{n\to\infty} B(s_n + t) - B(s_n)$ for a strictly decreasing sequence $\{s_n : n \in \mathbb{N}\}$ converging to s. By Theorem 2.3, for any $t_1, \ldots, t_m \geq 0$, the vector $(B(t_1+s) - B(s), \ldots, B(t_m+s) - B(s)) = \lim_{j\uparrow\infty}(B(t_1+s_j) - B(s_j), \ldots, B(t_m+s_j) - B(s_j))$ is independent of $\mathcal{F}^+(s)$, and so is the process $\{B(t+s) - B(s) : t \geq 0\}$. ∎

Remark 2.6. An alternative way of stating this is that conditional on $\mathcal{F}^+(s)$ the process $\{B(t+s) : t \geq 0\}$ is a Brownian motion started in $B(s)$. ◇

We now look at the **germ σ-algebra** $\mathcal{F}^+(0)$, which heuristically comprises all events defined in terms of Brownian motion on an infinitesimal small interval to the right of the origin.

Theorem 2.7 (Blumenthal's 0-1 law). *Let $x \in \mathbb{R}^d$ and $A \in \mathcal{F}^+(0)$. Then $\mathbb{P}_x(A) \in \{0,1\}$.*

Proof. Using Theorem 2.5 for $s = 0$ we see that any $A \in \sigma\{B(t) \colon t \geq 0\}$ is independent of $\mathcal{F}^+(0)$. This applies in particular to $A \in \mathcal{F}^+(0)$, which therefore is independent of itself, hence has probability zero or one. ∎

As a first application we show that a standard linear Brownian motion has positive and negative values and zeros in every small interval to the right of 0. We have studied this remarkable property of Brownian motion already by different means, in the discussion following Theorem 1.27.

Theorem 2.8. *Suppose $\{B(t) \colon t \geq 0\}$ is a linear Brownian motion. Define $\tau = \inf\{t > 0 : B(t) > 0\}$ and $\sigma = \inf\{t > 0 : B(t) = 0\}$. Then*

$$\mathbb{P}_0\{\tau = 0\} = \mathbb{P}_0\{\sigma = 0\} = 1\,.$$

Proof. The event

$$\{\tau = 0\} = \bigcap_{n=1}^{\infty} \Big\{ \text{ there is } 0 < \varepsilon < 1/n \text{ such that } B(\varepsilon) > 0 \Big\}$$

is clearly in $\mathcal{F}^+(0)$. Hence we just have to show that this event has positive probability. This follows, as, for $t > 0$, $\mathbb{P}_0\{\tau \leq t\} \geq \mathbb{P}_0\{B(t) > 0\} = 1/2$. Hence $\mathbb{P}_0\{\tau = 0\} \geq 1/2$ and we have shown the first part. The same argument works replacing $B(t) > 0$ by $B(t) < 0$ and from these two facts $\mathbb{P}_0\{\sigma = 0\} = 1$ follows, using the intermediate value property of continuous functions. ∎

A further application is a 0-1 law for the **tail σ-algebra** of a Brownian motion. Define $\mathcal{G}(t) = \sigma\{B(s) \colon s \geq t\}$. Let $\mathcal{T} = \bigcap_{t\geq 0} \mathcal{G}(t)$ be the σ-algebra of all **tail events**.

Theorem 2.9 (Kolmogorov's 0-1 Law). *Let $x \in \mathbb{R}^d$ and $A \in \mathcal{T}$. Then $\mathbb{P}_x(A) \in \{0,1\}$.*

Proof. It suffices to look at the case $x = 0$. Under the time inversion of Brownian motion, the tail σ-algebra is mapped on the germ σ-algebra, which is trivial by Blumenthal's 0-1 law. ∎

Remark 2.10. In Exercise 2.2 we shall see that, for any tail event $A \in \mathcal{T}$, the probability $\mathbb{P}_x(A)$ is independent of x. For a germ event $A \in \mathcal{F}^+(0)$, however, the probability $\mathbb{P}_x(A)$ may depend on x. $\diamond$

As final example of this section we now exploit the Markov property to show that the set of local extrema of a linear Brownian motion is a countable, dense subset of $[0, \infty)$. We shall use the easy fact, proved in Exercise 2.3, that every local maximum of Brownian motion is a strict local maximum.

Proposition 2.11. *The set M of times where a linear Brownian motion assumes its local maxima is almost surely countable and dense.*

Proof. Consider the function from the set of non-degenerate closed intervals with rational endpoints to $\mathbb{R}$ given by

$$[a, b] \mapsto \inf \big\{ t \geq a \, : \, B(t) = \max_{a \leq s \leq b} B(s) \big\}.$$

The image of this map contains the set M almost surely by Exercise 2.3. This shows that M is countable almost surely. We already know that Brownian motion has no interval of increase or decrease almost surely, by Proposition 1.22. It follows that it almost surely has a local maximum in every nondegenerate interval. ■

2. The strong Markov property and the reflection principle

Heuristically, the Markov property states that Brownian motion is started anew at each *deterministic time instance*. It is a crucial property of Brownian motion that this holds also for an important class of random times. These random times are called *stopping times*.

The basic idea is that a random time T is a stopping time if we can decide whether $\{T < t\}$ by just knowing the path of the stochastic process up to time t. Think of the situation that T is the first moment where some random event related to the process happens.

Definition 2.12. *A random variable T with values in $[0, \infty]$, defined on a probability space with filtration $(\mathcal{F}(t) : t \geq 0)$ is called a* **stopping time** *if $\{T < t\} \in \mathcal{F}(t)$, for every $t \geq 0$. It is called a* **strict stopping time** *if $\{T \leq t\} \in \mathcal{F}(t)$, for every $t \geq 0$.* $\diamond$

It is easy to see that every strict stopping time is also a stopping time. This follows from

$$\{T < t\} = \bigcup_{n=1}^{\infty} \{T \leq t - 1/n\} \in \mathcal{F}(t) \, .$$

For certain nice filtrations strict stopping times and stopping times agree. In order to come into this situation, we are going to work with the filtration $(\mathcal{F}^+(t) : t \geq 0)$ in the case of Brownian motion and refer the notions of stopping time, etc. always to this filtration. As this filtration is larger than $(\mathcal{F}^0(t) : t \geq 0)$, our choice produces *more stopping times*. The crucial property which distinguishes $\{\mathcal{F}^+(t) : t \geq 0\}$ from $\{\mathcal{F}^0(t) : t \geq 0\}$ is **right-continuity**, which means that

$$\bigcap_{\varepsilon>0} \mathcal{F}^+(t+\varepsilon) = \mathcal{F}^+(t) .$$

To see this note that

$$\bigcap_{\varepsilon>0} \mathcal{F}^+(t+\varepsilon) = \bigcap_{n=1}^{\infty} \bigcap_{k=1}^{\infty} \mathcal{F}^0(t+1/n+1/k) = \mathcal{F}^+(t) .$$

THEOREM 2.13. *Every stopping time T with respect to the filtration $(\mathcal{F}^+(t) : t \geq 0)$, or indeed with respect to any right-continuous filtration, is automatically a strict stopping time.*

Proof. Suppose that T is a stopping time. Then

$$\{T \leq t\} = \bigcap_{k=1}^{\infty} \{T < t+1/k\} \in \bigcap_{n=1}^{\infty} \mathcal{F}^+(t+1/n) = \mathcal{F}^+(t) .$$

Note that this argument uses only the right-continuity of $\{\mathcal{F}^+(t) : t \geq 0\}$. ■

We give some examples.

- Every deterministic time $t \geq 0$ is also a stopping time.
- Suppose $G \subset \mathbb{R}^d$ is an open set. Then $T = \inf\{t \geq 0 : B(t) \in G\}$ is a stopping time.

  Proof. Let $\mathbb{Q}$ be the rationals in $(0,t)$. Then, by continuity of B,

  $$\{T < t\} = \bigcup_{s\in\mathbb{Q}} \{B(s) \in G\} \in \mathcal{F}^+(t) .$$ ■

- If $T_n \uparrow T$ is an increasing sequence of stopping times, then T is also a stopping time.

  Proof. $\{T \leq t\} = \bigcap_{n=1}^{\infty} \{T_n \leq t\} \in \mathcal{F}^+(t)$. ■

- Suppose H is a closed set, for example a singleton. Then $T = \inf\{t \geq 0 : B(t) \in H\}$ is a stopping time.

  Proof. Let $G(n) = \{x \in \mathbb{R}^d : \exists y \in H \text{ with } |x-y| < 1/n\}$ so that $H = \bigcap G(n)$. Then $T_n := \inf\{t \geq 0 : B(t) \in G(n)\}$ are stopping times, which are increasing to T. ■

- Let T be a stopping time. Define stopping times T_n by

  $$T_n = (m+1)2^{-n} \text{ if } m2^{-n} \leq T < (m+1)2^{-n} .$$

  In other words, we stop at the first time of the form $k2^{-n}$ after T. It is easy to see that T_n is a stopping time. We will use it later as a discrete approximation to T.

We define, for every stopping time T, the σ-algebra

$$\mathcal{F}^+(T) = \{A \in \mathcal{A} : A \cap \{T < t\} \in \mathcal{F}^+(t) \text{ for all } t \geq 0\} .$$

This means that the part of A that lies in $\{T < t\}$ should be measurable with respect to the information available at time t. Heuristically, this is the collection of events that happened before the stopping time T. As in the proof of the last theorem we can infer that for right-continuous filtrations like our $(\mathcal{F}^+(t) : t \geq 0)$ the event $\{T \leq t\}$ may replace $\{T < t\}$ without changing the definition.

We can now state and prove the *strong Markov property* for Brownian motion, which was rigorously established by Hunt [**Hu56**] and Dynkin [**Dy57**].

Theorem 2.14 (Strong Markov property). *For every almost surely finite stopping time T, the process $\{B(T+t) - B(T) : t \geq 0\}$ is a standard Brownian motion independent of $\mathcal{F}^+(T)$.*

Proof. We first show our statement for the stopping times T_n which discretely approximate T from above, $T_n = (m+1)2^{-n}$ if $m2^{-n} \leq T < (m+1)2^{-n}$, see the examples above. Write $B_k = \{B_k(t) : t \geq 0\}$ for the Brownian motion defined by $B_k(t) = B(t + k/2^n) - B(k/2^n)$, and $B_* = \{B_*(t) : t \geq 0\}$ for the process defined by $B_*(t) = B(t + T_n) - B(T_n)$. Suppose that $E \in \mathcal{F}^+(T_n)$. Then, for every event $\{B_* \in A\}$, we have

$$\begin{aligned}\mathbb{P}\big(\{B_* \in A\} \cap E\big) &= \sum_{k=0}^{\infty} \mathbb{P}\big(\{B_k \in A\} \cap E \cap \{T_n = k2^{-n}\}\big) \\ &= \sum_{k=0}^{\infty} \mathbb{P}\{B_k \in A\}\, \mathbb{P}\big(E \cap \{T_n = k2^{-n}\}\big),\end{aligned}$$

using that $\{B_k \in A\}$ is independent of $E \cap \{T_n = k2^{-n}\} \in \mathcal{F}^+(k2^{-n})$ by Theorem 2.5. Now, by Theorem 2.3, $\mathbb{P}\{B_k \in A\} = \mathbb{P}\{B \in A\}$ does not depend on k, and hence we get

$$\begin{aligned}\sum_{k=0}^{\infty} \mathbb{P}\{B_k \in A\}\, \mathbb{P}\big(E \cap \{T_n = k2^{-n}\}\big) &= \mathbb{P}\{B \in A\} \sum_{k=0}^{\infty} \mathbb{P}\big(E \cap \{T_n = k2^{-n}\}\big) \\ &= \mathbb{P}\{B \in A\}\mathbb{P}(E),\end{aligned}$$

which shows that B_* is a Brownian motion and independent of E, hence of $\mathcal{F}^+(T_n)$, as claimed.

It remains to generalize this to general stopping times T. As $T_n \downarrow T$ we have that $\{B(s+T_n) - B(T_n) : s \geq 0\}$ is a Brownian motion independent of $\mathcal{F}^+(T_n) \supset \mathcal{F}^+(T)$. Hence the increments

$$B(s+t+T) - B(t+T) = \lim_{n\to\infty} B(s+t+T_n) - B(t+T_n)$$

of the process $\{B(r+T) - B(T) : r \geq 0\}$ are independent and normally distributed with mean zero and variance s. As the process is obviously almost surely continuous, it is a Brownian motion. Moreover all increments, $B(s+t+T) - B(t+T) = \lim B(s+t+T_n) - B(t+T_n)$, and hence the process itself, are independent of $\mathcal{F}^+(T)$. ■

Remark 2.15. Let $T = \inf\{t \geq 0 : B(t) = \max_{0 \leq s \leq 1} B(s)\}$. It is intuitively clear that T is *not* a stopping time. One way to prove it, is to first observe that almost surely $T < 1$. The increment $B(t+T) - B(T)$ does not take positive values in a small neighbourhood to the right of 0, which contradicts the strong Markov property and Theorem 2.8. ◇

2.1. The reflection principle. 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. More formally:

THEOREM 2.16 (Reflection principle). *If T is a stopping time and $\{B(t) : t \geq 0\}$ is a standard Brownian motion, then the process $\{B^*(t) : t \geq 0\}$ called* **Brownian motion reflected at** T *and defined by*

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

is also a standard Brownian motion.

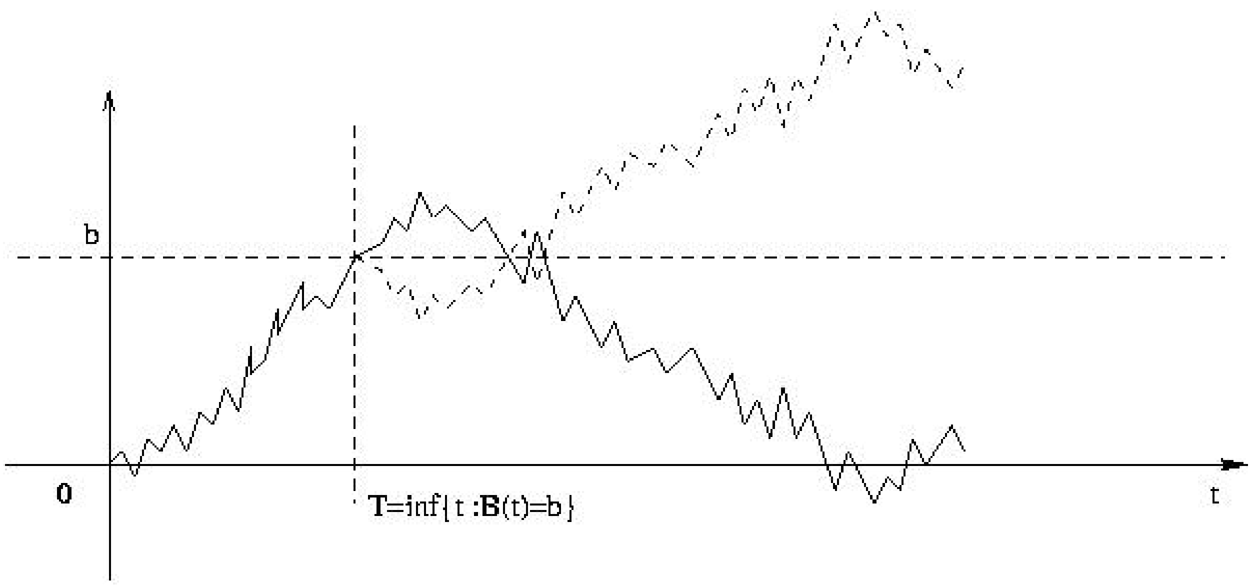

FIGURE 2. The reflection principle

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

$$\{B(t+T) - B(T) : t \geq 0\} \text{ and } \{-(B(t+T) - B(T)) : t \geq 0\}$$

are Brownian motions and independent of the beginning $\{B(t) : t \in [0,T]\}$. Hence the concatenation (gluing together) of the beginning with the first part and the concatenation with the second part have the same distribution. The first is just $\{B(t) : t \geq 0\}$, the second is the object $\{B^*(t) : t \geq 0\}$ introduced in the statement. ■

REMARK 2.17. For a linear Brownian motion, consider $\tau = \inf\{t : B(t) = \max_{0 \leq s \leq 1} B(s)\}$ and let $\{B^*(t) : t \geq 0\}$ be the reflection at τ defined as in Theorem 2.16. Almost surely $\{B(\tau + t) - B(\tau) : t \geq 0\}$ is non-positive on some right neighbourhood of $t = 0$, and hence is *not* Brownian motion. The strong Markov property does not apply here because τ is not a stopping time for the filtration $(\mathcal{F}^+(t) : t \geq 0)$. We shall see in Theorem 5.14 that Brownian motion almost surely has no point of increase. Since τ is a point of increase of the reflected process $\{B^*(t) : t \geq 0\}$, it follows that the distributions of $\{B(t) : t \geq 0\}$ and of $\{B^*(t) : t \geq 0\}$ are singular. ◇

Now specialise to the case of linear Brownian motion. Let $M(t) = \max_{0\le s\le t} B(s)$. 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.

THEOREM 2.18. *If $a > 0$ then* $\mathbb{P}_0\{M(t) > a\} = 2\mathbb{P}_0\{B(t) > a\} = \mathbb{P}_0\{|B(t)| > a\}$.

Proof. Let $T = \inf\{t \ge 0 \,:\, B(t) = a\}$ and let $\{B^*(t) \,:\, t \ge 0\}$ be Brownian motion reflected at T. Then $\{M(t) > a\}$ is the disjoint union of the events $\{B(t) > a\}$ and $\{M(t) > a,\ B(t) \le a\}$ and since the latter is exactly $\{B^*(t) \ge a\}$ the statement follows from the reflection principle.■

REMARK 2.19. Theorem 2.18 is most useful when combined with a tail estimate for the Gaussian as in Lemma II.3.1. For example, for an upper bound we obtain, for all $a > 0$,

$$\mathbb{P}_0\{M(t) > a\} \le \frac{\sqrt{2t}}{a\sqrt{\pi}} \exp\Big\{-\frac{a^2}{2t}\Big\}.$$

◇

2.2. The area of planar Brownian motion. Continuous curves in the plane can still be extremely wild. Space-filling curves, like the Peano curve, can map the time interval $[0,1]$ continuously on sets of positive area, see for example [**La98**]. We now show that the range of planar Brownian motion has zero area. The Markov property and the reflection principle play an important role in the proof.

Suppose $\{B(t) \,:\, t \ge 0\}$ is planar Brownian motion. We denote the Lebesgue measure on $\mathbb{R}^d$ by $\mathcal{L}_d$, and use the symbol $f * g$ to denote the **convolution** of the functions f and g given, whenever well-defined, by

$$f * g\,(x) := \int f(y)g(x-y)\,dy.$$

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

LEMMA 2.20. *If $A_1, A_2 \subset \mathbb{R}^2$ are Borel sets with positive area, then*

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

Proof. We may assume A_1 and A_2 are bounded. By Fubini's theorem,

$$\begin{aligned}\int_{\mathbb{R}^2} \mathbb{1}_{A_1} * \mathbb{1}_{-A_2}(x)\,dx &= \int_{\mathbb{R}^2}\int_{\mathbb{R}^2} \mathbb{1}_{A_1}(w)\mathbb{1}_{A_2}(w-x)\,dw\,dx \\ &= \int_{\mathbb{R}^2} \mathbb{1}_{A_1}(w)\left(\int_{\mathbb{R}^2} \mathbb{1}_{A_2}(w-x)\,dx\right)dw \\ &= \mathcal{L}_2(A_1)\mathcal{L}_2(A_2) > 0.\end{aligned}$$

Thus $\mathbb{1}_{A_1} * \mathbb{1}_{-A_2}(x) > 0$ on a set of positive area. But

$$\begin{aligned}\mathbb{1}_{A_1} * \mathbb{1}_{-A_2}(x) &= \int \mathbb{1}_{A_1}(y)\, \mathbb{1}_{-A_2}(x-y)\, dy = \int \mathbb{1}_{A_1}(y)\, \mathbb{1}_{A_2+x}(y)\, dy \\ &= \mathcal{L}_2(A_1 \cap (A_2 + x))\,,\end{aligned}$$

proving the lemma. ■

We are now ready to prove Lévy's theorem on the area of planar Brownian motion.

THEOREM 2.21 (Lévy 1940). *Almost surely,* $\mathcal{L}_2(B[0,1]) = 0$.

Proof. 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 sidelength $\sqrt{a}/2$. Hence, using Theorem 2.18,

$$\mathbb{P}\{X > a\} \le 2\,\mathbb{P}\{\max_{t\in[0,1]} |W(t)| > \sqrt{a}/2\,\} = 4\,\mathbb{P}\{W(1) > \sqrt{a}/2\} \le 4e^{-a/8},$$

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

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

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 $\mathcal{L}_2(B[0,3]) \le \sum_{j=0}^2 \mathcal{L}_2(B[j,j+1])$ with equality if and only if for $0 \le i < j \le 2$ 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

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

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, $\{B_1(t) : t \in [0,1]\}$ by $B_1(t) = B(t)$, and $\{B_2(t) : t \in [0,1]\}$ by $B_2(t) = B(t+2) - B(2) + B(1)$. Both Brownian motions are independent of the random variable $Y := B(2) - B(1)$. 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 $\{R(x) : x \in \mathbb{R}^2\}$ 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 $R(x) = 0$ almost surely for $\mathcal{L}_2$-almost all x, and hence

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

From Lemma 2.20 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 2.22. How big is the range, or 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 Chapter 4. ◇

Corollary 2.23. *For any points $x, y \in \mathbb{R}^d$, $d \geq 2$, 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 2.21 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,

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

where we have used the Markov property in the second step. ■

Remark 2.24. Loosely speaking, planar Brownian motion almost surely does not hit singletons. Which other sets are not hit by Brownian motion? This clearly depends on the size and shape of the set in some intricate way, and a precise answer will use the notion of capacity, which we study in Chapter 8. ◇

2.3. The zero set of Brownian motion. As a further application of the strong Markov property we have a first look at the properties of the zero set $\{t \geq 0 : B(t) = 0\}$ of one-dimensional Brownian motion. We prove that this set is a closed set with no isolated points (sometimes called a **perfect** set). This is perhaps surprising since, almost surely, a Brownian motion has isolated zeros from the left, for instance the first zero after $1/2$, or from the right, like the last zero before $1/2$.

Theorem 2.25. *Let $\{B(t) : t \geq 0\}$ be a one dimensional Brownian motion and*

$$\mathsf{Zero} = \{t \geq 0 : B(t) = 0\}$$

its **zero set**. *Then, almost surely,* Zero *is a closed set with no isolated points.*

Proof. Clearly, with probability one, Zero is closed because Brownian motion is continuous almost surely. To prove that no point of Zero is isolated we consider the following construction: For each rational $q \in [0, \infty)$ consider the first zero after q, i.e.,

$$\tau_q = \inf\{t \geq q \colon B(t) = 0\}.$$

Note that τ_q is an almost surely finite stopping time. Since Zero is closed, the inf is almost surely a minimum. By the strong Markov property, applied to τ_q, we have that for each q, almost surely τ_q is not an isolated zero from the right. But, since there are only countably many rationals, we conclude that almost surely, for all q rational, τ_q is not an isolated zero from the right.

Our next task is to prove that the remaining points of Z are not isolated from the left. So we claim that any $0 < t \in Z$ which is different from τ_q for all rational q is not an isolated point from the left. To see this take a sequence $q_n \uparrow t$, $q_n \in \mathbb{Q}$. Define $t_n = \tau_{q_n}$. Clearly $q_n \leq t_n < t$ and so $t_n \uparrow t$. Thus t is not isolated from the left. ■

REMARK 2.26. Theorem 2.25 implies that Zero is uncountable, see Exercise 2.7. ◇

3. Markov processes derived from Brownian motion

In this section, we define the concept of a Markov process. Our motivation is that various processes derived from Brownian motion are Markov processes. Among the examples are the reflection of Brownian motion in zero, and the process $\{T_a : a \geq 0\}$ of times T_a when a Brownian motion reaches level a for the first time. We assume that the reader is familiar with the notion of conditional expectation given a σ-algebra, see [**Wi91**] for a reference.

DEFINITION 2.27. *A function $p\colon [0, \infty) \times \mathbb{R}^d \times \mathfrak{B} \to \mathbb{R}$, where $\mathfrak{B}$ is the Borel σ-algebra in $\mathbb{R}^d$, is a* **Markov transition kernel** *provided*

(1) *$p(\,\cdot\,, \cdot\,, A)$ is measurable as a function of (t, x), for each $A \in \mathfrak{B}$;*

(2) *$p(t, x, \cdot\,)$ is a Borel probability measure on $\mathbb{R}^d$ for all $t \geq 0$ and $x \in \mathbb{R}^d$, when integrating a function f with respect to this measure we write*

$$\int f(y)\, p(t, x, dy)\,;$$

(3) *for all $A \in \mathfrak{B}$, $x \in \mathbb{R}^d$ and $t, s > 0$,*

$$p(t + s, x, A) = \int_{\mathbb{R}^d} p(t, y, A)\, p(s, x, dy).$$

An adapted process $\{X(t) : t \geq 0\}$ is a (time-homogeneous) **Markov process** *with transition kernel p with respect to a filtration $(\mathcal{F}(t) : t \geq 0)$, if for all $t \geq s$ and Borel sets $A \in \mathfrak{B}$ we have*

$$\mathbb{P}\{X(t) \in A \,|\, \mathcal{F}(s)\} = p(t - s, X(s), A)\,. \quad ◇$$

Observe that $p(t, x, A)$ is the probability that the process takes a value in A at time t, if it is started at the point x. Readers familiar with *Markov chains* can recognise the pattern behind this definition: The Markov transition kernel p plays the role of the transition matrix P in this setup. The next two examples are trivial consequences of the Markov property for Brownian motion.

EXAMPLE 2.28. Brownian motion is a Markov process and for its transition kernel p the distribution $p(t, x, \cdot\,)$ is a normal distribution with mean x and variance t. Similarly, d-dimensional Brownian motion is a Markov process and $p(t, x, \cdot\,)$ is a Gaussian with mean x and covariance matrix t times identity. Note that property (3) in the definition of the Markov transition kernel is just the fact that the sum of two independent Gaussian random vectors is a Gaussian random vector with the sum of the covariance matrices. $\diamond$

NOTATION 2.29. The transition kernel of d-dimensional Brownian motion is described by probability measures $p(t, x, \cdot\,)$ with densities denoted throughout this book by

$$\mathfrak{p}(t, x, y) = (2\pi t)^{-d/2} \exp\Big(- \frac{|x-y|^2}{2t} \Big). \qquad \diamond$$

EXAMPLE 2.30. The *reflected one-dimensional Brownian motion* $\{X(t) : t \geq 0\}$ defined by $X(t) = |B(t)|$ is a Markov process. Moreover, its transition kernel $p(t, x, \cdot)$ is the law of $|Y|$ for Y normally distributed with mean x and variance t, which we call the *modulus normal distribution* with parameters x and t. $\diamond$

We now prove a famous theorem of Paul Lévy, which shows that the difference of the maximum process of a Brownian motion and the Brownian motion itself is a reflected Brownian motion. To be precise, this means that the difference of the processes has the same marginal distributions as a reflected Brownian motion, and is also almost surely continuous.

THEOREM 2.31 (Lévy 1948). *Let $\{M(t)\colon t \geq 0\}$ be the maximum process of a linear standard Brownian motion $\{B(t)\colon t \geq 0\}$, i.e. the process defined by*

$$M(t) = \max_{0 \leq s \leq t} B(s).$$

Then, the process $\{Y(t) : t \geq 0\}$ defined by $Y(t) = M(t) - B(t)$ is a reflected Brownian motion.

Proof. The main step is to show that the process $Y = \{Y(t) : t \geq 0\}$ is a Markov process and its Markov transition kernel $p(t, x, \cdot\,)$ has modulus normal distribution with parameters x and t. Once this is established, it is immediate that the marginal distributions of Y agree with those of a reflected Brownian motion. Obviously, Y has almost surely continuous paths.

For the main step, fix $s > 0$, consider the two processes $\{\hat{B}(t) : t \geq 0\}$ defined by

$$\hat{B}(t) = B(s+t) - B(s) \text{ for } t \geq 0,$$

and $\{\hat{M}(t) : t \geq 0\}$ defined by

$$\hat{M}(t) = \max_{0 \leq u \leq t} \hat{B}(u) \text{ for } t \geq 0.$$

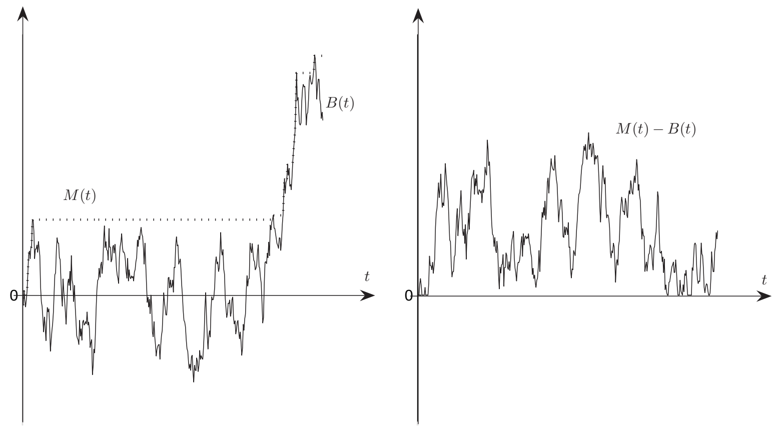

FIGURE 3. On the left, the processes $\{B(t)\colon t \geq 0\}$ with associated maximum process $\{M(t)\colon t \geq 0\}$ indicated by the dashed curve. On the right the process $\{M(t) - B(t)\colon t \geq 0\}$.

Because $Y(s)$ is $\mathcal{F}^+(s)$-measurable, it suffices to check that conditional on $\mathcal{F}^+(s)$, for every $t \geq 0$, the random variable $Y(s+t)$ has the same distribution as $|Y(s) + \hat{B}(t)|$. Indeed, this directly implies that $\{Y(t) : t \geq 0\}$ is a Markov process with the same transition kernel as the reflected Brownian motion.

To prove the claim fix $s, t \geq 0$ and observe that $M(s+t) = M(s) \vee (B(s) + \hat{M}(t))$, and so we have $Y(s+t) = (M(s) \vee (B(s) + \hat{M}(t))) - (B(s) + \hat{B}(t))$. Using the fact that $(a \vee b) - c = (a-c) \vee (b-c)$, we have

$$Y(s+t) = \big(Y(s) \vee \hat{M}(t)\big) - \hat{B}(t).$$

To finish, it suffices to check, for every $y \geq 0$, that $y \vee \hat{M}(t) - \hat{B}(t)$ has the same distribution as $|y + \hat{B}(t)|$. For any $a \geq 0$ write

$$P_1 = \mathbb{P}\{y - \hat{B}(t) > a\}, \qquad P_2 = \mathbb{P}\big\{y - \hat{B}(t) \leq a \text{ and } \hat{M}(t) - \hat{B}(t) > a\big\}.$$

Then $\mathbb{P}\{y \vee \hat{M}(t) - \hat{B}(t) > a\} = P_1 + P_2$. Since $\hat{B}$ has the same distribution as $-\hat{B}$ we have $P_1 = \mathbb{P}\{y + \hat{B}(t) > a\}$. To study the second term it is useful to define the time reversed Brownian motion $\{W(u) : 0 \leq u \leq t\}$ by

$$W(u) := \hat{B}(t-u) - \hat{B}(t) \text{ for } 0 \leq u \leq t.$$

Note that W is also a Brownian motion for $0 \leq u \leq t$ since it is continuous and its finite dimensional distributions are Gaussian with the right covariances.

Let $M_W(t) = \max_{0 \leq u \leq t} W(u)$. Then $M_W(t) = \hat{M}(t) - \hat{B}(t)$. Since $W(t) = -\hat{B}(t)$, we have

$$P_2 = \mathbb{P}\{y + W(t) \leq a \text{ and } M_W(t) > a\}.$$

If we use the reflection principle by reflecting $\{W(u) : 0 \leq u \leq t\}$ at the first time it hits a, we get another Brownian motion $\{W^*(u) : 0 \leq u \leq t\}$. In terms of this Brownian motion we have $P_2 = \mathbb{P}\{W^*(t) \geq a + y\}$. Since W^* has the same distribution as $-\hat{B}$, it follows that

$P_2 = \mathbb{P}\{y + \hat{B}(t) \leq -a\}$. The Brownian motion $\hat{B}$ has continuous distribution, and so, by adding P_1 and P_2, we get

$$\mathbb{P}\{y \vee \hat{M}(t) - \hat{B}(t) > a\} = \mathbb{P}\{|y + \hat{B}(t)| > a\}.$$

This proves the main step and, consequently, the theorem. ■

While, as seen above, $\{M(t) - B(t) : t \geq 0\}$ is a Markov process, it is important to note that the maximum process $\{M(t) : t \geq 0\}$ itself is not a Markov process. However the times when new maxima are achieved form a Markov process, as the following theorem shows.

THEOREM 2.32. *For any $a \geq 0$ define the stopping times*

$$T_a = \inf\{t \geq 0 : B(t) = a\}.$$

Then $\{T_a : a \geq 0\}$ is an increasing Markov process with transition kernels given by the densities

$$s \mapsto p(a,t,s) = \tfrac{a}{\sqrt{2\pi(s-t)^3}} \exp\big(-\tfrac{a^2}{2(s-t)}\big) \, \mathbb{1}\{s > t\}, \qquad \textit{for } a > 0.$$

This process is called the **stable subordinator** *of index $\frac{1}{2}$.*

Proof. Fix $a \geq b \geq 0$ and note that for all $t \geq 0$ we have

$$\{T_a - T_b = t\} = \{B(T_b + s) - B(T_b) < a - b, \text{ for } s < t, \text{ and } B(T_b + t) - B(T_b) = a - b\}.$$

By the strong Markov property of Brownian motion this event is independent of $\mathcal{F}^+(T_b)$ and therefore in particular of $\{T_d : d \leq b\}$. This proves the Markov property of $\{T_a : a \geq 0\}$. The form of the transition kernel follows from the reflection principle,

$$\begin{aligned}
\mathbb{P}\{T_a - T_b \leq t\} &= \mathbb{P}\{T_{a-b} \leq t\} = \mathbb{P}\big\{\max_{0 \leq s \leq t} B(s) \geq a - b\big\} \\
&= 2\mathbb{P}\big\{B(t) \geq a - b\big\} = 2 \int_{a-b}^{\infty} \frac{1}{\sqrt{2\pi t}} \exp\big(-\tfrac{x^2}{2t}\big)\, dx \\
&= \int_0^t \frac{1}{\sqrt{2\pi s^3}}\, (a-b)\, \exp\big(-\tfrac{(a-b)^2}{2s}\big)\, ds,
\end{aligned}$$

where we used the substitution $x = \sqrt{t/s}\,(a-b)$ in the last step. ■

In a similar way there is another important Markov process, the Cauchy process, hidden in the planar Brownian motion, see Figure 4.

THEOREM 2.33. *Let $\{B(t) \colon t \geq 0\}$ be a planar Brownian motion, define a family $(V(a) \colon a \geq 0)$ of vertical lines $V(a) = \{(x,y) \in \mathbb{R}^2 : x = a\}$, and let $T(a) = \tau(V(a))$ be the first hitting time of $V(a)$. Then the process $\{X(a) \colon a \geq 0\}$ defined by $X(a) := B_2(T(a))$ for $B(t) = (B_1(t), B_2(t))^{\mathrm{T}}$ is a Markov process with transition kernel*

$$p(a,x,dy) = \frac{a}{\pi(a^2 + (x-y)^2)}\, dy\,.$$

This process is called the **Cauchy process**.

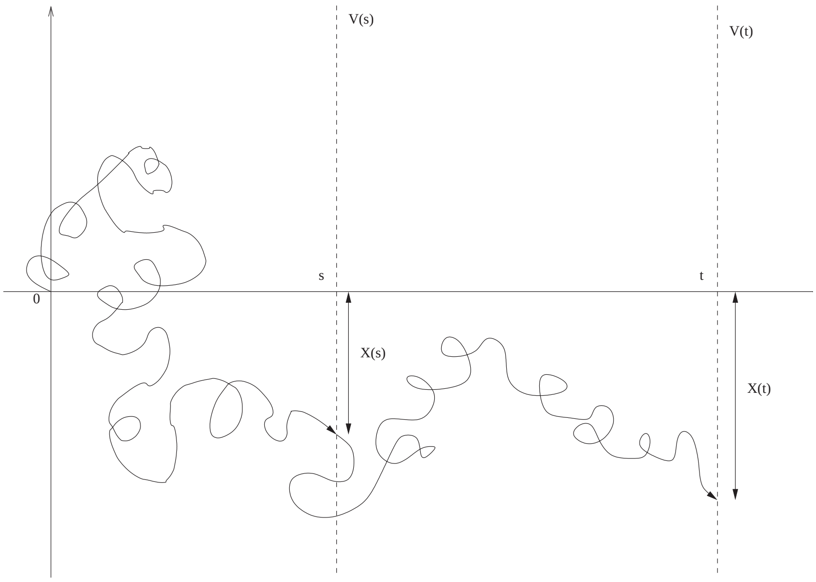

FIGURE 4. The Cauchy process embedded in planar Brownian motion

Proof. The Markov property of $\{X(a)\colon a \geq 0\}$ is a consequence of the strong Markov property of Brownian motion for the stopping times $T(a)$, and the fact that $T(a) < T(b)$ for all $a < b$. In order to calculate the transition density recall from Theorem 2.32 that $T(a)$, which is the first time when the one-dimensional Brownian motion $\{B_1(s)\colon s \geq 0\}$ hits level a, has density

$$\frac{a}{\sqrt{2\pi s^3}}\,\exp\big(-\tfrac{a^2}{2s}\big)\,.$$

$T(a)$ is independent of $\{B_2(s)\colon s \geq 0\}$ and therefore the density of $B_2(T(a))$ is (in the variable x)

$$\int_0^\infty \frac{1}{\sqrt{2\pi s}}\,\exp\big(-\tfrac{x^2}{2s}\big)\,\frac{a}{\sqrt{2\pi s^3}}\,\exp\big(-\tfrac{a^2}{2s}\big)\,ds = \int_0^\infty \frac{a}{\pi(a^2+x^2)}\,e^{-\sigma}\,d\sigma = \frac{a}{\pi(a^2+x^2)}\,,$$

where the integral is evaluated using the substitution $\sigma = \frac{1}{2s}\,(a^2+x^2)$. ■

REMARK 2.34. Alternative proofs of Theorem 2.33, avoiding the explicit evaluation of integrals will be given in Exercise 2.17 and Exercise 7.4. ◇

4. The martingale property of Brownian motion

In the previous section we have taken a particular feature of Brownian motion, the Markov property, and introduced an abstract class of processes, the Markov processes, which share this feature. We have seen that a number of process derived from Brownian motion are again Markov processes and this insight helped us getting new information about Brownian motion. In this section we follow a similar plan, taking a different feature of Brownian motion, the martingale property, as a starting point.

Definition 2.35. *A real-valued stochastic process* $\{X(t) : t \geq 0\}$ *is a* **martingale** *with respect to a filtration* $(\mathcal{F}(t) : t \geq 0)$ *if it is adapted to the filtration,* $\mathbb{E}|X(t)| < \infty$ *for all* $t \geq 0$ *and, for any pair of times* $0 \leq s \leq t$,

$$\mathbb{E}\big[X(t) \,\big|\, \mathcal{F}(s)\big] = X(s) \text{ almost surely.}$$

The process is called a **submartingale** *if* $\geq$ *holds, and a* **supermartingale** *if* $\leq$ *holds in the display above.* ◇

Remark 2.36. Intuitively, a martingale is a process where the current state $X(t)$ is always the best prediction for its further states. In this sense, martingales describe *fair games.* If $\{X(t) : t \geq 0\}$ is a martingale, the process $\{|X(t)| : t \geq 0\}$ need not be a martingale, but it still is a submartingale, as a simple application of the triangle inequality shows. ◇

Example 2.37. For a one-dimensional Brownian motion $\{B(t) : t \geq 0\}$ we have

$$\begin{aligned}\mathbb{E}\big[B(t) \,\big|\, \mathcal{F}^+(s)\big] &= \mathbb{E}\big[B(t) - B(s) \,\big|\, \mathcal{F}^+(s)\big] + B(s) \\ &= \mathbb{E}\big[B(t) - B(s)\big] + B(s) = B(s)\,,\end{aligned}$$

using Theorem 2.5 in the second step. Hence Brownian motion is a martingale. ◇

We now state two useful facts about martingales, which we will exploit extensively: The *optional stopping theorem* and *Doob's maximal inequality.* Both of these results are well-known in the discrete time setting and there is a reminder in Appendix II.3. The natural extension of these results to the continuous time setting is the content of our propositions.

The optional stopping theorem provides a condition under which the defining equation for martingales can be extended from fixed times $0 \leq s \leq t$ to stopping times $0 \leq S \leq T$.

Proposition 2.38 (Optional stopping theorem). *Suppose* $\{X(t) : t \geq 0\}$ *is a continuous martingale, and* $0 \leq S \leq T$ *are stopping times. If the process* $\{X(t \wedge T) : t \geq 0\}$ *is dominated by an integrable random variable* X, *i.e.* $|X(t \wedge T)| \leq X$ *almost surely, for all* $t \geq 0$, *then*

$$\mathbb{E}\big[X(T) \,\big|\mathcal{F}(S)\big] = X(S), \text{ almost surely.}$$

Proof. The best way to prove this is to prove the result first for martingales in discrete time, and then extend the result by approximation. The result for discrete time is provided in our appendix, see Theorem II.4.11. Let us explain the approximation step here.

Fix $N \in \mathbb{N}$ and define a discrete time martingale by $X_n = X(T \wedge n2^{-N})$ and stopping times $S' = \lfloor 2^N S \rfloor + 1$ and $T' = \lfloor 2^N T \rfloor + 1$, with respect to the filtration $(\mathcal{G}(n) : n \in \mathbb{N})$ given by $\mathcal{G}(n) = \mathcal{F}(n2^{-N})$. Obviously X_n is dominated by an integrable random variable and hence the discrete time result gives $\mathbb{E}\big[X_{T'} \,\big|\, \mathcal{G}(S')\big] = X_{S'}$, which translates as $\mathbb{E}\big[X(T) \big| \mathcal{F}(S_N)\big] = X(T \wedge S_N)$, for $S_N = 2^{-N}(\lfloor 2^N S \rfloor + 1)$. Letting $N \uparrow \infty$ and using dominated convergence for conditional expectations gives the result. ■

The following inequality will also be of great use to us.

Proposition 2.39 (Doob's maximal inequality). *Suppose $\{X(t) : t \geq 0\}$ is a continuous submartingale and $p > 1$. Then, for any $t \geq 0$,*

$$\mathbb{E}\Big[\big(\sup_{0\leq s\leq t}|X(t)|\big)^p\Big] \leq \big(\tfrac{p}{p-1}\big)^p\,\mathbb{E}\big[|X(t)|^p\big]\,.$$

Proof. Again this is proved for martingales in discrete time in our appendix, see Theorem II.4.14, and can be extended by approximation. Fix $N \in \mathbb{N}$ and define a discrete time martingale by $X_n = X(tn2^{-N})$ with respect to the filtration $(\mathcal{G}(n) : n \in \mathbb{N})$ given by $\mathcal{G}(n) = \mathcal{F}(tn2^{-N})$. By the discrete version of Doob's maximal inequality,

$$\mathbb{E}\Big[\big(\sup_{1\leq k\leq 2^N} X_k\big)^p\Big] \leq \big(\tfrac{p}{p-1}\big)^p\,\mathbb{E}\big[X_{2^N}^p\big] = \big(\tfrac{p}{p-1}\big)^p\,\mathbb{E}\big[X(t)^p\big]\,.$$

Letting $N \uparrow \infty$ and using monotone convergence gives the claim. ■

We now use the martingale property and the optional stopping theorem to prove Wald's lemmas for Brownian motion. These results identify the first and second moments of the value of Brownian motion at well-behaved stopping times.

Theorem 2.40 (Wald's lemma for Brownian motion). *Let $\{B(t) : t \geq 0\}$ be a standard linear Brownian motion, and T be a stopping time such that either*

(i) $\mathbb{E}[T] < \infty$, *or*
(ii) $\big\{B(t \wedge T) : t \geq 0\big\}$ *is L^1-bounded .*

Then we have $\mathbb{E}[B(T)] = 0$.

Remark 2.41. The proof of Wald's lemma is based on an optional stopping argument. An alternative proof of (i), which uses only the strong Markov property and the law of large numbers, is suggested in Exercise 2.5. Also, the moment condition (i) in Theorem 2.40 can be relaxed, see Theorem 2.46 for an optimal criterion. ◇

Proof. We first show that a stopping time satisfying condition (i), also satisfies condition (ii). So suppose $\mathbb{E}[T] < \infty$, and define

$$M_k = \max_{0\leq t\leq 1}|B(t+k) - B(k)| \quad \text{and } M = \sum_{k=1}^{\lceil T\rceil} M_k.$$

Then

$$\begin{aligned}\mathbb{E}[M] = \mathbb{E}\Big[\sum_{k=1}^{\lceil T\rceil} M_k\Big] &= \sum_{k=1}^{\infty}\mathbb{E}\big[\mathbb{1}\{T\geq k\}\,M_k\big] = \sum_{k=1}^{\infty}\mathbb{P}\{T\geq k\}\,\mathbb{E}[M_k]\\ &= \mathbb{E}[M_0]\,\mathbb{E}[T] < \infty\,,\end{aligned}$$

where, using Fubini's theorem and Remark 2.19,

$$\mathbb{E}[M_0] = \int_0^\infty \mathbb{P}\big\{\max_{0\leq t\leq 1}|B(t)| > x\big\}\,dx \leq \int_0^\infty \tfrac{2\sqrt{2}}{x\sqrt{\pi}}\,\exp\big\{-\tfrac{x^2}{2t}\big\} < \infty\,.$$

Now note that $|B(t\wedge T)| \leq M$, so that (ii) holds. It remains to observe that under condition (ii) we can apply the optional stopping theorem with $S = 0$, which yields $\mathbb{E}[B(T)] = 0$. ■

COROLLARY 2.42. *Let $S \leq T$ be stopping times and $\mathbb{E}[T] < \infty$. Then*
$$\mathbb{E}\big[(B(S))^2\big] = \mathbb{E}\big[(B(T))^2\big] + \mathbb{E}\big[\big(B(T) - B(S)\big)^2\big].$$

Proof. The tower property of conditional expectation gives
$$\mathbb{E}\big[\big(B(T)\big)^2\big] = \mathbb{E}\big[\big(B(S)\big)^2\big] + 2\mathbb{E}\big[B(S)\mathbb{E}\big[B(T) - B(S) \mid \mathcal{F}(S)\big] + \mathbb{E}\big[\big(B(T) - B(S)\big)^2\big].$$
Note that $\mathbb{E}[T] < \infty$ implies $\mathbb{E}[T - S \mid \mathcal{F}(S)] < \infty$ almost surely. Hence the strong Markov property at time S together with Wald's lemma imply $\mathbb{E}[B(T) - B(S) \mid \mathcal{F}(S)] = 0$ almost surely, so that the middle term vanishes. ■

To find the second moment of $B(T)$ and thus prove Wald's second lemma, we identify a further martingale derived from Brownian motion.

LEMMA 2.43. *Suppose $\{B(t) : t \geq 0\}$ is a linear Brownian motion. Then the process*
$$\big\{B(t)^2 - t \,:\, t \geq 0\big\}$$
is a martingale.

Proof. The process is adapted to the natural filtration of Brownian motion and
$$\begin{aligned}\mathbb{E}\big[B(t)^2 - t \,\big|\, \mathcal{F}^+(s)\big] &= \mathbb{E}\big[\big(B(t) - B(s)\big)^2 \,\big|\, \mathcal{F}^+(s)\big] + 2\,\mathbb{E}\big[B(t)B(s) \,\big|\, \mathcal{F}^+(s)\big] - B(s)^2 - t \\ &= (t - s) + 2B(s)^2 - B(s)^2 - t = B(s)^2 - s\,,\end{aligned}$$
which completes the proof. ■

THEOREM 2.44 (Wald's second lemma). *Let T be a stopping time for standard Brownian motion such that $\mathbb{E}[T] < \infty$. Then*
$$\mathbb{E}\big[B(T)^2\big] = \mathbb{E}[T].$$

Proof. Look at the martingale $\{B(t)^2 - t\colon t \geq 0\}$ and define stopping times
$$T_n = \inf\{t \geq 0\colon |B(t)| = n\}$$
so that $\{B(t \wedge T \wedge T_n)^2 - t \wedge T \wedge T_n \,:\, t \geq 0\}$ is dominated by the integrable random variable $n^2 + T$. By the optional stopping theorem we get $\mathbb{E}[B(T \wedge T_n)^2] = \mathbb{E}[T \wedge T_n]$.
By Corollary 2.42 we have $\mathbb{E}[B(T)^2] \geq \mathbb{E}[B(T \wedge T_n)^2]$. Hence, by monotone convergence,
$$\mathbb{E}\big[B(T)^2\big] \geq \lim_{n\to\infty} \mathbb{E}\big[B(T \wedge T_n)^2\big] = \lim_{n\to\infty} \mathbb{E}\big[T \wedge T_n\big] = \mathbb{E}[T]\,.$$
Conversely, now using Fatou's lemma in the first step,
$$\mathbb{E}\big[B(T)^2\big] \leq \liminf_{n\to\infty} \mathbb{E}\big[B(T \wedge T_n)^2\big] = \liminf_{n\to\infty} \mathbb{E}\big[T \wedge T_n\big] \leq \mathbb{E}[T]\,.$$ ■

Wald's lemmas suffice to obtain exit probabilities and expected exit times for a linear Brownian motion. In Chapter 3 we shall explore the corresponding problem for higher-dimensional Brownian motion using harmonic functions.

THEOREM 2.45. *Let $a < 0 < b$ and, for a standard linear Brownian motion $\{B(t) : t \geq 0\}$, define $T = \min\{t \geq 0 : B(t) \in \{a, b\}\}$. Then*

- $\mathbb{P}\{B(T) = a\} = \dfrac{b}{|a| + b}$ *and* $\mathbb{P}\{B(T) = b\} = \dfrac{|a|}{|a| + b}$.
- $\mathbb{E}[T] = |a|b$.

Proof. Let $T = \tau(\{a, b\})$ be the first exit time from the interval $[a, b]$. This stopping time satisfies the condition of the optional stopping theorem, as $|B(t \wedge T)| \leq |a| \vee b$. Hence, by Wald's first lemma,

$$0 = \mathbb{E}[B(T)] = a\mathbb{P}\{B(T) = a\} + b\mathbb{P}\{B(T) = b\}.$$

Together with the trivial equation $\mathbb{P}\{B(T) = a\} + \mathbb{P}\{B(T) = b\} = 1$ one can solve this, and obtain $\mathbb{P}\{B(T) = a\} = b/(|a| + b)$, and $\mathbb{P}\{B(T) = b\} = |a|/(|a| + b)$. To use Wald's second lemma, we check that $\mathbb{E}[T] < \infty$. For this purpose note that

$$\mathbb{E}[T] = \int_0^\infty \mathbb{P}\{T > t\}\, dt = \int_0^\infty \mathbb{P}\{B(s) \in (a, b) \text{ for all } s \in [0, t]\}\, dt \quad ,$$

and that, for $t \geq k \in \mathbb{N}$ the integrand is bounded by the k^{th} power of $\max_{x \in (a,b)} \mathbb{P}_x\{B(1) \in (a, b)\}$, i.e. decreases exponentially. Hence the integral is finite.

Now, by Wald's second lemma and the exit probabilities, we obtain

$$\mathbb{E}[T] = \mathbb{E}[B(T)^2] = \frac{a^2 b}{|a| + b} + \frac{b^2 |a|}{|a| + b} = |a|b. \qquad \blacksquare$$

We now discuss a strengtening of Theorem 2.40, which works with a weaker moment condition. This theorem will not be used in the remainder of the book and can be skipped on first reading. We shall see in Exercise 2.11 that the condition we give is in some sense optimal.

THEOREM* 2.46. *Let $\{B(t) \colon t \geq 0\}$ be a standard linear Brownian motion and T a stopping time with $\mathbb{E}[T^{1/2}] < \infty$. Then $\mathbb{E}[B(T)] = 0$.*

Proof. Let $\{M(t) \colon t \geq 0\}$ be the maximum process of $\{B(t) \colon t \geq 0\}$ and T a stopping time with $\mathbb{E}[T^{1/2}] < \infty$. Let $\tau = \lceil \log_4 T \rceil$, so that $B(t \wedge T) \leq M(4^\tau)$. In order to get $\mathbb{E}[B(T)] = 0$ from the optional stopping theorem it suffices to show that the majorant is integrable, i.e. that

$$\mathbb{E}M(4^\tau) < \infty.$$

Define a discrete time stochastic process $\{X_k \colon k \in \mathbb{N}\}$ by $X_k = M(4^k) - 2^{k+2}$, and observe that τ is a stopping time with respect to the natural filtration $(\mathcal{F}_k \colon k \in \mathbb{N})$ of this process. Moreover, the process is a supermartingale. Indeed,

$$\mathbb{E}\big[X_k \,\big|\, \mathcal{F}_{k-1}\big] \leq M(4^{k-1}) + \mathbb{E}\Big[\max_{0 \leq t \leq 4^k - 4^{k-1}} B(t)\Big] - 2^{k+2}.$$

and the supermartingale property follows as

$$\mathbb{E}\Big[\max_{0\le t\le 4^k-4^{k-1}} B(t)\Big] = \sqrt{4^k-4^{k-1}}\,\mathbb{E}\Big[\max_{0\le t\le 1} B(t)\Big] = 2\sqrt{4^k-4^{k-1}} \le 2^{k+2}-2^{k+1}\,.$$

Now let $t=4^\ell$ and use the supermartingale property for $\tau\wedge\ell$ to get

$$\mathbb{E}\big[M(4^\tau\wedge t)\big] = \mathbb{E}\big[X_{\tau\wedge\ell}\big] + \mathbb{E}\big[2^{\tau\wedge\ell+2}\big] \le \mathbb{E}[X_0] + 4\,\mathbb{E}\big[2^\tau\big]\,.$$

Note that $X_0 = M(1)-4$, which has finite expectation and, by our assumption on the moments of T, we have $\mathbb{E}[2^\tau]<\infty$. Thus, by monotone convergence,

$$\mathbb{E}\big[M(4^\tau)\big] = \lim_{t\uparrow\infty}\big[M(4^\tau\wedge t)\big] < \infty\,,$$

which completes the proof of the theorem. ■

Given a function $f\colon\mathbb{R}\to\mathbb{R}$ we were able, in Lemma 2.43, to subtract a suitable term from $f(B(t))$ to obtain a martingale. To get a feeling what we have to subtract in the general case, we look at the analogous problem for the simple random walk $\{S_n : n\in\mathbb{N}\}$. A straightforward calculation gives, for $f\colon\mathbb{Z}\to\mathbb{R}$,

$$\begin{aligned}\mathbb{E}\big[f(S_{n+1})\,\big|\,\sigma\{S_1,\dots,S_n\}\big] - f(S_n) &= \tfrac12\big(f(S_n+1)-2f(S_n)+f(S_n-1)\big)\\ &= \tfrac12\,\tilde\Delta f(S_n)\,,\end{aligned}$$

where $\tilde\Delta$ is the second difference operator

$$\tilde\Delta f(x) := f(x+1)-2f(x)+f(x-1)\,.$$

Hence

$$f(S_n) - \tfrac12\sum_{k=0}^{n-1}\tilde\Delta f(S_k)$$

defines a (discrete time) martingale. In the Brownian motion case, one would expect a similar result with $\tilde\Delta f$ replaced by its continuous analogue, the Laplacian

$$\Delta f(x) = \sum_{i=1}^d \frac{\partial^2 f}{\partial x_i^2}\,.$$

THEOREM 2.47. *Let $f\colon\mathbb{R}^d\to\mathbb{R}$ be twice continuously differentiable, and $\{B(t)\colon t\ge0\}$ be a d-dimensional Brownian motion. Further suppose that, for all $t>0$ and $x\in\mathbb{R}^d$, we have $\mathbb{E}_x|f(B(t))|<\infty$ and $\mathbb{E}_x\int_0^t|\Delta f(B(s))|\,ds<\infty$. Then the process $\{X(t) : t\ge0\}$ defined by*

$$X(t) = f(B(t)) - \tfrac12\int_0^t \Delta f(B(s))\,ds$$

is a martingale.

Proof. For any $0\le s<t$,

$$\mathbb{E}\big[X(t)\,\big|\,\mathcal{F}(s)\big] = \mathbb{E}_{B(s)}\big[f(B(t))\big] - \tfrac12\int_0^s \Delta f(B(u))\,du - \int_s^t \mathbb{E}_{B(s)}\big[\tfrac12\,\Delta f(B(u))\big]\,du\,.$$

Now, using integration by parts and $\frac{1}{2}\,\Delta\mathfrak{p}(t,x,y) = \frac{\partial}{\partial t}\mathfrak{p}(t,x,y)$, we find

$$\begin{aligned}\mathbb{E}_{B(s)}\big[\tfrac{1}{2}\,\Delta f(B(u))\big] &= \tfrac{1}{2}\int \mathfrak{p}(u-s,B(s),x)\,\Delta f(x)\,dx \\ &= \tfrac{1}{2}\int \Delta\mathfrak{p}(u-s,B(s),x)\,f(x)\,dx = \int \tfrac{\partial}{\partial u}\mathfrak{p}(u-s,B(s),x)\,f(x)\,dx\,,\end{aligned}$$

and hence

$$\begin{aligned}\int_s^t \mathbb{E}_{B(s)}\big[\tfrac{1}{2}\,\Delta f(B(u))\big]\,du &= \lim_{\varepsilon\downarrow 0}\int\Big[\int_{s+\varepsilon}^t \tfrac{\partial}{\partial u}\mathfrak{p}(u-s,B(s),x)\,du\Big]\,f(x)\,dx \\ &= \int \mathfrak{p}(t-s,B(s),x)\,f(x)\,dx - \lim_{\varepsilon\downarrow 0}\int \mathfrak{p}(\varepsilon,B(s),x)\,f(x)\,dx \\ &= \mathbb{E}_{B(s)}\big[f(B(t))\big] - f(B(s))\,,\end{aligned}$$

and this confirms the martingale property. ∎

Example 2.48. Using $f(x) = x^2$ in Theorem 2.47 yields the familiar martingale $\{B(t)^2 - t\colon t \geq 0\}$. Using $f(x) = x^3$ we obtain the martingale $\{B(t)^3 - 3\int_0^t B(s)\,ds\colon t \geq 0\}$ and not the familiar martingale $\{B(t)^3 - 3tB(t)\colon t \geq 0\}$. Of course, the difference $\{\int_0^t (B(t) - B(s))\,ds \,:\, t \geq 0\}$ is a martingale. ◇

The next lemma states a fundamental principle, which we will discuss further in Chapter 7, see in particular Theorem 7.17.

Corollary 2.49. *Suppose $f\colon \mathbb{R}^d \to \mathbb{R}$ satisfies $\Delta f(x) = 0$ and $\mathbb{E}_x|f(B(t))| < \infty$, for every $x \in \mathbb{R}^d$ and $t > 0$. Then the process $\{f(B(t)) \,:\, t \geq 0\}$ is a martingale.*

Example 2.50. The function $f\colon \mathbb{R}^2 \to \mathbb{R}$ given by $f(x_1,x_2) = e^{x_1}\cos x_2$ satisfies $\Delta f(x) = 0$. Hence $X(t) = e^{B_1(t)}\cos B_2(t)$ defines a martingale, where $\{B_1(t)\colon t \geq 0\}$ and $\{B_2(t)\colon t \geq 0\}$ are independent linear Brownian motions. ◇

Exercises

Exercise 2.1. Show that the definition of d-dimensional Brownian motion is invariant under an orthogonal change of coordinates.

Exercise 2.2. Show that, for any tail event $A \in \mathcal{T}$, the probability $\mathbb{P}_x(A)$ is independent of x. Also show that, for a germ event $A \in \mathcal{F}^+(0)$, the probability $\mathbb{P}_x(A)$ may depend on x.

Exercise 2.3 (*). For a linear Brownian motion, almost surely, every local maximum is a strict local maximum.

Hint. Use the Markov property to show that, given two disjoint closed time intervals, the maxima of Brownian motion on them are different almost surely. Then show that every local maximum of Brownian motion is a strict local maximum if this holds simultaneously for all pairs of disjoint rational intervals.

Exercise 2.4 (*).

(i) If $S \leq T$ are stopping times, then $\mathcal{F}^+(S) \subset \mathcal{F}^+(T)$.
(ii) If $T_n \downarrow T$ are stopping times, then $\mathcal{F}^+(T) = \bigcap_{n=1}^{\infty} \mathcal{F}^+(T_n)$.
(iii) If T is a stopping time, then the random variable $B(T)$ is $\mathcal{F}^+(T)$-measurable.

Exercise 2.5 (*). Let $\{B(t)\colon t \geq 0\}$ be a standard Brownian motion on the line, and T be a stopping time with $\mathbb{E}[T] < \infty$. Define an increasing sequence of stopping times by $T_1 = T$ and $T_n = T(B_n) + T_{n-1}$ where $T(B_n)$ is the same stopping time, but associated with the Brownian motion $\{B_n(t)\colon t \geq 0\}$ given by

$$B_n(t) = B(t + T_{n-1}) - B(T_{n-1}).$$

(a) Show that, almost surely,
$$\lim_{n\uparrow\infty} \frac{B(T_n)}{n} = 0.$$
(b) Show that $B(T)$ is integrable.
(c) Show that, almost surely,
$$\lim_{n\uparrow\infty} \frac{B(T_n)}{n} = \mathbb{E}\big[B(T)\big].$$

Combining (a) and (c) implies that $\mathbb{E}\big[B(T)\big] = 0$, which is Wald's lemma.

Exercise 2.6. Show that, for any $x > 0$ and measurable set $A \subset [0, \infty)$,

$$\mathbb{P}_x\big\{B(s) \geq 0 \text{ for all } 0 \leq s \leq t \text{ and } B(t) \in A\big\} = \mathbb{P}_x\{B(t) \in A\} - \mathbb{P}_{-x}\{B(t) \in A\}.$$

Exercise 2.7 ($*$). Show that any nonempty, closed set with no isolated points is uncountable. Note that this applies, in particular, to the zero set of (linear) Brownian motion.

Exercise 2.8. The *Ornstein-Uhlenbeck diffusion* is the process $\{X(t) : t \in \mathbb{R}\}$, given by

$$X(t) = e^{-t} B(e^{2t}) \text{ for all } t \in \mathbb{R},$$

see also Remark 1.10. Show that $\{X(t) \colon t \geq 0\}$ and $\{X(-t) \colon t \geq 0\}$ are Markov processes and find their Markov transition kernels.

Exercise 2.9. Let $x, y \in \mathbb{R}^d$ and $\{B(t) \colon t \geq 0\}$ a d-dimensional Brownian motion started in x. Define the d-dimensional *Brownian bridge* $\{X(t) \colon 0 \leq t \leq 1\}$ with start in x and end in y by

$$X(t) = B(t) - t\big(B(1) - y\big), \qquad \text{for } 0 \leq t \leq 1\,.$$

Show that the Brownian bridge is *not* a homogeneous Markov process.

Exercise 2.10. Find two stopping times $S \leq T$ with $\mathbb{E}[S] < \infty$ such that

$$\mathbb{E}[(B(S))^2] > \mathbb{E}[(B(T))^2].$$

Exercise 2.11 ($*$). The purpose of this exercise is to show that the moment condition in Theorem 2.46 is optimal. Let $\{B(t) \colon t \geq 0\}$ be a standard linear Brownian motion and define

$$T = \inf\{t \geq 0 \colon B(t) = 1\},$$

so that $B(T) = 1$ almost surely. Show that

$$\mathbb{E}[T^\alpha] < \infty \qquad \text{for all } \alpha < 1/2.$$

Exercise 2.12. Let $\{B(t) \colon t \geq 0\}$ be a standard linear Brownian motion

(a) Show that there exists a stopping time T with $\mathbb{E}T = \infty$ but $\mathbb{E}[(B(T))^2] < \infty$.
(b) Show that, for every stopping time T with $\mathbb{E}T = \infty$ and $\mathbb{E}\sqrt{T} < \infty$, we have

$$\mathbb{E}\big[B(T)^2\big] = \infty.$$

Exercise 2.13. Let $\{B(t)\colon t \geq 0\}$ be a linear Brownian motion.

(a) Show that, for $\sigma > 0$, the process $\{\exp(\sigma B(t) - \frac{\sigma^2 t}{2}) \,:\, t \geq 0\}$ is a martingale.

(b) By taking derivatives $\frac{\partial^n}{\partial \sigma^n}$ at zero, derive that the following processes are martingales.

- $\{B(t)^2 - t \,:\, t \geq 0\}$,
- $\{B(t)^3 - 3tB(t) \,:\, t \geq 0\}$, and
- $\{B(t)^4 - 6tB(t)^2 + 3t^2 \,:\, t \geq 0\}$.

(c) Find $\mathbb{E}[T^2]$ for $T = \min\{t \geq 0\colon B(t) \in \{a, b\}\}$ and $a < 0 < b$.

Exercise 2.14 (∗). Let $\{B(t)\colon t \geq 0\}$ be a linear Brownian motion and $a, b > 0$. Show that

$$\mathbb{P}_0\big\{B(t) = a + bt \text{ for some } t > 0\big\} = e^{-2ab}.$$

Exercise 2.15 (∗). Let $R > 0$ and $A = \{-R, R\}$. Denote by $\tau(A)$ the first hitting time of A, and by T_x the first hitting times of the point $x \in \mathbb{R}$. Consider a linear Brownian motion started at $x \in [0, R]$, and prove that

(a) $\mathbb{E}_x[\tau(A)] = R^2 - x^2$.

(b) $\mathbb{E}_x\big[T_R \,\big|\, T_R < T_0\big] = \frac{R^2 - x^2}{3}$.

Hint. In (b) use one of the martingales of Exercise 2.13(b).

Exercise 2.16. Let $\{B(t)\colon t \geq 0\}$ be a Brownian motion in dimension one.

(a) Use the optional stopping theorem for the martingale in Exercise 2.13(a) to show that, with $\tau_a = \inf\{t \geq 0 \,:\, B(t) = a\}$,

$$\mathbb{E}_0\big[e^{-\lambda \tau_a}\big] = e^{-a\sqrt{2\lambda}}, \qquad \text{for all } \lambda, a > 0\,.$$

(b) Use the reflection principle to show that, with $\tau_{-a} = \inf\{t \geq 0 \,:\, B(t) = -a\}$, we have

$$\mathbb{E}_0\big[e^{-\lambda \tau_a}\big] = \mathbb{E}_0\big[e^{-\lambda \tau_a}\, \mathbb{1}\{\tau_a < \tau_{-a}\}\big] + \mathbb{E}_0\big[e^{-\lambda \tau_{-a}}\, \mathbb{1}\{\tau_{-a} < \tau_a\}\big]\, e^{-2a\sqrt{2\lambda}}\,.$$

(c) Deduce that $\tau = \tau_a \wedge \tau_{-a}$ satisfies

$$\mathbb{E}_0\big[e^{-\lambda \tau}\big] = \operatorname{sech}(a\sqrt{2\lambda})\,,$$

where $\operatorname{sech}(x) = \frac{2}{e^x + e^{-x}}$.

Exercise 2.17. In this exercise we interpret $\mathbb{R}^2$ as the complex plane. Hence a planar Brownian motion becomes a complex Brownian motion. A complex-valued stochastic process is called a martingale, if its real and imaginary parts are martingales.

Let $\{B(t) : t \geq 0\}$ be a complex Brownian motion started in $\mathbf{i}$.

(a) Show that $\{e^{\mathbf{i}\lambda B(t)} : t \geq 0\}$ is a martingale, for any $\lambda \in \mathbb{R}$.

(b) Let T be the first time when $\{B(t) : t \geq 0\}$ hits the real axis. Using the optional stopping theorem at T, show that

$$\mathbb{E}\big[e^{\mathbf{i}\lambda B(T)}\big] = e^{-\lambda}.$$

Inverting the Fourier transform, the statement of (b) means that $B(T)$ is Cauchy distributed, a fact we already know from an explicit calculation, see Theorem 2.33.

Exercise 2.18 (∗). Let $f : \mathbb{R}^d \to \mathbb{R}$ be twice continuously differentiable, $\{B(t) : t \geq 0\}$ a d-dimensional Brownian motion such that $\mathbb{E}_x \int_0^t e^{-\lambda s} |f(B(s))|\, ds < \infty$ and $\mathbb{E}_x \int_0^t e^{-\lambda s} |\Delta f(B(s))|\, ds < \infty$, for any $x \in \mathbb{R}^d$ and $t > 0$.

(a) Show that the process $\{X(t) : t \geq 0\}$ defined by

$$X(t) = e^{-\lambda t} f(B(t)) - \int_0^t e^{-\lambda s} \big(\tfrac{1}{2} \Delta f(B(s)) - \lambda f(B(s))\big)\, ds$$

is a martingale.

(b) Suppose U is a bounded open set, $\lambda \geq 0$, and $u : U \to \mathbb{R}$ is a bounded solution of

$$\tfrac{1}{2}\Delta u(x) = \lambda\, u(x), \qquad \text{for } x \in U,$$

and $\lim_{x \to x_0} u(x) = f(x_0)$ for all $x_0 \in \partial U$. Show that,

$$u(x) = \mathbb{E}_x\Big[f(B(\tau))\, e^{-\lambda \tau}\Big],$$

where $\tau = \inf\{t \geq 0 : B(t) \notin U\}$.

Notes and Comments

The Markov property is central to any discussion of Brownian motion. The discussion of this chapter is only a small fraction of what has to be said, and the Markov property will be omnipresent in the rest of the book. The name goes back to Markov's paper [**Ma06**] where the Markovian dependence structure was introduced and a law of large numbers for dependent random variables was proved. The strong Markov property had been used for special stopping times, like hitting times of a point, since the 1930s. Hunt [**Hu56**] formalised the idea and gave rigorous proofs, and so did, independently, Dynkin [**Dy57**].

Zero-one laws are classics in probability theory. We have already encountered the powerful Hewitt-Savage law. Blumenthal's zero-one law was first proved in [**Bl57**].

The reflection principle is usually attributed to D. André [**An87**], who stated a variant for random walks. His concern was the ballot problem: if two candidates in a ballot receive a, respectively b votes, with $a > b$, what is the probability that the first candidate was always in the lead during the counting of the votes? See the classical text of Feller [**Fe68**] for more on this problem. A formulation of the reflection principle for Brownian motion was given by Lévy [**Le39**], though apparently not based on the rigorous foundation of the strong Markov property. We shall later use a higher-dimensional version of the reflection principle, where a Brownian motion in $\mathbb{R}^d$ is reflected in a hyperplane.

The class of Markov processes, defined in this chapter, has a rich and fascinating theory of its own, and some aspects are discussed in the books [**RW00**] and [**Ch82**]. A typical feature of this theory is its strong connection to analysis and potential theory, which stems from the key rôle played by the transition semigroup in their definition. This aspect is emphasised in the book [**BG68**]. Many of the important examples of Markov processes can be derived from Brownian motion in one way or the other, and this is an excellent motivation for further study of the theory. Amongst them are stable Lévy processes, like the Cauchy process or stable subordinators, the Bessel processes, and diffusions.

The intriguing relationship uncovered in Theorem 2.31 has found numerous extensions and complementary results, among them Pitman's $2M - X$ theorem, see [**Pi75**], which describes the process $\{2M(t) - B(t) : t \geq 0\}$ as a 3-dimensional Bessel process.

The concept of martingales is due to Doob, see [**Do53**]. They are an important class of stochastic processes in their own right and one of the gems of modern probability theory. A gentle introduction, mostly in discrete time, is [**Wi91**], while [**RY94**] discusses continuous martingales and the rich relations to Brownian motion. A fascinating fact, due to Dambis [**Da65**], Dubins, and Schwarz [**DS65**], is that for every continuous martingale there exists a time-change, i.e. a reparametrisation $t \mapsto T_t$ such that $T_t, t \geq 0$ are stopping times, such that $t \mapsto M(T_t)$ is a Brownian motion. Exercise 2.15 appears in similar form in [**St75**].

CHAPTER 3

Harmonic functions, transience and recurrence

In this chapter we explore the relation of harmonic functions and Brownian motion. This approach will be particularly useful for d-dimensional Brownian motion for $d > 1$. It allows us to study the fundamental questions of transience and recurrence of Brownian motion, investigate the classical Dirichlet problem of electrostatics, and provide the background for the deeper investigations of probabilistic potential theory, which will follow in Chapter 8.

1. Harmonic functions and the Dirichlet problem

Let U be a **domain**, i.e. a connected open set $U \subset \mathbb{R}^d$, and ∂U be its boundary. Suppose that its closure $\overline{U}$ is a homogeneous body and its boundary is electrically charged, the charge given by some continuous function $\varphi\colon \partial U \to \mathbb{R}$. The *Dirichlet problem* asks for the voltage $u(x)$ at some point $x \in U$. Kirchhoff's laws state that u must be a *harmonic function* in U. We therefore start by discussing the basic features of harmonic functions.

Definition 3.1. *Let $U \subset \mathbb{R}^d$ be a domain. A function $u\colon U \to \mathbb{R}$ is* harmonic *(on U) if it is twice continuously differentiable and, for any $x \in U$,*

$$\Delta u(x) := \sum_{j=1}^{d} \frac{\partial^2 u}{\partial x_j^2}(x) = 0.$$

If instead of the last condition only $\Delta u(x) \geq 0$, then the function u is called **subharmonic**. $\diamond$

To begin with we give two useful reformulations of the harmonicity condition, called the **mean value properties**, which do not make explicit reference to differentiability.

Theorem 3.2. *Let $U \subset \mathbb{R}^d$ be a domain and $u\colon U \to \mathbb{R}$ measurable and locally bounded. The following conditions are equivalent:*

(i) *u is harmonic;*
(ii) *for any ball $\mathcal{B}(x,r) \subset U$,*

$$u(x) = \frac{1}{\mathcal{L}(\mathcal{B}(x,r))} \int_{\mathcal{B}(x,r)} u(y)\, dy;$$

(iii) *for any ball $\mathcal{B}(x,r) \subset U$,*

$$u(x) = \frac{1}{\sigma_{x,r}(\partial\mathcal{B}(x,r))} \int_{\partial\mathcal{B}(x,r)} u(y)\, d\sigma_{x,r}(y),$$

where $\sigma_{x,r}$ is the surface measure on $\partial\mathcal{B}(x,r)$.

Remark 3.3. We use the following version of Green's identity,

$$\int_{\partial\mathcal{B}(x,r)} \frac{\partial u}{\partial n}(y)\, d\sigma_{x,r}(y) = \int_{\mathcal{B}(x,r)} \Delta u(y)\, dy, \tag{1.1}$$

where $n(y)$ is the outward normal vector of the ball at y. The result can also be proved by purely probabilistic means, see Exercise 8.1. $\diamond$

Proof. **(ii) $\Rightarrow$ (iii)** Assume u has the mean value property (ii). Define $\psi\colon (0,\infty) \to \mathbb{R}$ by

$$\psi(r) = r^{1-d} \int_{\partial\mathcal{B}(x,r)} u(y)\, d\sigma_{x,r}(y).$$

We show that ψ is constant. Indeed, for any $r > 0$,

$$r^d\, \mathcal{L}(\mathcal{B}(x,1))\, u(x) = \mathcal{L}(\mathcal{B}(x,r))\, u(x) = \int_{\mathcal{B}(x,r)} u(y)\, dy = \int_0^r \psi(s)\, s^{d-1}\, ds.$$

Differentiating with respect to r gives $d\mathcal{L}(\mathcal{B}(x,1))u(x) = \psi(r)$, and therefore $\psi(r)$ is constant. Now (iii) follows from the well known identity $d\mathcal{L}(\mathcal{B}(x,r))/r = \sigma_{x,r}(\partial\mathcal{B}(x,r))$.

(iii) $\Rightarrow$ (ii) Fix $s > 0$, multiply (ii) by $\sigma_{x,r}(\partial\mathcal{B}(x,r))$ and integrate over all radii $0 < r < s$.

(iii) $\Rightarrow$ (i) Suppose $g\colon [0,\infty) \to [0,\infty)$ is a smooth function with compact support in $[0,\varepsilon)$ and $\int g(|\,x|)\, dx = 1$. Integrating (iii) one obtains

$$u(x) = \int u(y) g(|\,x-y|\,)\, dy$$

for all $x \in U$ and sufficiently small $\varepsilon > 0$. As convolution of a smooth function with a bounded function produces a smooth function, we observe that u is infinitely often differentiable in U.

Now suppose that $\Delta u \neq 0$, so that there exists a small ball $\mathcal{B}(x,\varepsilon) \subset U$ such that either $\Delta u(x) > 0$ on $\mathcal{B}(x,\varepsilon)$, or $\Delta u(x) < 0$ on $\mathcal{B}(x,\varepsilon)$. Using the notation from above, we obtain that

$$0 = \psi'(r) = r^{1-d} \int_{\partial\mathcal{B}(x,r)} \frac{\partial u}{\partial n}(y)\, d\sigma_{x,r}(y) = r^{1-d} \int_{\mathcal{B}(x,r)} \Delta u(y)\, dy,$$

using (1.1). This is a contradiction.

(i) $\Rightarrow$ (iii) Suppose that u is harmonic and $\mathcal{B}(x,r) \subset U$. With the notation from above and (1.1), we obtain that

$$\psi'(r) = r^{1-d} \int_{\partial\mathcal{B}(x,r)} \frac{\partial u}{\partial n}(y)\, d\sigma_{x,r}(y) = r^{1-d} \int_{\mathcal{B}(x,r)} \Delta u(y)\, dy = 0.$$

Hence ψ is constant, and as $\lim_{r\downarrow 0} \psi(r) = \sigma_{0,1}(\mathcal{B}(0,1))\, u(x)$, we obtain (iii). $\blacksquare$

Remark 3.4. A twice differentiable function $u\colon U \to \mathbb{R}$ is subharmonic if and only if

$$u(x) \le \frac{1}{\mathcal{L}(\mathcal{B}(x,r))} \int_{\mathcal{B}(x,r)} u(y)\,dy \qquad \text{for any ball } \mathcal{B}(x,r) \subset U. \tag{1.2}$$

This can be obtained in a way very similar to Theorem 3.2, see also Exercise 3.1. $\diamond$

An important property satisfied by harmonic, and in fact subharmonic, functions is the maximum principle. This is one of the key principles of analysis.

Theorem 3.5 (Maximum principle). *Suppose $u\colon \mathbb{R}^d \to \mathbb{R}$ is a function, which is subharmonic on an open connected set $U \subset \mathbb{R}^d$.*

(i) *If u attains its maximum in U, then u is a constant.*
(ii) *If u is continuous on $\bar{U}$ and U is bounded, then*

$$\max_{x\in\bar{U}} u(x) = \max_{x\in\partial U} u(x).$$

Remark 3.6. If u is harmonic, the theorem may be applied to both u and $-u$. Hence the conclusions of the theorem also hold with 'maximum' replaced by 'minimum'. $\diamond$

Proof. (i) Let M be the maximum. Note that $V = \{x \in U : u(x) = M\}$ is relatively closed in U. Since U is open, for any $x \in V$, there is a ball $\mathcal{B}(x,r) \subset U$. By the mean-value property of u, see Remark 3.4,

$$M = u(x) \le \frac{1}{\mathcal{L}(\mathcal{B}(x,r))} \int_{\mathcal{B}(x,r)} u(y)\,dy \le M.$$

Equality holds everywhere, and as $u(y) \le M$ for all $y \in \mathcal{B}(x,r)$, we infer that $u(y) = M$ almost everywhere on $\mathcal{B}(x,r)$. By continuity this implies $\mathcal{B}(x,r) \subset V$. Hence V is also open, and by assumption nonempty. Since U is connected we get that $V = U$. Therefore, u is constant on U.

(ii) Since u is continuous and $\bar{U}$ is closed and bounded, u attains a maximum on $\bar{U}$. By (i) the maximum has to be attained on ∂U. ■

Corollary 3.7. *Suppose $u_1, u_2\colon \mathbb{R}^d \to \mathbb{R}$ are functions, which are harmonic on a bounded domain $U \subset \mathbb{R}^d$ and continuous on $\bar{U}$. If u_1 and u_2 agree on ∂U, then they are identical.*

Proof. By Theorem 3.5(ii) applied to $u_1 - u_2$ we obtain that

$$\sup_{x\in\bar{U}} \big\{u_1(x) - u_2(x)\big\} = \sup_{x\in\partial U} \big\{u_1(x) - u_2(x)\big\} = 0.$$

Hence $u_1(x) \le u_2(x)$ for all $x \in \bar{U}$. Applying the same argument to $u_2 - u_1$, one sees that $\sup_{x\in\bar{U}}\{u_2(x) - u_1(x)\} = 0$. Hence $u_1(x) = u_2(x)$ for all $x \in \bar{U}$. ■

We can now formulate the basic fact on which the relationship of Brownian motion and harmonic functions rests.

THEOREM 3.8. *Suppose U is a domain, $\{B(t) : t \geq 0\}$ a Brownian motion started inside U and $\tau = \tau(\partial U) = \min\{t \geq 0 : B(t) \in \partial U\}$ the first hitting time of its boundary. Let $\varphi \colon \partial U \to \mathbb{R}$ be measurable, and such that the function $u \colon U \to \mathbb{R}$ with*

$$u(x) = \mathbb{E}_x\big[\varphi(B(\tau))\, \mathbb{1}\{\tau < \infty\}\big], \qquad \text{for every } x \in U, \tag{1.3}$$

is locally bounded. Then u is a harmonic function.

Proof. The proof uses only the strong Markov property of Brownian motion and the mean value characterisation of harmonic functions. For a ball $\mathcal{B}(x,\delta) \subset U$ let $\tilde{\tau} = \inf\{t > 0 \colon B(t) \notin \mathcal{B}(x,\delta)\}$, then the strong Markov property implies that

$$u(x) = \mathbb{E}_x\big[\mathbb{E}_x\big[\varphi(B(\tau))\, \mathbb{1}\{\tau < \infty\} \,\big|\, \mathcal{F}^+(\tilde{\tau})\big]\big] = \mathbb{E}_x\big[u\big(B(\tilde{\tau})\big)\big] = \int_{\partial \mathcal{B}(x,r)} u(y)\, \varpi_{x,\delta}(dy),$$

where $\varpi_{x,\delta}$ is the uniform distribution on the sphere $\partial\mathcal{B}(x,\delta)$. Therefore, u has the mean value property and hence it is harmonic on U by Theorem 3.2. ■

DEFINITION 3.9. *Let U be a domain in $\mathbb{R}^d$ and let ∂U be its boundary. Suppose $\varphi \colon \partial U \to \mathbb{R}$ is a continuous function on its boundary. A continuous function $v \colon \overline{U} \to \mathbb{R}$ is a* **solution to the Dirichlet problem** *with boundary value φ, if it is harmonic on U and $v(x) = \varphi(x)$ for $x \in \partial U$.* ◇

The Dirichlet problem was posed by Gauss in 1840. In fact Gauss thought he showed that there is always a solution, but his reasoning was wrong and Zaremba in 1911 and Lebesgue in 1924 gave counterexamples. However, if the domain is sufficiently nice there is a solution, as we will see below.

DEFINITION 3.10. *Let $U \subset \mathbb{R}^d$ be a domain. We say that U satisfies the* **Poincaré cone condition** *at $x \in \partial U$ if there exists a cone V based at x with opening angle $\alpha > 0$, and $h > 0$ such that $V \cap \mathcal{B}(x,h) \subset U^{\mathrm{c}}$.* ◇

The following lemma, which is illustrated by Figure 1, will prepare us to solve the Dirichlet problem for 'nice' domains. Recall that we denote, for any open or closed set $A \subset \mathbb{R}^d$, by $\tau(A)$ the first hitting time of the set A by Brownian motion, $\tau(A) = \inf\{t \geq 0 \colon B(t) \in A\}$.

LEMMA 3.11. *Let $0 < \alpha < 2\pi$ and $C_0(\alpha) \subset \mathbb{R}^d$ is a cone based at the origin with opening angle α, and*

$$a = \sup_{x \in \mathcal{B}(0,\frac{1}{2})} \mathbb{P}_x\big\{\tau(\partial\mathcal{B}(0,1)) < \tau(C_0(\alpha))\big\}.$$

Then $a < 1$ and, for any positive integer k and $h' > 0$, we have

$$\mathbb{P}_x\big\{\tau(\partial\mathcal{B}(z,h')) < \tau(C_z(\alpha))\big\} \leq a^k,$$

for all $x, z \in \mathbb{R}^d$ with $|x - z| < 2^{-k}h'$, where $C_z(\alpha)$ is a cone based at z with opening angle α.

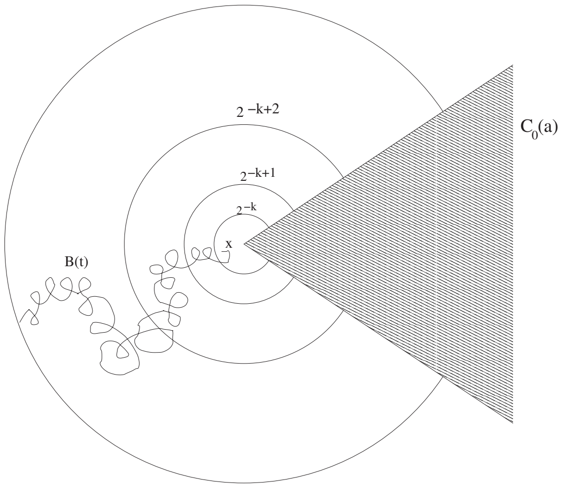

FIGURE 1. Brownian motion avoiding a cone

Proof. Obviously $a < 1$. If $x \in \mathcal{B}(0, 2^{-k})$ then by the strong Markov property

$$\mathbb{P}_x\big\{\tau(\partial\mathcal{B}(0,1)) < \tau(C_0(\alpha))\big\}$$

$$\le \prod_{i=0}^{k-1} \sup_{x \in \mathcal{B}(0,2^{-k+i})} \mathbb{P}_x\big\{\tau(\partial\mathcal{B}(0,2^{-k+i+1})) < \tau(C_0(\alpha))\big\} = a^k.$$

Therefore, for any positive integer k and $h' > 0$, we have by scaling $\mathbb{P}_x\big\{\tau(\partial\mathcal{B}(z,h')) < \tau(C_z(\alpha))\big\} \le a^k$, for all x with $|\,x - z| < 2^{-k}h'$. ■

THEOREM 3.12 (Dirichlet Problem). *Suppose $U \subset \mathbb{R}^d$ is a bounded domain such that every boundary point satisfies the Poincaré cone condition, and suppose φ is a continuous function on ∂U. Let $\tau(\partial U) = \inf\{t > 0 \colon B(t) \in \partial U\}$. Then the function $u \colon \overline{U} \to \mathbb{R}$ given by*

$$u(x) = \mathbb{E}_x\big[\varphi(B(\tau(\partial U)))\big], \qquad \text{for } x \in \overline{U},$$

is the unique continuous function harmonic on U with $u(x) = \varphi(x)$ for all $x \in \partial U$.

Proof. The uniqueness claim follows from Corollary 3.7. The function u is bounded and hence harmonic on U by Theorem 3.8. It remains to show that the Poincaré cone condition implies the boundary condition. Fix $z \in \partial U$, then there is a cone $C_z(\alpha)$ based at z with angle $\alpha > 0$ with $C_z(\alpha) \cap \mathcal{B}(z,h) \subset U^c$. By Lemma 3.11, for any positive integer k and $h' > 0$, we have

$$\mathbb{P}_x\big\{\tau(\partial\mathcal{B}(z,h')) < \tau(C_z(\alpha))\big\} \le a^k$$

for all x with $|\,x - z| < 2^{-k}h'$. Given $\varepsilon > 0$, there is a $0 < \delta \le h$ such that $|\varphi(y) - \varphi(z)| < \varepsilon$ for all $y \in \partial U$ with $|y - z| < \delta$. For all $x \in \overline{U}$ with $|z - x| < 2^{-k}\delta$,

$$|u(x) - u(z)| = \big|\mathbb{E}_x\varphi(B(\tau(\partial U))) - \varphi(z)\big| \le \mathbb{E}_x\big|\varphi(B(\tau(\partial U))) - \varphi(z)\big|. \tag{1.4}$$

If the Brownian motion hits the cone $C_z(\alpha)$, which is outside the domain U, before the sphere $\partial\mathcal{B}(z,\delta)$, then $|z - B(\tau(\partial U))| < \delta$, and $\varphi(B(\tau(\partial U)))$ is close to $\varphi(z)$. The complement has small probability. More precisely, (1.4) is bounded above by

$$2\|\varphi\|_\infty \mathbb{P}_x\big\{\tau(\partial\mathcal{B}(z,\delta)) < \tau(C_z(\alpha))\big\} + \varepsilon \mathbb{P}_x\{\tau(\partial U) < \tau(\partial\mathcal{B}(z,\delta))\} \le 2\|\varphi\|_\infty a^k + \varepsilon.$$

This implies that u is continuous on $\overline{U}$. ■

Remark 3.13. If the Poincaré cone condition holds at every boundary point, we can simulate the solution of the Dirichlet problem by running many independent Brownian motions, starting in $x \in U$ until they hit the boundary of U and letting $u(x)$ be the average of the values of φ on the hitting points. ◇

Remark 3.14. In Chapter 8 we will improve the results on the Dirichlet problem significantly and give sharp criteria for the existence of solutions. ◇

To justify the introduction of conditions on the domain we now give an example where the function u of Theorem 3.12 fails to solve the Dirichlet problem.

Example 3.15. Take a solution $v\colon \mathcal{B}(0,1) \to \mathbb{R}$ of the Dirichlet problem on the planar disc $\mathcal{B}(0,1)$ with boundary condition $\varphi\colon \partial\mathcal{B}(0,1) \to \mathbb{R}$. Let $U = \{x \in \mathbb{R}^2 : 0 < |x| < 1\}$ be the punctured disc. We claim that $u(x) = \mathbb{E}_x\big[\varphi(B(\tau(\partial U)))\big]$ fails to solve the Dirichlet problem on U with boundary condition $\varphi\colon \partial\mathcal{B}(0,1) \cup \{0\} \to \mathbb{R}$ if $\varphi(0) \neq v(0)$. Indeed, as planar Brownian motion does not hit points, by Corollary 2.23, the first hitting time τ of $\partial U = \partial\mathcal{B}(0,1) \cup \{0\}$ agrees almost surely with the first hitting time of $\partial\mathcal{B}(0,1)$. Then, by Theorem 3.12, $u(0) = \mathbb{E}_0[\varphi(B(\tau))] = v(0) \neq \varphi(0)$. ◇

We now show how the techniques we have developed so far can be used to prove a classical result from harmonic analysis, Liouville's theorem, by probabilistic means. The proof uses the reflection principle for higher-dimensional Brownian motion.

Theorem 3.16 (Liouville's theorem). *Any bounded harmonic function on $\mathbb{R}^d$ is constant.*

Proof. Let $u\colon \mathbb{R}^d \to [-M, M]$ be a harmonic function, x, y two distinct points in $\mathbb{R}^d$, and H the hyperplane so that the reflection in H takes x to y. Let $\{B(t) : t \ge 0\}$ be Brownian motion started at x, and $\{\overline{B}(t)\colon t \ge 0\}$ its reflection in H. Let $\tau(H) = \min\{t\colon B(t) \in H\}$ and note that

$$\{B(t)\colon t \ge \tau(H)\} \overset{\mathrm{d}}{=} \{\overline{B}(t)\colon t \ge \tau(H)\}. \tag{1.5}$$

Harmonicity implies that $\mathbb{E}_x[u(B(t))] = u(x)$ and decomposing the above into $t < \tau(H)$ and $t \ge \tau(H)$ we get

$$u(x) = \mathbb{E}_x\big[u(B(t))\mathbb{1}_{\{t<\tau(H)\}}\big] + \mathbb{E}_x\big[u(B(t))\mathbb{1}_{\{t\ge\tau(H)\}}\big].$$

A similar equality holds for $u(y)$. Now, using (1.5),

$$|u(x) - u(y)| = \big|\mathbb{E}\big[u(B(t))\mathbb{1}_{\{t<\tau(H)\}}\big] - \mathbb{E}\big[u(\overline{B}(t))\mathbb{1}_{\{t<\tau(H)\}}\big]\big| \le 2M\mathbb{P}\{t < \tau(H)\} \to 0,$$

as $t \to \infty$. Thus $u(x) = u(y)$, and since x and y were chosen arbitrarily, u must be constant. ■

2. Recurrence and transience of Brownian motion

A Brownian motion $\{B(t)\colon t \ge 0\}$ in dimension d is called *transient* if

$$\lim_{t\uparrow\infty} |B(t)| = \infty \qquad \text{almost surely.}$$

Note that the event $\{\lim_{t\uparrow\infty} |B(t)| = \infty\}$ is a tail event and hence, by Kolmogorov's zero-one law, it must have probability zero or one. In this section we decide in which dimensions d the Brownian motion is transient, and in which it is not. This question is intimately related to the exit probabilities of the Brownian motion from an annulus: Suppose the motion starts at a point x inside an annulus

$$A = \{x \in \mathbb{R}^d \,:\, r \le |\,x| \le R\}, \qquad \text{for } 0 < r < R < \infty.$$

What is the probability that the Brownian motion hits $\partial\mathcal{B}(0,r)$ before $\partial\mathcal{B}(0,R)$? The answer is given in terms of harmonic functions on the annulus and is therefore closely related to the Dirichlet problem.

To find explicit solutions $u\colon \bar{A} \to \mathbb{R}$ of the Dirichlet problem on an annulus it is first reasonable to assume that u is spherically symmetric, i.e. there is a function $\psi\colon [r,R] \to \mathbb{R}$ such that $u(x) = \psi(|\,x|^2)$. We can express derivatives of u in terms of ψ as

$$\partial_i \psi(|\,x|^2) = \psi'(|\,x|^2)2x_i \text{ and } \partial_{ii}\psi(|\,x|^2) = \psi''(|\,x|^2)4x_i^2 + 2\psi'(|\,x|^2).$$

Therefore, $\Delta u = 0$ means

$$0 = \sum_{i=1}^d \Big(\psi''(|\,x|^2)4x_i^2 + 2\psi'(|\,x|^2)\Big) = 4|\,x|^2\psi''(|\,x|^2) + 2d\psi'(|\,x|^2).$$

Letting $y = |\,x|^2 > 0$ we can write this as

$$\psi''(y) = \frac{-d}{2y}\psi'(y).$$

This is solved by every ψ satisfying $\psi'(y) = y^{-d/2}$ and thus $\Delta u = 0$ holds on $\{|\,x| \ne 0\}$ for

$$(2.1) \qquad u(x) = \begin{cases} |\,x| & \text{if } d = 1, \\ 2\log|\,x| & \text{if } d = 2, \\ |\,x|^{2-d} & \text{if } d \ge 3. \end{cases}$$

We write $u(r)$ for the value of $u(x)$ for all $x \in \partial\mathcal{B}(0,r)$. Now define stopping times

$$T_r = \tau(\partial\mathcal{B}(0,r)) = \inf\{t > 0 \,:\, |B(t)| = r\} \text{ for } r > 0,$$

and denote by $T = T_r \wedge T_R$ the first exit time from A. By Theorem 3.12 we have

$$u(x) = \mathbb{E}_x\big[u(B(T))\big] = u(r)\mathbb{P}_x\{T_r < T_R\} + u(R)(1 - \mathbb{P}_x\{T_r < T_R\}).$$

This formula can be solved

$$\mathbb{P}_x\{T_r < T_R\} = \frac{u(R) - u(x)}{u(R) - u(r)}$$

and we get an explicit solution for the exit problem.

Theorem 3.17. *Suppose $\{B(t) : t \geq 0\}$ is a Brownian motion in dimension $d \geq 1$ started in*

$$x \in A := \{x \in \mathbb{R}^d : r \leq |x| \leq R\}$$

inside an annulus A with radii $0 < r < R < \infty$. Then,

$$\mathbb{P}_x\{T_r < T_R\} = \begin{cases} \frac{R-|x|}{R-r} & \text{if } d = 1, \\ \frac{\log R - \log|x|}{\log R - \log r} & \text{if } d = 2, \\ \frac{R^{2-d}-|x|^{2-d}}{R^{2-d}-r^{2-d}} & \text{if } d \geq 3. \end{cases}$$

Letting $R \uparrow \infty$ in Theorem 3.17 leads to the following corollary.

Corollary 3.18. *For any $x \notin \mathcal{B}(0, r)$, we have*

$$\mathbb{P}_x\{T_r < \infty\} = \begin{cases} 1 & \text{if } d \leq 2, \\ \frac{r^{d-2}}{|x|^{d-2}} & \text{if } d \geq 3. \end{cases}$$

We now apply this to the problem of *recurrence* and *transience* of Brownian motion in various dimensions. Generally speaking, we call a Markov process $\{X(t) : t \geq 0\}$ with values in $\mathbb{R}^d$

- **point recurrent**, if for every $x \in \mathbb{R}^d$, almost surely, there is a (random) sequence $t_n \uparrow \infty$ such that $X(t_n) = x$ for all $n \in \mathbb{N}$,
- **neighbourhood recurrent**, if, for every $x \in \mathbb{R}^d$ and $\varepsilon > 0$, almost surely, there exists a (random) sequence $t_n \uparrow \infty$ such that $X(t_n) \in \mathcal{B}(x, \varepsilon)$ for all $n \in \mathbb{N}$.
- **transient**, if it converges to infinity almost surely.

Theorem 3.19. *Brownian motion is*

- *point recurrent in dimension $d = 1$,*
- *neighbourhood recurrent, but not point recurrent, in $d = 2$,*
- *transient in dimension $d \geq 3$.*

Proof. We leave the case $d = 1$ as Exercise 3.3, and look at dimension $d = 2$. Fix $\varepsilon > 0$ and $x \in \mathbb{R}^d$. By Corollary 3.18 and shift-invariance the stopping time $t_1 = \inf\{t > 0 : B(t) \in \mathcal{B}(x, \varepsilon)\}$ is almost surely finite. Using the strong Markov property at time $t_1 + 1$ we see that this also applies to $t_2 = \inf\{t > t_1 + 1 : B(t) \in \mathcal{B}(x, \varepsilon)\}$, and continuing like this, we obtain a sequence of times $t_n \uparrow \infty$ such that, almost surely, $B(t_n) \in \mathcal{B}(x, \varepsilon)$ for all $n \in \mathbb{N}$. Taking a union over a countable family of small balls, which form a basis of the Eulidean topology, implies that in $d = 2$ Brownian motion is neighbourhood recurrent. Recall from Corollary 2.23 that planar Brownian motion does not hit points, hence it cannot be point recurrent.

It remains to show that Brownian motion is transient in dimensions $d \geq 3$. Look at the events $A_n := \{|B(t)| > n \text{ for all } t \geq T_{n^3}\}$. Recall from Proposition 1.23 that $T_{n^3} < \infty$ almost surely. By the strong Markov property, for every $n \geq |x|^{1/3}$,

$$\mathbb{P}_x(A_n^{\mathrm{c}}) = \mathbb{E}_x\Big[\mathbb{P}_{B(T_{n^3})}\{T_n < \infty\}\Big] = \Big(\frac{1}{n^2}\Big)^{d-2}.$$

Note that the right hand side is summable, and hence the Borel-Cantelli lemma shows that only finitely many of the events A_n^c occur, which implies that $|B(t)|$ diverges to infinity, almost surely, and hence that Brownian motion in $d \geq 3$ is transient. ∎

Remark 3.20. Neighbourhood recurrence, in particular, implies that the path of a planar Brownian motion (running for an infinite amount of time) is dense in the plane. ◇

We now have a qualitative look at the transience of Brownian motion in $\mathbb{R}^d$, $d \geq 3$, and ask for the speed of escape to infinity. This material is slightly more advanced and can be skipped on first reading.

Consider a standard Brownian motion $\{B(t) \colon t \geq 0\}$ in $\mathbb{R}^d$, for $d \geq 3$, and fix a sequence $t_n \uparrow \infty$. For any $\varepsilon > 0$, by Fatou's lemma,

$$\mathbb{P}\big\{\big|B(t_n)\big| < \varepsilon\sqrt{t_n} \text{ infinitely often } \big\} \geq \limsup_{n\to\infty} \mathbb{P}\big\{\big|B(t_n)\big| < \varepsilon\sqrt{t_n}\big\} > 0.$$

By the Hewitt Savage zero-one law, the probability on the left-hand side must therefore be one, whence

$$\liminf_{n\to\infty} \frac{|B(t_n)|}{\sqrt{t_n}} = 0, \qquad \text{almost surely.} \tag{2.2}$$

This statement is refined by the Dvoretzky-Erdős test.

Theorem* 3.21 (Dvoretzky-Erdős test). *Let $\{B(t) \colon t \geq 0\}$ be Brownian motion in $\mathbb{R}^d$ for $d \geq 3$ and $f \colon (0,\infty) \to (0,\infty)$ increasing. Then*

$$\int_1^\infty f(r)^{d-2} r^{-d/2}\, dr < \infty \qquad \textit{if and only if} \qquad \liminf_{t\uparrow\infty} \frac{|B(t)|}{f(t)} = \infty \textit{ almost surely.}$$

Conversely, if the integral diverges, then $\liminf_{t\uparrow\infty} |B(t)|/f(t) = 0$ almost surely.

For the proof we first recall two generally useful tools. The first is an easy case of the Paley-Zygmund inequality, see Exercise 3.4 for the full statement.

Lemma 3.22 (Paley-Zygmund inequality). *For any nonnegative random variable X with $\mathbb{E}[X^2] < \infty$,*

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

Proof. The Cauchy-Schwarz inequality gives

$$\mathbb{E}[X] = \mathbb{E}[X\, \mathbb{1}\{X > 0\}] \leq \mathbb{E}[X^2]^{1/2} \big(\mathbb{P}\{X > 0\}\big)^{1/2},$$

and the required inequality follows immediately. ∎

The second tool is a version of the Borel-Cantelli lemma, which allows some dependence of the events. This is known as the Kochen-Stone lemma, and is a consequence of the Paley-Zygmund inequality, see Exercise 3.5 or [**FG97**].

LEMMA 3.23. *Suppose $E_1, E_2, \dots$ are events with*
$$\sum_{n=1}^{\infty} \mathbb{P}(E_n) = \infty \qquad \text{and} \qquad \liminf_{k\to\infty} \frac{\sum_{m=1}^{k}\sum_{n=1}^{k} \mathbb{P}(E_n \cap E_m)}{\big(\sum_{n=1}^{k} \mathbb{P}(E_n)\big)^2} < \infty\,.$$
Then, with positive probability, infinitely many of the events take place.

A core estimate in the proof of the Dvoretzky-Erdős test is the following lemma, which is based on the hitting probabilities of the previous paragraphs.

LEMMA 3.24. *There exists a constant $C_1 > 0$ depending only on the dimension d such that, for any $\rho > 0$, we have*
$$\sup_{x\in\mathbb{R}^d} \mathbb{P}_x\big\{ \text{ there exists } t > 1 \text{ with } |B(t)| \le \rho \big\} \le C_1\, \rho^{d-2}\,.$$

Proof. We use Corollary 3.18 for the probability that the motion started at time one hits $\mathcal{B}(0,\rho)$, to see that
$$\begin{aligned} \mathbb{P}_x\big\{ \text{ there exists } t > 1 \text{ with } |B(t)| \le \rho \big\} &\le \mathbb{E}_0\Big[\Big(\frac{\rho}{|B(1)+x|}\Big)^{d-2}\Big] \\ &\le \rho^{d-2}\,\frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} |y+x|^{2-d}\ \exp\big\{-\tfrac{|y|^2}{2}\big\}\, dy. \end{aligned}$$
By considering the integration domains $|y+x| \ge |y|$ and $|y+x| \le |y|$ separately, it is easy to see that the integral on the right is uniformly bounded in x. ■

Proof of Theorem 3.21. Define events
$$A_n = \big\{ \text{ there exists } t \in (2^n, 2^{n+1}] \text{ with } |B(t)| \le f(t) \big\}\,.$$
By Brownian scaling, monotonicity of f, and Lemma 3.24,
$$\mathbb{P}(A_n) \le \mathbb{P}\big\{ \text{ there exists } t > 1 \text{ with } |B(t)| \le f(2^{n+1})2^{-n/2} \big\} \le C_1 \Big(f(2^{n+1})\, 2^{-n/2}\Big)^{d-2}\,.$$
Now assume that the integral converges, or equivalently, that

$$\sum_{n=1}^{\infty}\Big(f(2^n)\,2^{-n/2}\Big)^{d-2} < \infty\,. \tag{2.3}$$

Then the Borel Cantelli lemma and (2.3) imply that, almost surely, the set $\{t > 0 : |B(t)| \le f(t)\}$ is bounded. Since (2.3) also applies to any constant multiple of f in place of f, it follows that $\liminf_{t\uparrow\infty} |B(t)|/f(t) = \infty$ almost surely.

For the converse, suppose that the integral diverges, whence

$$\sum_{n=1}^{\infty}\Big(f(2^n)\,2^{-n/2}\Big)^{d-2} = \infty\,. \tag{2.4}$$

In view of (2.2), we may assume that $f(t) < \sqrt{t}$ for all large enough t. Changing f on a finite interval, we may assume that this inequality holds for all $t > 0$.

For $\rho \in (0,1)$, consider the random variable $I_\rho = \int_1^2 \mathbb{1}\{|B(t)| \le \rho\}$. Since the density of $|B(t)|$ on the unit ball is bounded from above and also away from zero for $t \in [1,2]$, we infer that

$$C_2 \rho^d \le \mathbb{E}[I_\rho] \le C_3 \rho^d$$

for suitable constants depending only on the dimension. To complement this by an estimate of the second moment, we use the Markov property to see that

$$\begin{aligned}\mathbb{E}[I_\rho^2] &= 2\mathbb{E}\Big[\int_1^2 \mathbb{1}\{|B(t)| \le \rho\} \int_t^2 \mathbb{1}\{|B(s)| \le \rho\}\, ds\, dt\Big] \\ &\le 2\mathbb{E}\Big[\int_1^2 \mathbb{1}\{|B(t)| \le \rho\}\, \mathbb{E}_{B(t)} \int_0^\infty \mathbb{1}\{|\tilde{B}(s)| \le \rho\}\, ds\, dt\Big],\end{aligned}$$

where the inner expectation is with respect to a Brownian motion $\{\tilde{B}(t) \colon t \ge 0\}$ started in the fixed point $B(t)$, whereas the outer expectation is with respect to $B(t)$. We analyse the dependence of the inner expectation on the starting point. Given $x \ne 0$, we let $T = \inf\{t > 0 \colon |B(t)| = x\}$ and use the strong Markov property to see that

$$\mathbb{E}_0 \int_0^\infty \mathbb{1}\{|B(s)| \le \rho\}\, ds \ge \mathbb{E} \int_T^\infty \mathbb{1}\{|B(s)| \le \rho\}\, ds = \mathbb{E}_x \int_0^\infty \mathbb{1}\{|B(s)| \le \rho\}\, ds,$$

so that the expectation is maximal if the process is started at the origin. Hence we obtain that

$$\mathbb{E}[I_\rho^2] \le 2C_3\, \rho^d\, \mathbb{E}_0 \int_0^\infty \mathbb{1}\{|B(s)| \le \rho\}\, ds\,.$$

Moreover, by Brownian scaling, we obtain

$$\begin{aligned}\mathbb{E}_0 \int_0^\infty \mathbb{1}\{|B(s)| \le \rho\}\, ds &= \rho^2 \int_0^\infty \mathbb{P}\{|B(s)| \le 1\}\, ds \\ &= \rho^2 \int_{\mathcal{B}(0,1)} dx \int_0^\infty \frac{1}{(2\pi s)^{d/2}} \exp\big\{-x^2/s\big\}\, ds = C_4\, \rho^2,\end{aligned}$$

where the finiteness of the constant C_4 is easily checked by substituting sx^2 for s in the inner integral. In summary, we have $\mathbb{E}[I_\rho^2] \le 2C_3C_4\, \rho^{d+2}$. By the Paley-Zygmund inequality, for a suitable constant $C_5 > 0$,

$$\mathbb{P}\{I_\rho > 0\} \ge \frac{\mathbb{E}[I_\rho]^2}{\mathbb{E}[I_\rho^2]} \ge C_5 \rho^{d-2}\,.$$

Now choose $\rho = f(2^n)2^{-n/2}$, which is smaller than one, as $f(t) < \sqrt{t}$. By Brownian scaling and monotonicity of f, we have

$$\mathbb{P}(A_n) \ge \mathbb{P}\{I_\rho > 0\} \ge C_5 \Big(f(2^n)\, 2^{-n/2}\Big)^{d-2},$$

so $\sum_n \mathbb{P}(A_n) = \infty$ by (2.4). For $m < n-1$, the Markov property at time 2^{n-1}, Brownian scaling and Lemma 3.24 yield that

$$\begin{aligned}\mathbb{P}[A_n \mid A_m] &\le \sup_{x\in\mathbb{R}^d} \mathbb{P}_x\big\{ \text{ there exists } t>1 \text{ with } |B(t)| \le f(2^{n+1})2^{(1-n)/2}\big\}\\ &\le C_1 \Big(f(2^{n+1})\, 2^{(1-n)/2}\Big)^{d-2}.\end{aligned}$$

From this we get that

$$\begin{aligned}\liminf_{k\to\infty} \frac{\sum_{m=1}^k \sum_{n=1}^k \mathbb{P}(A_n\cap A_m)}{\big(\sum_{n=1}^k \mathbb{P}(A_n)\big)^2} &\le 2 \liminf_{k\to\infty} \frac{\sum_{m=1}^k \mathbb{P}(A_m) \sum_{n=m+2}^k \mathbb{P}[A_n \mid A_m]}{\big(\sum_{n=1}^k \mathbb{P}(A_n)\big)^2}\\ &\le 2\frac{C_1}{C_5} \liminf_{k\to\infty} \frac{\sum_{n=1}^k (f(2^{n+1})\, 2^{(1-n)/2})^{d-2}}{\sum_{n=1}^k (f(2^n)\, 2^{-n/2})^{d-2}} < \infty\,.\end{aligned}$$

The Kochen-Stone lemma now yields that $\mathbb{P}\{A_n \text{ infinitely often}\} > 0$, whence by the Hewitt-Savage 0-1-law this probability is 1. Thus the set $\{t>0\colon |B(t)| \le f(t)\}$ is almost surely unbounded. Since (2.4) also applies to εf in place of f for any $\varepsilon > 0$, it follows that $\liminf_{t\uparrow\infty} |B(t)|/f(t) = \infty$ almost surely ■

3. Occupation measures and Green's functions

We now address the following question: Given a bounded domain $U \subset \mathbb{R}^d$, how much time does Brownian motion spend in U? Our first result states that for a linear Brownian motion running for a finite amount of time, this time is comparable to the Lebesgue measure of U.

THEOREM 3.25. *Let $\{B(s)\colon s\ge 0\}$ be a linear Brownian motion and $t>0$. Define the occupation measure μ_t by*

$$\mu_t(A) = \int_0^t \mathbb{1}_A(B(s))\, ds \qquad \textit{for } A\subset\mathbb{R} \textit{ Borel.}$$

Then, almost surely, μ_t is absolutely continuous with respect to the Lebesgue measure.

Proof. By Lebesgue's theorem absolute continuity with respect to the Lebesgue measure means that, for μ_t-almost very $x\in\mathbb{R}$,

$$\liminf_{r\downarrow 0} \frac{\mu_t(\mathcal{B}(x,r))}{\mathcal{L}(\mathcal{B}(x,r))} < \infty\,.$$

To see this we use first Fatou's lemma and then Fubini's theorem,

$$\begin{aligned}\mathbb{E}\int \liminf_{r\downarrow 0} \frac{\mu_t(\mathcal{B}(x,r))}{\mathcal{L}(\mathcal{B}(x,r))}\, d\mu_t(x) &\le \liminf_{r\downarrow 0} \frac{1}{2r}\, \mathbb{E}\int \mu_t(\mathcal{B}(x,r))\, d\mu_t(x)\\ &= \liminf_{r\downarrow 0} \frac{1}{2r} \int_0^t \int_0^t \mathbb{P}\big\{|B(s_1)-B(s_2)| \le r\big\}\, ds_1\, ds_2\,.\end{aligned}$$

Using that the density of a standard normal random variable X is bounded by one, we get

$$\mathbb{P}\big\{|B(s_1)-B(s_2)|\le r\big\} = \mathbb{P}\big\{|X| \le \tfrac{r}{\sqrt{|s_1-s_2|}}\big\} \le \tfrac{r}{\sqrt{|s_1-s_2|}},$$

and this implies that

$$\liminf_{r\downarrow 0} \frac{1}{2r} \int_0^t \int_0^t \mathbb{P}\big\{|B(s_1) - B(s_2)| \le r\big\}\, ds_1\, ds_2 \le \frac{1}{2} \int_0^t \int_0^t \frac{ds_1\, ds_2}{\sqrt{|s_1 - s_2|}} < \infty.$$

This implies that μ_t is absolutely continuous with respect to $\mathcal{L}$. ■

We now turn to higher dimensions $d \ge 2$. A first simple result shows that whether the overall time spent in a bounded set is finite or not depends just on transience or recurrence of the process.

THEOREM 3.26. *Let $U \subset \mathbb{R}^d$ be a nonempty bounded open set and $x \in \mathbb{R}^d$ arbitrary.*

- *If $d = 2$, then $\mathbb{P}_x$-almost surely,* $\displaystyle\int_0^\infty \mathbb{1}_U(B(t))\, dt = \infty\,.$
- *If $d \ge 3$, then* $\displaystyle\mathbb{E} \int_0^\infty \mathbb{1}_U(B(t))\, dt < \infty\,.$

Proof. As U is contained in a ball and contains a ball, it suffices to show this for balls. By shifting, we can even restrict to balls $U = \mathcal{B}(0, r)$ centred in the origin. Let us start with the first claim. We let $d \le 2$ and let $G = \mathcal{B}(0, 2r)$. Let $T_0 = 0$ and, for all $k \ge 1$, let

$$S_k = \inf\{t > T_{k-1} \,:\, B(t) \in U\} \text{ and } T_k = \inf\{t > S_k \,:\, B(t) \notin G\}.$$

Recall that, almost surely, these stopping times are finite. From the strong Markov property we infer, for $k \ge 1$,

$$\mathbb{P}_x\Big\{ \int_{S_k}^{T_k} \mathbb{1}_U(B(t))\, dt \ge s \,\Big|\, \mathcal{F}^+(S_k)\Big\} = \mathbb{P}_{B(S_k)}\Big\{ \int_0^{T_1} \mathbb{1}_U(B(t))\, dt \ge s\Big\}$$
$$= \mathbb{E}_x\Big[\mathbb{P}_{B(S_k)}\Big\{ \int_0^{T_1} \mathbb{1}_U(B(t))\, dt \ge s\Big\}\Big] = \mathbb{P}_x\Big\{ \int_{S_k}^{T_k} \mathbb{1}_U(B(t))\, dt \ge s\Big\},$$

by rotation invariance. The second expression does not depend on k, so that the random variables

$$\int_{S_k}^{T_k} \mathbb{1}_U(B(t))\, dt, \qquad \text{for } k = 1, 2, \ldots$$

are independent and identically distributed. As they are not identically zero, but nonnegative, they have positive expectation and, by the strong law of large numbers we infer

$$\int_0^\infty \mathbb{1}_U(B(t))\, dt = \lim_{n\to\infty} \sum_{k=1}^n \int_{S_k}^{T_k} \mathbb{1}_U(B(t))\, dt = \infty,$$

which proves the first claim. For the second claim, we first look at Brownian motion started in the origin and obtain, making good use of Fubini's theorem and denoting by $\mathfrak{p}\colon [0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d \to [0, 1]$ the transition density of Brownian motion,

$$\mathbb{E}_0 \int_0^\infty \mathbb{1}_{\mathcal{B}(0,r)}(B(s))\, ds = \int_0^\infty \mathbb{P}_0\{B(s) \in \mathcal{B}(0, r)\}\, ds = \int_0^\infty \int_{\mathcal{B}(0,r)} \mathfrak{p}(s, 0, y)\, dy\, ds$$
$$= \int_{\mathcal{B}(0,r)} \int_0^\infty \mathfrak{p}(s, 0, y)\, ds\, dy = \sigma(\partial\mathcal{B}(0, 1)) \int_0^r \rho^{d-1} \int_0^\infty \Big(\tfrac{1}{\sqrt{2\pi s}}\Big)^d e^{\frac{-\rho^2}{2s}}\, ds\, d\rho.$$

Now we can use the substitution $t = \rho^2/s$ and obtain, for a suitable constant $C(d) < \infty$,

$$= C(d) \int_0^r \rho^{d-1} \rho^{2-d} \, d\rho = \tfrac{C(d)}{2} \, r^2 < \infty.$$

For start in an arbitrary $x \neq 0$, we look at a Brownian motion started in 0 and a stopping time T, which is the first hitting time of the sphere $\partial\mathcal{B}(0, |\,x|\,)$. Using spherical symmetry and the strong Markov property we obtain

$$\mathbb{E}_x \int_0^\infty \mathbb{1}_{\mathcal{B}(0,r)}(B(s)) \, ds = \mathbb{E}_0 \int_T^\infty \mathbb{1}_{\mathcal{B}(0,r)}(B(s)) \, ds \leq \mathbb{E}_0 \int_0^\infty \mathbb{1}_{\mathcal{B}(0,r)}(B(s)) \, ds < \infty. \qquad \blacksquare$$

In the case when Brownian motion is transient it is interesting to ask further for the expected time the process spends in a bounded open set. In order not to confine this discussion to the case $d \geq 3$ we introduce suitable stopping rules for Brownian motion in $d = 2$.

Definition 3.27. *Suppose that $\{B(t) : 0 \leq t \leq T\}$ is a d-dimensional Brownian motion and one of the following three cases holds:*

(1) *$d \geq 3$ and $T = \infty$,*
(2) *$d \geq 2$ and T is an independent exponential time with parameter $\lambda > 0$,*
(3) *$d \geq 2$ and T is the first exit time from a bounded domain D containing* 0.

We use the convention that $D = \mathbb{R}^d$ in cases (1), (2). We refer to these three cases by saying that $\{B(t) : 0 \leq t \leq T\}$ is a **transient Brownian motion**. $\diamond$

Remark 3.28. For a transient Brownian motion $\{B(t) : 0 \leq t \leq T\}$, given $\mathcal{F}^+(t)$, on the event $\{B(t) = y, \, t < T\}$, the process $\{B(s + t) : 0 \leq s \leq T\}$ is again a transient Brownian motion of the same type, started in y. We do not consider Brownian motion stopped at a *fixed* time, because this model lacks exactly this form of the Markov property. $\diamond$

Proposition 3.29. *For transient Brownian motion $\{B(t) : 0 \leq t \leq T\}$ there exists a transition (sub-)density $\mathfrak{p}^* \colon [0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d \to [0, 1]$ such that, for any $t > 0$,*

$$\mathbb{P}_x\big\{B(t) \in A \text{ and } t \leq T\big\} = \int_A \mathfrak{p}^*(t, x, y) \, dy \qquad \text{for every } A \subset \mathbb{R}^d \text{ Borel.}$$

Moreover, for all $t \geq 0$ and $\mathcal{L}$-almost every $x, y \in D$ we have $\mathfrak{p}^*(t, x, y) = \mathfrak{p}^*(t, y, x)$.

Proof. Fix t throughout the proof. For the existence of the density, by the Radon-Nikodym theorem, it suffices to check that $\mathbb{P}_x\{B(t) \in A \text{ and } t \leq T\} = 0$, if A is a Borel set of Lebesgue measure zero. This is obvious, by just dropping the requirement $t \leq T$, and recalling that $B(t)$ is normally distributed. If $d \geq 3$ and $T = \infty$, or if $d \geq 2$ and T is independent, exponentially distributed symmetry is obvious.

Hence we can now concentrate on the case $d \geq 2$ and a bounded domain D. We fix a compact set $K \subset D$ and define, for every $x \in K$ and $n \in \mathbb{N}$, a measure $\mu_x^{(n)}$ on the Borel sets $A \subset D$,

$$\mu_x^{(n)}(A) = \mathbb{P}_x\big\{B(\tfrac{kt}{2^n}) \in K \text{ for all } k = 0, \ldots, 2^n \text{ and } B(t) \in A\big\} .$$

Then $\mu_x^{(n)}$ has a density

$$\mathfrak{p}_n^*(t,x,y) = \int_K \cdots \int_K \prod_{i=1}^{2^n} \mathfrak{p}\big(\tfrac{t}{2^n}, z_{i-1}, z_i\big)\, dz_1 \dots dz_{2^n-1}\,,$$

where $z_0 = x$, $z_{2^n} = y$ and $\mathfrak{p}$ is the transition density of d-dimensional Brownian motion. As $\mathfrak{p}$ is symmetric in the space variables, so is $\mathfrak{p}_n^*$ for every n. Note that $\mathfrak{p}_n^*$ is decreasing in n. From the monotone convergence theorem one can see that $\mathfrak{p}_K^*(t,x,y) := \lim \mathfrak{p}_n^*(t,x,y)$ is a transition subdensity of Brownian motion stopped upon leaving K. The symmetry of $\mathfrak{p}_n^*$ gives $\mathfrak{p}_K^*(t,x,y) = \mathfrak{p}_K^*(t,y,x)$. Choosing an increasing sequence of compact sets exhausting U and taking a monotone limit yields a symmetric version $\mathfrak{p}^*(t,x,y)$ of the transition density. ■

Definition 3.30. *For transient Brownian motion $\{B(t) : 0 \le t \le T\}$ we define the* **Green's function** $G\colon \mathbb{R}^d \times \mathbb{R}^d \to [0,\infty]$ *by*

$$G(x,y) = \int_0^\infty \mathfrak{p}^*(t,x,y)\, dt\,.$$

The Green's function is also called the **Green kernel**. *Sometimes it is also called the* potential kernel, *but we shall reserve this terminus for a closely related concept, see Remark 8.20.* ◇

In probabilistic terms G is the density of the *expected* occupation measure for the transient Brownian motion started in x.

Theorem 3.31. *If $f\colon \mathbb{R}^d \to [0,\infty]$ is measurable, then*

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

Proof. Fubini's theorem implies

$$\begin{aligned} \mathbb{E}_x \int_0^T f(B(t))\, dt &= \int_0^\infty \mathbb{E}_x\big[f(B(t))\, \mathbb{1}_{\{t \le T\}}\big]\, dt = \int_0^\infty \int \mathfrak{p}^*(t,x,y)\, f(y)\, dy\, dt \\ &= \int \int_0^\infty \mathfrak{p}^*(t,x,y)\, dt\, f(y)\, dy = \int G(x,y) f(y)\, dy, \end{aligned}$$

by definition of the Green's function. ■

In case (1), i.e. if $T = \infty$, Green's function can be calculated explicitly.

Theorem 3.32. *If $d \ge 3$ and $T = \infty$, then*

$$G(x,y) = c(d)\, |\, x - y|^{2-d}, \qquad \textit{where } c(d) = \tfrac{\Gamma(d/2-1)}{2\pi^{d/2}}.$$

Proof. Assume $d \geq 3$ and use the substitution $s = |\,x-y|^2/2t$ to obtain,

$$\begin{aligned} G(x,y) &= \int_0^\infty \frac{1}{(2\pi t)^{-d/2}}\, e^{-|\,x-y|^2/2t}\, dt = \int_\infty^0 \Big(\frac{s}{\pi|\,x-y|^2}\Big)^{d/2} e^{-s} \Big(-\frac{|\,x-y|^2}{2s^2}\Big)\, ds \\ &= \frac{|\,x-y|^{2-d}}{2\pi^{d/2}} \int_0^\infty s^{(d/2)-2}\, e^{-s}\, ds = \frac{\Gamma(d/2-1)}{2\pi^{d/2}}\, |\,x-y|^{2-d}, \end{aligned}$$

where $\Gamma(x) = \int_0^\infty s^{x-1} e^{-s}\, ds$ is the Gamma function. This proves that G has the given form and the calculation above also shows that the integral is infinite if $d \leq 2$. ■

In case (2), if Brownian motion is stopped at an independent exponential time, one can find the asymptotics of $G(x,y)$ for $x \to y$.

THEOREM 3.33. *If $d = 2$ and T is an independent exponential time with parameter $\lambda > 0$, then*

$$G(x,y) \sim -\frac{1}{\pi}\log|\,x-y| \qquad \text{for } |x-y| \downarrow 0\,.$$

Proof. Note that the transition sub-density of Brownian motion stopped at an independent exponential time with parameter $\lambda > 0$ equals

$$\mathfrak{p}^*(t,x,y) = e^{-\lambda t}\, \mathfrak{p}(t,x,y),$$

where $\mathfrak{p}$ is the transition density for (unstopped) Brownian motion. Hence

$$G(x,y) = G_\lambda(x,y) := \int_0^\infty \frac{1}{2\pi t}\, \exp\big\{-\tfrac{|x-y|^2}{2t} - \lambda t\big\}\, dt\,.$$

We thus get $G_\lambda(x-y) = G_1(\sqrt{\lambda}(x-y))$ and may assume without loss of generality that $\lambda = 1$. Then

$$G(x,y) = \frac{1}{2\pi}\int_0^\infty \frac{e^{-t}}{t} \int_{|x-y|^2/(2t)}^\infty e^{-s}\, ds\, dt = \frac{1}{2\pi}\int_0^\infty e^{-s} \int_{|x-y|^2/(2s)}^\infty \frac{e^{-t}}{t}\, dt\, ds\,.$$

For an upper bound we use that, for $|x-y| \leq 1$, that

$$\int_{|x-y|^2/(2s)}^\infty \frac{e^{-t}}{t}\, dt \leq \begin{cases} \log \frac{2s}{|x-y|^2} + 1, & \text{if } |x-y|^2 \leq 2s, \\ 1, & \text{if } |x-y|^2 > 2s. \end{cases}$$

This gives, with $0 < \gamma := -\int_0^\infty e^{-s} \log s\, ds$ denoting Euler's constant,

$$G(x,y) \leq \frac{1}{2\pi}\big(1 + \log 2 - \gamma - 2\log|x-y|\big),$$

from which the upper bound follows. For the lower bound we use

$$\int_{|x-y|^2/(2s)}^\infty \frac{e^{-t}}{t}\, dt \geq \int_{|x-y|^2/(2s)}^1 \frac{1-t}{t}\, dt \geq \log\frac{2s}{|x-y|^2} - 1,$$

and thus

$$G(x,y) \geq \frac{1}{2\pi}\big(-1 + \log 2 - \gamma - 2\,\log|x-y|\big),$$

and again this is asymptotically equal to $-\frac{1}{\pi}\log|\,x-y|$. ■

We now explore some of the major analytic properties of Green's function.

THEOREM 3.34. *In all three cases of transient Brownian motion in $d \geq 2$, the Green's function has the following properties:*

(i) *$G(x, y)$ is finite if $x \neq y$.*

(ii) *$G(x, y) = G(y, x)$ for all $x, y \in D$.*

(iii) *for any $y \in D$ the Green's function $G(\,\cdot\,, y)$ is harmonic on $D \setminus \{y\}$.*

This result is easy in the case $d \geq 3$, $T = \infty$, where the Green's function is explicitly known by Theorem 3.32. We therefore focus on the case $d = 2$ and prepare the proof by two lemmas, which are of some independent interest.

LEMMA 3.35. *If $d = 2$, for $x, z \in \mathbb{R}^2$ with $|x - z| = 1$,*

$$-\frac{1}{\pi} \log |\,x - y| = \int_0^\infty \mathfrak{p}(s, x, y) - \mathfrak{p}(s, x, z)\, ds\,,$$

where $\mathfrak{p}$ is the transition kernel for the (unstopped) Brownian motion.

Proof. For $|x - z| = 1$, we obtain

$$\int_0^\infty \mathfrak{p}(t, x, y) - \mathfrak{p}(t, x, z)\, dt = \frac{1}{2\pi} \int_0^\infty \big(e^{-\frac{|x-y|^2}{2t}} - e^{-\frac{1}{2t}}\big)\, \frac{dt}{t} = \frac{1}{2\pi} \int_0^\infty \Big(\int_{|x-y|^2/2t}^{1/2t} e^{-s}\, ds\Big)\, \frac{dt}{t}\,,$$

and by changing the order of integration this equals

$$\frac{1}{2\pi} \int_0^\infty e^{-s} \Big(\int_{|\,x-y|^2/2s}^{1/2s} \frac{dt}{t}\Big)\, ds = -\frac{1}{\pi}\, \log |\,x - y|\,,$$

which completes the proof. ■

LEMMA 3.36. *Let $D \subset \mathbb{R}^2$ be a bounded domain and $x, y \in D$. Then with $u(x) = 2 \log |x|$,*

$$G(x, y) = \frac{-1}{2\pi}\, u(x - y) - \mathbb{E}_x \Big[\frac{-1}{2\pi}\, u\big(B(\tau) - y\big)\Big]\,,$$

where τ is the first exit time from D.

Proof. Let $f \colon D \to [0, \infty)$ be continuous with compact support. Picking $v \in \partial\mathcal{B}(0, 1)$, we obtain for any $x \in D$,

$$\begin{aligned}\int G(x, y)\, f(y)\, dy &= \mathbb{E}_x \int_0^\infty f(B(t)) - \mathbb{1}\{t \geq \tau\} f(B(t))\, dt \\ &= \int_0^\infty \int \big(\mathfrak{p}(t, x, y) - \mathfrak{p}(t, x, x + v)\big)\, f(y)\, dy\, dt \\ &\quad - \mathbb{E}_x \int_0^\infty \int \big(\mathfrak{p}(t, B(\tau), y) - \mathfrak{p}(t, B(\tau), B(\tau) + v)\big)\, f(y)\, dy\, dt\,.\end{aligned}$$

This implies the statament for every $x \in D$ and almost every $y \in D$. It remains to show that we can choose a version of the density $\mathfrak{p}^*(t, \cdot, \cdot)$ such that the statement holds for *every* $x, y \in D$. Indeed, let

$$\mathfrak{p}^*(t,x,y) = \mathfrak{p}(t,x,y) - \mathbb{E}_x\big[\mathfrak{p}(t-\tau, B(\tau), y)\mathbb{1}\{\tau < t\}\big] .$$

Integrating over all $y \in A$ gives that this is indeed a transition density for the stopped process. Moreover, adding and subtracting $\mathfrak{p}(t,x,x+v) = \mathfrak{p}(t, B(\tau), B(\tau)+v)$ on the right hand side and integrating over $t \in (0,\infty)$ yields the statement for the Green's function associated to this particular choice of transition kernel $\mathfrak{p}^*$, which therefore holds for all $x, y \in D$. ■

Proof of Theorem 3.34. We first look at properties (i), (ii) and continuity of $G(\cdot\,, y)$ on $D \setminus \{y\}$. These three properties are obvious in the case $d \geq 3$, $T = \infty$, by the explicit form of the Green's function uncovered in Theorem 3.32. If Brownian motion is stopped at an independent exponential time, it is easy to see from $\mathfrak{p}^*(t,x,y) = e^{-\lambda t}\,\mathfrak{p}(t,x,y)$ that the Green's function is finite everywhere except on the diagonal $\Delta = \{(x,y)\colon x = y\}$, and symmetric. Moreover, continuity is easy to check using dominated convergence. We can therefore focus on the case where the Brownian motion is stopped upon leaving a bounded domain D.

First let $d = 2$. Lemma 3.35 gives, for $x \neq y$, that $G(x,y) < \infty$. However, we have $\mathbb{E}_x[-1/(2\pi)\, u(B(\tau) - x)] < \infty$, hence $G(x,x) = \infty$ by Lemma 3.36. If $x \in D$, then $G(x, \cdot)$ is continuous on $D \setminus \{x\}$, because the right hand side of the equation in Lemma 3.36 is continuous. Similarly, if $y \in D$ the right hand side is continuous in x on $D \setminus \{y\}$, as $\mathbb{E}_x[u(B(\tau) - y)]$ is harmonic in x. Hence $G(\cdot\,, x)$ is also continuous on $D \setminus \{x\}$. The symmetry follows from the almost-everywhere symmetry of $\mathfrak{p}^*(t, \cdot, \cdot)$ together with the continuity.

Next, if $d \geq 3$ we can carry out the same proof replacing $-1/(2\pi)u(x,y)$ by $\ell(x,y) = c(d)|x - y|^{2-d}$. In fact the arguments become significantly easier because

$$\ell(x,y) = \int_0^\infty \mathfrak{p}(t,x,y)\,dt\,, \text{ for all } x, y \in \mathbb{R}^d\,,$$

and there is no need to subtract a 'renormalisation' term.

Finally, we show property (iii) in all cases. Define

$$G_\varepsilon(x,y) := \int_{\mathcal{B}(y,\varepsilon)} G(x,z)\,dz, \qquad \text{for } \mathcal{B}(y,\varepsilon) \subset D \text{ and } x \in D.$$

We prove that $G_\varepsilon(\cdot\,, y)$ satisfies the mean value property on $D \setminus \mathcal{B}(y,\varepsilon)$, i.e.

$$G_\varepsilon(x,y) = \frac{1}{\mathcal{L}(\mathcal{B}(x,r))} \int_{\mathcal{B}(x,r)} G_\varepsilon(z,y)\,dz, \qquad \text{for } 0 < r < |\,x - y| - \varepsilon.$$

The result follows from this since, using continuity of G, for $x, y \in D$ with $|x - y| > r$,

$$\begin{aligned} G(x,y) &= \lim_{\varepsilon \downarrow 0} \frac{G_\varepsilon(x,y)}{\mathcal{L}(\mathcal{B}(y,\epsilon))} = \lim_{\varepsilon \downarrow 0} \frac{1}{\mathcal{L}(\mathcal{B}(x,r))} \int_{\mathcal{B}(x,r)} \frac{G_\varepsilon(z,y)}{\mathcal{L}(\mathcal{B}(y,\varepsilon))}\,dz \\ &= \frac{1}{\mathcal{L}(\mathcal{B}(x,r))} \int_{\mathcal{B}(x,r)} G(z,y)\,dz, \end{aligned}$$

where the last equality follows from the bounded convergence theorem.

Fix $x \neq y$ in D, let $0 < r < |\,x - y|$ and let $\varepsilon < |\,x - y| - r$. Denote $\tau = T \wedge \inf\{t\colon |B(t) - x| = r\}$. As a Brownian motion started in x spends no time in $\mathcal{B}(y, \varepsilon)$ before time τ, we can write

$$G_\varepsilon(x,y) = \mathbb{E}_x \int_\tau^T \mathbb{1}\{B(t) \in \mathcal{B}(y,\varepsilon)\}\,dt = \mathbb{E}_x\Big[\mathbb{E}_{B(\tau)} \int_0^T \mathbb{1}\{\tilde{B}(t) \in \mathcal{B}(y,\varepsilon)\}\,dt\Big],$$

where the inner expectation is with respect to a Brownian motion $\{\tilde{B}(t)\colon t \geq 0\}$ started in the fixed point $B(\tau)$, whereas the outer expectation is with respect to $B(\tau)$. By the strong Markov property and since, given $\tau < T$, the random variable $B(\tau)$ is uniformly distributed on $\partial\mathcal{B}(x, r)$, by rotational symmetry, we conclude,

$$G_\varepsilon(x,y) = \mathbb{E}_x G_\varepsilon(B(\tau), y) = \int_{\partial\mathcal{B}(x,r)} G_\varepsilon(z,y)\,d\varpi_{x,r}(z),$$

so that G_ε satisfies the mean value property and hence is harmonic by Theorem 3.2. ■

Remark 3.37. Suppose $d \geq 3$ and $T = \infty$. Let K be a compact set and μ a measure on ∂K. Then

$$u(x) = \int_{\partial K} G(x,y)\,d\mu(y), \qquad \text{for } x \in K^{\mathrm{c}}$$

is a harmonic function on K^{c}. This can be verified easily from the harmonicity of $G(\,\cdot\,, y)$ and the mean value property. Physically, $u(x)$ is the electrostatic (or Newtonian) potential at x resulting from a charge represented by μ. In particular, the Green function $G(\,\cdot\,, y)$ can be interpreted as the electrostatic potential induced by a unit charge in the point y. An interesting question is whether every positive harmonic function on K^{c} can be represented in such a way by a suitable μ. We will come back to this question in Chapter 8. ◇

4. The harmonic measure

A particularly appealing way of writing the harmonic function u in Theorem 3.8 is in terms of the harmonic measure on ∂U.

Definition 3.38. *Let $\{B(t)\colon t \geq 0\}$ be a d-dimensional Brownian motion, $d \geq 2$, started in some point x and fix a closed set $A \subset \mathbb{R}^d$. Define a measure $\mu_A(x, \,\cdot\,)$ by*

$$\mu_A(x, B) = \mathbb{P}\big\{B(\tau) \in B,\, \tau < \infty\big\} \quad \text{where } \tau = \inf\{t \geq 0 \,:\, B(t) \in A\},$$

for $B \subset A$ Borel. In other words, $\mu_A(x, \,\cdot\,)$ is the distribution of the first hitting point of A, and the total mass of the measure is the probability that a Brownian motion started in x ever hits the set A. The harmonic measure is supported by ∂A. ◇

The following corollary is only an equivalent reformulation of Theorem 3.12.

COROLLARY 3.39. *If the Poincaré cone condition is satisfied at every point $x \in \partial U$ on the boundary of a bounded domain U, then the solution of the Dirichlet problem with boundary condition $\varphi \colon \partial U \to \mathbb{R}$, can be written as*

$$u(x) = \int \varphi(y)\, \mu_{\partial U}(x, dy) \qquad \text{for all } x \in \overline{U}.$$

REMARK 3.40. Of course, the harmonicity of u does not rely on the Poincaré cone condition. In fact, by Theorem 3.8, for any compact $A \subset \mathbb{R}^d$ and Borel set $B \subset \partial A$, the function $x \mapsto \mu_A(x, B)$ is harmonic on A^{c}. ◇

Besides its value in the discussion of the Dirichlet problem, the harmonic measure is also interesting in its own right, as it intuitively weighs the points of A according to their accessibility from x. We now show that the measures $\mu_A(x, \cdot\,)$ for different values of $x \in A^{\mathrm{c}}$ are mutually absolutely continuous. This is a form of the famous *Harnack principle*.

THEOREM 3.41 (Harnack principle). *Suppose $A \subset \mathbb{R}^d$ is compact and x, y are in the unbounded component of A^{c}. Then $\mu_A(x, \cdot\,) \ll \mu_A(y, \cdot\,)$.*

Proof. Given $B \subset \partial A$ Borel, by Remark 3.40, the mapping $x \mapsto \mu_A(x, B)$ is a harmonic function on A^{c}. If it takes the value zero for some $y \in A^{\mathrm{c}}$, then y is a minimum and the maximum principle, Theorem 3.5, together with the subsequent remark, imply that $\mu_A(x, B) = 0$ for all $x \in A^{\mathrm{c}}$, as required. ■

The Harnack principle allows to formulate the following definition.

DEFINITION 3.42. *A compact set A is called* **nonpolar for Brownian motion**, *or simply* **nonpolar**, *if $\mu_A(x, A) > 0$ for one (and hence for all) $x \in A^{\mathrm{c}}$. Otherwise, the set A is called* **polar for Brownian motion**. ◇

We now give an explicit formula for the harmonic measures on the unit sphere $\partial\mathcal{B}(0,1)$. Note that if $x = 0$ then the distribution of $B(\tau)$ is (by symmetry) the uniform distribution, but if x is another point it is an interesting problem to determine this distribution in terms of a probability density.

THEOREM 3.43 (Poisson's formula). *Suppose that $A \subset \partial\mathcal{B}(0,1)$ is a Borel subset of the unit sphere for $d \geq 2$. Let ϖ denote the uniform distribution on the unit sphere. Then, for all $x \notin \partial\mathcal{B}(0,1)$,*

$$\mu_{\partial\mathcal{B}(0,1)}(x, A) = \int_A \frac{\big|\, 1 - |\,x|^2\, \big|}{|\,x - y|^d}\, d\varpi(y).$$

REMARK 3.44. The density appearing in the theorem is usually called the *Poisson kernel* and appears frequently in potential theory. ◇

Proof. We start by looking at the case $|x| < 1$. To prove the theorem we indeed show that for every bounded measurable $f\colon \mathbb{R}^d \to \mathbb{R}$ we have

$$\mathbb{E}_x[f(B(\tau))] = \int_{\partial\mathcal{B}(0,1)} \frac{1-|x|^2}{|x-y|^d} f(y)\, d\varpi(y), \tag{4.1}$$

which on the one hand implies the formula by choosing indicator functions, on the other hand, by the monotone class theorem, see e.g. [**Du95**, Chapter 5, (1.5)], it suffices to show this for smooth functions. To prove (4.1) we recall Theorem 3.12, which tells us that we just have to show that the right hand side as a function in $x \in \mathcal{B}(0,1)$ defines a solution of the Dirichlet problem on $\mathcal{B}(0,1)$ with boundary value f.

To check this, one first checks that $\frac{1-|x|^2}{|x-y|^d}$ is harmonic in x on $\mathcal{B}(0,1)$, which is a straightforward calculation, and then argues that it is allowed to differentiate twice under the integral sign. To check the boundary condition first look at the case $f \equiv 1$, in which case we have to show that, for all $x \in \mathcal{B}(0,1)$,

$$I(x) := \int_{\partial\mathcal{B}(0,1)} \frac{1-|x|^2}{|x-y|^d}\, \varpi(dy) \equiv 1.$$

We use Theorem 3.12 to show this. Indeed, observe that $I(0) = 1$, I is invariant under rotation and $\Delta I = 0$ on $\mathcal{B}(0,1)$, by the first part. Now let $x \in \mathcal{B}(0,1)$ with $|x| = r < 1$ and let $\tau := \inf\{t\colon |B(t)| > r\}$. By Theorem 3.12,

$$I(0) = \mathbb{E}_0\big[I(B(\tau))\big] = I(x)\,,$$

using rotation invariance in the second step. Hence $I \equiv 1$. Now we show that the right hand side in the theorem can be extended continuously to all points $y \in \partial\mathcal{B}(0,1)$ by $f(y)$. We write D_0 for $\partial\mathcal{B}(0,1)$ with a δ-neighbourhood $\mathcal{B}(y,\delta)$ removed and $D_1 = \partial\mathcal{B}(0,1) \setminus D_0$. We have, using that $I \equiv 1$, for all $x \in \mathcal{B}(y,\delta/2) \cap U$,

$$\begin{aligned}
&\Big|f(y) - \int_{\partial\mathcal{B}(0,1)} \frac{1-|x|^2}{|x-z|^d} f(z)\, d\varpi(z)\Big| \\
&= \Big|\int_{\partial\mathcal{B}(0,1)} \frac{1-|x|^2}{|x-z|^d} (f(y)-f(z))\, d\varpi(z)\Big| \\
&\leq 2\|f\|_\infty \int_{D_0} \frac{1-|x|^2}{|x-z|^d}\, d\varpi(z) + \sup_{z\in D_1} |f(y)-f(z)|.
\end{aligned}$$

For fixed $\delta > 0$ the first term goes to 0 as $x \to y$ by dominated convergence, whereas the second can be made arbitrarily small by choice of δ. This completes the proof if $x \in \mathcal{B}(0,1)$.

If $|x| > 1$ we use *inversion at the unit circle* to transfer the problem to the case studied before. By a straightforward calculation, one can check that a function $u\colon \overline{\mathcal{B}(0,1)}^{\mathrm{c}} \to \mathbb{R}$ is harmonic if and only if its inversion $u^*\colon \mathcal{B}(0,1) \setminus \{0\} \to \mathbb{R}$ defined by

$$u^*\big(\tfrac{x}{|x|^2}\big) = u(x)|x|^{d-2}$$

is harmonic. Now suppose that $f\colon \partial\mathcal{B}(0,1) \to R$ is a smooth function on the boundary. Then define a harmonic function $u\colon \overline{\mathcal{B}(0,1)}^{\mathrm{c}} \to \mathbb{R}$ by

$$u(x) = \mathbb{E}_x\big[f\big(B(\tau(\partial\mathcal{B}(0,1)))\big)\, \mathbb{1}\big\{\tau(\partial\mathcal{B}(0,1)) < \infty\big\}\big]\,.$$

Then $u^*\colon \mathcal{B}(0,1) \setminus \{0\} \to \mathbb{R}$ is bounded and harmonic. By Exercise 3.8 we can extend it to the origin, so that the extension is harmonic on $\mathcal{B}(0,1)$. In fact, this extension is obviously given by $u^*(0) = \int \varphi \, d\varpi$. The harmonic extension is continuous on the closure, with boundary values given by f. Hence it agrees with the function of the first part, and $u = u^{**}$ must be its inversion, which gives the claimed formula. ■

We now fix a compact nonpolar set $A \subset \mathbb{R}^d$, and look at the harmonic measure $\mu_A(x, \cdot\,)$ when $x \to \infty$. The first task is to make sure that this object is well-defined.

THEOREM 3.45. *Let $A \subset \mathbb{R}^d$ be a compact, nonpolar set, then there exists a probability measure μ_A on A, given by*

$$\mu_A(B) = \lim_{x\to\infty} \mathbb{P}_x\big\{B(\tau(A)) \in B \,|\, \tau(A) < \infty\big\}\,.$$

This measure is called the **harmonic measure** *(from infinity).*

REMARK 3.46. The harmonic measure weighs the points of A according to their accessibility from infinity. It is naturally supported by the *outer boundary* of A, which is the boundary of the infinite connected component of $\mathbb{R}^d \setminus A$. ◇

The proof is prepared by a lemma, which is yet another example how the strong Markov property can be exploited to great effect.

LEMMA 3.47. *For $A \subset \mathbb{R}^d$ compact and nonpolar and every $\varepsilon > 0$, there exists a large $R > 0$ such that, for all $x \in \partial\mathcal{B}(0,R)$ and any hyperplane $H \subset \mathbb{R}^d$ containing the origin,*

$$\mathbb{P}_x\big\{\tau(A) < \tau(H)\big\} < \varepsilon\,\mathbb{P}_x\big\{\tau(A) < \infty\big\}.$$

Proof. Pick a radius $r > 0$ such that $A \subset \mathcal{B}(0,r)$ and note from Remark 3.40 that $x \mapsto \mathbb{P}_x\{\tau(A) < \infty\}$ is harmonic on $A^{\rm c}$. Therefore the minimum of this function on the compact set $\partial\mathcal{B}(0,r)$ is positive, say $\delta > 0$. It therefore suffices to show that

$$\mathbb{P}_x\big\{\tau(\mathcal{B}(0,r) < \tau(H)\big\} < \varepsilon\delta\,\mathbb{P}_x\big\{\tau(\mathcal{B}(0,r) < \infty\big\}.$$

Now there exists an absolute constant $q < 1$ such that, for any $x \in \partial\mathcal{B}(0,2)$ and hyperplane H,

$$\mathbb{P}_x\big\{\tau(\mathcal{B}(0,1)) < \tau(H)\big\} < q\,\mathbb{P}_x\big\{\tau(\mathcal{B}(0,1)) < \infty\big\}.$$

Let k be large enough to ensure that $q^k < \varepsilon\delta$. Then, by the strong Markov property and Brownian scaling,

$$\begin{aligned}
&\sup_{x\in\partial\mathcal{B}(0,r2^k)} \mathbb{P}_x\big\{\tau(\mathcal{B}(0,r)) < \tau(H)\big\}\\
&\qquad\le \sup_{x\in\partial\mathcal{B}(0,r2^k)} \mathbb{E}_x\Big[\mathbb{1}\{\tau(\mathcal{B}(0,r2^{k-1})) < \tau(H)\}\,\mathbb{P}_{B(\tau(\mathcal{B}(0,r2^{k-1})))}\big\{\tau(\mathcal{B}(0,r)) < \tau(H)\big\}\Big]\\
&\qquad\le q \sup_{x\in\partial\mathcal{B}(0,r2^k)} \mathbb{P}_x\big\{\tau(\mathcal{B}(0,r2^{k-1})) < \infty\big\} \sup_{x\in\partial\mathcal{B}(0,r2^{k-1})} \mathbb{P}_x\big\{\tau(\mathcal{B}(0,r)) < \tau(H)\big\}.
\end{aligned}$$

Iterating this and letting $R = r2^k$ gives

$$\sup_{x\in\partial\mathcal{B}(0,R)} \mathbb{P}_x\big\{\tau(\mathcal{B}(0,r)) < \tau(H)\big\} \le q^k \sup_{x\in\partial\mathcal{B}(0,R)} \mathbb{P}_x\big\{\tau(\mathcal{B}(0,r)) < \infty\big\},$$

as required to complete the proof. ■

Proof of Theorem 3.45. Let $x, y \in \partial\mathcal{B}(0,r)$ and H be the hyperplane through the origin, which is orthogonal to $x - y$. If $\{B(t) : t \ge 0\}$ is a Brownian motion started in x, define $\{\overline{B}(t) : t \ge 0\}$ the Brownian motion started in y, obtained by defining $\overline{B}(t)$ as the reflection of $B(t)$ at H, for all times $t \le \tau(H)$, and $\overline{B}(t) = B(t)$ for all $t \ge \tau(H)$. This coupling gives, for every $\varepsilon > 0$ and sufficiently large r,

$$\big|\mu_A(x,B) - \mu_A(y,B)\big| \le \mathbb{P}_x\big\{\tau(A) < \tau(H)\big\} \le \varepsilon\mu_A(x,A),$$

using Lemma 3.47 for the last inequality. In particular, we get $|\mu_A(x,A) - \mu_A(y,A)| \le \varepsilon\mu_A(x,A)$. Next, let $|z| > r$ and apply the strong Markov property to obtain

$$\begin{aligned}\frac{\mu_A(x,B)}{\mu_A(x,A)} - \frac{\mu_A(z,B)}{\mu_A(z,A)} &= \int \Big(\frac{\mu_A(x,B)}{\mu_{\mathcal{B}(0,r)}(z,\mathcal{B}(0,r))\mu_A(x,A)} - \frac{\mu_A(y,B)}{\mu_A(z,A)}\Big)\mu_{\mathcal{B}(0,r)}(z,dy)\\ &= \frac{1}{\mu_A(z,A)} \int \Big(\mu_A(x,B)\,\frac{\mu_A(z,A)}{\mu_{\mathcal{B}(0,r)}(z,\mathcal{B}(0,r))\mu_A(x,A)} - \mu_A(y,B)\Big)\mu_{\mathcal{B}(0,r)}(z,dy).\end{aligned}$$

We note that $\mu_A(z,A) = \int \mu_{\mathcal{B}(0,r)}(z,dy)\mu_A(y,A) \le (1+\varepsilon)\mu_{\mathcal{B}(0,r)}(z,\mathcal{B}(0,r))\mu_A(x,A)$, which leads to the estimate

$$\frac{\mu_A(x,B)}{\mu_A(x,A)} - \frac{\mu_A(z,B)}{\mu_A(z,A)} \le \varepsilon + \varepsilon(1+\varepsilon),$$

and the same estimate can be performed with the roles of x and z reversed. As $\varepsilon > 0$ was arbitrary, this implies that $\mu_A(x,B)/\mu_A(x,A)$ converges as $x \to \infty$. ■

Example 3.48. For any ball $\mathcal{B}(x,r)$ we have $\mu_{\mathcal{B}(x,r)} = \varpi_{x,r}$, the uniform distribution. Indeed, $\varpi_{x,r}(\,\cdot\,) = c(R)\int_{\partial\mathcal{B}(x,R)} \mu_{\mathcal{B}(x,r)}(y,\,\cdot\,)d\varpi_{x,R}(y)$, for all $R > r$ and a suitable constant $C(R)$, as the two balls are concentric, and both sides of the equation are rotationally invariant measures on the sphere $\partial\mathcal{B}(x,r)$. Letting $R \uparrow \infty$, we obtain from Theorem 3.45, that $\varpi_{x,r} = \mu_{\mathcal{B}(x,r)}$. ◇

The following surprising proposition shows that the harmonic measure from infinity can also be obtained without this limiting procedure.

Theorem 3.49. *Let $A \subset \mathbb{R}^d$ be compact and suppose $\mathcal{B}(x,r) \supset A$, let $\varpi_{x,r}$ be the uniform distribution on $\partial\mathcal{B}(x,r)$. Then we have, for any Borel set $B \subset A$,*

$$\mu_A(B) = \frac{\int \mu_A(a,B)\,\varpi_{x,r}(da)}{\int \mu_A(a,A)\,\varpi_{x,r}(da)}.$$

Remark 3.50. The surprising fact here is that the right hand side does *not depend* on the choice of the ball $\mathcal{B}(x, r)$. ◇

The crucial observation behind this result is that, starting a Brownian motion in a uniformly chosen point on the boundary of a sphere, the first hitting point of any ball inside that sphere, if it exists, is again uniformly distributed, see Figure 2.

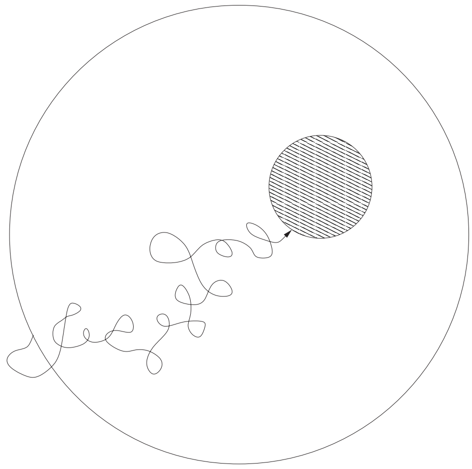

Figure 2. Starting Brownian motion uniformly on the big circle, the distribution of the first hitting point on the small circle is also uniform.

Lemma 3.51. *Let $\mathcal{B}(x, r) \subset \mathcal{B}(y, s)$ and $B \subset \partial\mathcal{B}(x, r)$ Borel. Then*

$$\frac{\int \mu_{\mathcal{B}(x,r)}(a, B)\, \varpi_{y,s}(da)}{\int \mu_{\mathcal{B}(x,r)}(a, \mathcal{B}(x, r))\, \varpi_{y,s}(da)} = \varpi_{x,r}(B).$$

Proof. By Example 3.48 we have $\varpi_{y,s} = \mu_{\mathcal{B}(y,s)}$ and hence, for the normalization constant $c(R) = 1/\int \mu_{\mathcal{B}(y,s)}(a, \mathcal{B}(y, s))\, \varpi_{x,R}(da)$, we have

$$\varpi_{y,s}(\,\cdot\,) = \lim_{R\uparrow\infty} c(R) \int \mu_{\mathcal{B}(y,s)}(a,\,\cdot\,)\, \varpi_{x,R}(da)\,.$$

Hence, for any $B \subset \mathcal{B}(x, r)$ Borel, using the Markov property in the second step,

$$\begin{aligned}\int \mu_{\mathcal{B}(x,r)}(a, B)\, \varpi_{y,s}(da) &= \lim_{R\uparrow\infty} c(R) \int\int \mu_{\mathcal{B}(x,r)}(a, B)\mu_{\mathcal{B}(y,s)}(b, da)\, \varpi_{x,R}(db)\\ &= \lim_{R\uparrow\infty} c(R) \int \mu_{\mathcal{B}(x,r)}(a, B)\, \varpi_{x,R}(da) = \Big(\lim_{R\uparrow\infty} c(R)\Big)\, \varpi_{x,r}(B)\,,\end{aligned}$$

because $\mathcal{B}(x, R)$ and $\mathcal{B}(x, r)$ are concentric. By substituting $B = \mathcal{B}(x, r)$ into the equation, we see that the limit must be equal to the stated constant. ■

Proof of Theorem 3.49. Assume that $\mathcal{B}(x,r)$ and $\mathcal{B}(y,s)$ are two balls containing A. We may then find a ball $\mathcal{B}(z,t)$ containing both these balls. Using Lemma 3.51 and the strong Markov property applied to the first hitting of $\mathcal{B}(x,r)$ we obtain, for any Borel set $B \subset A$,

$$\begin{aligned}\int \mu_A(a,B)\,\varpi_{x,r}(da) &= c_1 \int\!\!\int \mu_A(a,B)\mu_{\mathcal{B}(x,r)}(b,da)\,\varpi_{z,t}(db) = c_1 \int \mu_A(b,B)\,\varpi_{z,t}(db)\\ &= c_2 \int\!\!\int \mu_A(a,B)\mu_{\mathcal{B}(y,s)}(b,da)\,\varpi_{z,t}(db) = c_2 \int \mu_A(a,B)\,\varpi_{y,s}(da),\end{aligned}$$

for suitable constants c_1, c_2 depending only on the choice of the balls. Choosing $B = A$ gives the normalisation constant

$$c_2 = \frac{\int \mu_A(a,A)\,\varpi_{x,r}(da)}{\int \mu_A(a,A)\,\varpi_{y,s}(da)},$$

and this shows that the right hand side in Theorem 3.49 is independent of the choice of the enclosing ball. It therefore must stay constant when its radius goes to infinity, thus completing the proof. ■

Exercises

Exercise 3.1. Show that, if $u\colon U \to \mathbb{R}$ is subharmonic, then

$$u(x) \leq \frac{1}{\mathcal{L}(\mathcal{B}(x,r))} \int_{\mathcal{B}(x,r)} u(y)\, dy \qquad \text{for any ball } \mathcal{B}(x,r) \subset U.$$

Conversely, show that any twice differentiable function $u\colon U \to \mathbb{R}$ satisfying (1.2) is subharmonic. Also give an example of a discontinuous function u satisfying (1.2).

Exercise 3.2 ($*$). Suppose $u\colon \mathcal{B}(x,r) \to \mathbb{R}$ is harmonic and bounded by M. Show that the k^{th} order partial derivatives are bounded by a constant multiple of Mr^{-k}.

Exercise 3.3. Prove the case $d = 1$ in Theorem 3.19.

Exercise 3.4 ($*$). Prove the strong form of the *Paley-Zygmund inequality*:
For any nonnegative random variable X with $\mathbb{E}[X^2] < \infty$ and $\lambda \in [0,1)$,

$$\mathbb{P}\big\{X > \lambda\, \mathbb{E}[X]\big\} \geq (1-\lambda)^2 \, \frac{\mathbb{E}[X]^2}{\mathbb{E}[X^2]}\,.$$

Exercise 3.5. Prove the *Kochen-Stone lemma*:
Suppose $E_1, E_2, \ldots$ are events with

$$\sum_{n=1}^{\infty} \mathbb{P}(E_n) = \infty \qquad \text{and} \qquad \liminf_{k\to\infty} \frac{\sum_{m=1}^{k}\sum_{n=1}^{k} \mathbb{P}(E_n \cap E_m)}{\big(\sum_{n=1}^{k} \mathbb{P}(E_n)\big)^2} < \infty\,.$$

Then, with positive probability, infinitely many of the events take place.
Hint. Apply the Paley-Zygmund inequality to $X = \liminf_{n\to\infty} \mathbb{1}_{E_n}$.

Exercise 3.6 ($*$). Suppose that u is a radial harmonic function on the annulus $D = \{x \in \mathbb{R}^d\colon r < |\,x| < R\}$, where radial means $u(x) = \tilde{u}(|\,x|\,)$ for some function $\tilde{u}\colon (r,R) \to R$ and all x. Suppose further that u is continuous on $\bar{D}$. Show that,

- if $d \geq 3$, there exist constants a and b such that $u(x) = a + b|\,x|^{2-d}$;
- if $d = 2$, there exist constants a and b such that $u(x) = a + b\log|\,x|$.

Exercise 3.7 ($*$). Show that any positive harmonic function on $\mathbb{R}^d$ is constant.

Exercise 3.8 (∗). Let $D \subset \mathbb{R}^d$ be a domain and $x \in D$. Suppose $u\colon D \setminus \{x\} \to \mathbb{R}$ is bounded and harmonic. Show that there exists a unique harmonic continuation $u\colon D \to \mathbb{R}$.

Exercise 3.9. Let $f\colon (0,1) \to (0,\infty)$ with $t \mapsto f(t)/t$ decreasing. Then

$$\int_0^1 f(r)^{d-2} r^{-d/2}\, dr < \infty \qquad \text{if and only if} \qquad \liminf_{t\downarrow 0} \frac{|B(t)|}{f(t)} = \infty \text{ almost surely.}$$

Conversely, if the integral diverges, then $\liminf_{t\downarrow 0} |B(t)|/f(t) = 0$ almost surely.

Exercise 3.10. Show that, if $d \geq 3$ and T is an independent exponential time with parameter $\lambda > 0$, then

$$G(x,y) \sim c(d)\, |\, x - y|^{2-d} \qquad \text{for } |x-y| \downarrow 0\,,$$

where $c(d)$ is as in Theorem 3.32.

Exercise 3.11. Show that

- if $d \geq 2$ and T exponential time with parameter $\lambda > 0$, then $G(\,\cdot\,, y)$ is subharmonic on $\mathbb{R}^d \setminus \{y\}$;

- if $d \geq 2$ and T the first exit time from the domain D, then $G(\,\cdot\,, y)$ is harmonic on $D \setminus \{y\}$.

Exercise 3.12 (∗). Show that if D is a bounded domain, than $G\colon D \times D \setminus \Delta$ is continuous, where $\Delta = \{(x,x) \,:\, x \in D\}$ is the diagonal.

Exercise 3.13 (∗). Find the Green's function for the planar Brownian motion stopped when leaving the domain $\mathcal{B}(0,r)$.

Exercise 3.14 (∗). Suppose $x, y \notin \overline{\mathcal{B}(0,r)}$ and $A \subset \mathcal{B}(0,r)$ is a compact, nonpolar set. Show that $\mu_A(x,\,\cdot\,)$ and $\mu_A(y,\,\cdot\,)$ are mutually absolutely continuous with a density bounded away from zero and infinity.

Notes and Comments

Gauss discusses the Dirichlet problem in [**Ga40**] in a paper on electrostatics. Examples which show that a solution may not exist for certain domains were given by Zaremba [**Za11**] and Lebesgue [**Le24**]. Zaremba's example is the punctured disk we discuss in Example 3.15, and Lebesgue's example is the thorn, which we will discuss in Example 8.32. For domains with smooth boundary the problem was solved by Poincaré [**Po90**].

Bachelier [**Ba00, Ba01**] was the first to note a connection of Brownian motion and the Laplace operator. The first probabilistic approaches to the Dirichlet problem were made by Phillips and Wiener [**PW23**] and Courant, Friedrichs and Lewy [**CFL28**]. These proofs used probability in a discrete setting and approximation. The treatment of the Dirichlet problem using Brownian motion and the probabilistic definition of the harmonic measure are due to the pioneering work of Kakutani [**Ka44a, Ka44b, Ka45**]. A current survey of probabilistic methods in analysis can be found in the book of Bass [**Ba95**], see also Port and Stone [**PS78**] for a classical reference.

Pólya [**Po21**] discovered that a simple symmetric random walk on $\mathbb{Z}^d$ is recurrent for $d \leq 2$ and transient otherwise. His result was later extended to Brownian motion by Lévy [**Le40**] and Kakutani [**Ka44a**]. Neighbourhood recurrence implies, in particular, that the path of a planar Brownian motion (running for an infinite amount of time) is dense in the plane. A more subtle question is whether in $d \geq 3$ all orthogonal projections of a d-dimensional Brownian motion are neighbourhood recurrent, or equivalently whether there is an infinite cylinder avoided by its range. In fact, an avoided cylinder does exist almost surely. This result is due to Adelman, Burdzy and Pemantle [**ABP98**]. The Dvoretzky-Erdős test is originally from [**DE51**] and more information and additional references can be found in [**Pr90**]. There is also an analogous result for planar Brownian motion (with shrinking balls) which is due to Spitzer [**Sp58**].

Green introduced the function named after him in [**Gr28**]. Its probabilistic interpretation appears in Kac's paper [**Ka51**] and is investigated thoroughly by Hunt [**Hu56**]. Quite a lot can be said about the transition densities: $\mathfrak{p}^*(t,\,\cdot\,,\,\cdot\,)$ is jointly continuous on $\overline{D} \times \overline{D}$ and symmetric in the space variables. Moreover, $\mathfrak{p}^*(t,x,y)$ vanishes if either x or y is on the boundary of D, if this boundary is sufficiently regular. This is, of course, only nontrivial in case (3) and full proofs for this case can be found in [**Ba95**] or [**PS78**].

CHAPTER 4

Hausdorff dimension: Techniques and applications

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.

In this chapter we first introduce a suitably general notion of dimension, the Hausdorff dimension. We then describe general techniques to calculate the Hausdorff dimension of arbitrary subsets of $\mathbb{R}^d$, and apply these techniques to the graph and zero set of Brownian motion in dimension one, and to the range of higher dimensional Brownian motion. Lots of further examples will follow in subsequent chapters.

1. Minkowski and Hausdorff dimension

1.1. The Minkowski dimension. How can we capture the dimension of a geometric object? One requirement for a useful definition of dimension is that it should be *intrinsic*. This means that it should be independent of an embedding of the object in an ambient space like $\mathbb{R}^d$. Intrinsic notions of dimension can be defined in arbitrary metric spaces.

Suppose E is a bounded metric space with metric ρ. Here bounded means that the diameter $|E| = \sup\{\rho(x,y) : x, y \in E\}$ of E is finite. The example we have in mind is a bounded subset of $\mathbb{R}^d$. The definition of Minkowski dimension is based on the notion of a *covering* of the metric space E. A **covering** of E is a finite or countable collection of sets

$$E_1, E_2, E_3, \dots \text{ with } E \subset \bigcup_{i=1}^{\infty} E_i\,.$$

Define, for $\varepsilon > 0$,

$$M(E,\varepsilon) = \min\Big\{k \geq 1 : \text{there exists a finite covering } E_1, \dots, E_k \text{ of } E \text{ with } |E_i| \leq \varepsilon \text{ for } i = 1, \dots, k\Big\}\,,$$

where $|A|$ is the diameter of a set $A \subset E$. Intuitively, when E has dimension s the number $M(E,\varepsilon)$ should be of order ε^{-s}. This can be verified in simple cases like line segments, planar squares, etc. This intuition motivates the definition of *Minkowski dimension.*

Definition 4.1. *For a bounded metric space E we define the* **lower Minkowski dimension** *as*

$$\underline{\dim}_M E := \liminf_{\varepsilon \downarrow 0} \frac{\log M(E,\varepsilon)}{\log(1/\varepsilon)}\,,$$

and the **upper Minkowski dimension** *as*

$$\overline{\dim}_M E := \limsup_{\varepsilon \downarrow 0} \frac{\log M(E,\varepsilon)}{\log(1/\varepsilon)}\,.$$

We always have $\underline{\dim}_M E \le \overline{\dim}_M E$*, but equality need not hold. If it holds we write*

$$\dim_M E := \underline{\dim}_M E = \overline{\dim}_M E. \qquad \diamond$$

Remark 4.2. If E is a subset of the unit cube $[0,1]^d \subset \mathbb{R}^d$ then let

$$\tilde{M}_n(E) = \#\big\{Q \in \mathfrak{D}_n \,:\, Q \cap E \neq \emptyset\big\}$$

be the number of dyadic cubes of sidelength 2^{-n} which hit E. Then there exists a constant $C(d) > 0$ depending only on the dimension, such that $\tilde{M}_n(E) \ge M(E, \sqrt{d}\,2^{-n}) \ge C(d)\,\tilde{M}_n(E)$. Hence

$$\overline{\dim}_M E := \limsup_{n\uparrow\infty} \frac{\log \tilde{M}_n(E)}{n \log 2} \qquad \text{and} \qquad \underline{\dim}_M E := \liminf_{n\uparrow\infty} \frac{\log \tilde{M}_n(E)}{n \log 2}\,. \qquad \diamond$$

Example 4.3. In Exercise 4.1, we calculate the Minkowski dimension of a deterministic ‘fractal’, the (ternary) Cantor set,

$$C = \Big\{ \sum_{i=1}^{\infty} \frac{x_i}{3^i} \,:\, x_i \in \{0,2\} \Big\} \subset [0,1]\,.$$

This set is obtained from the unit interval $[0,1]$ by first removing the middle third, and the successively the middle third out of each remaining interval ad infinitum, see Figure 1 for the first three stages of the construction. $\diamond$

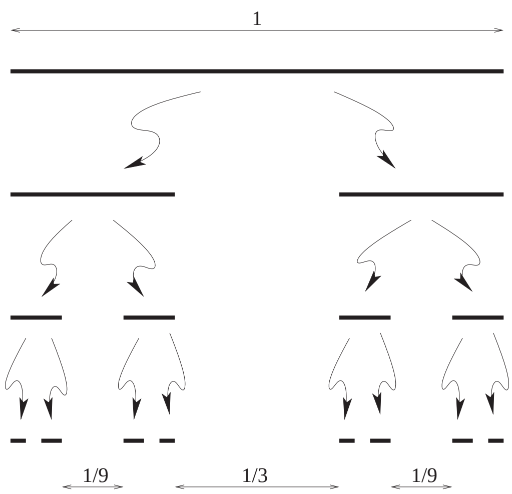

Figure 1. The ternary Cantor set is obtained by removing the middle third from each interval. The figure shows the first three steps of the infinite procedure.

Remark 4.4. There is an unpleasant limitation of Minkowski dimension: Observe that singletons $S = \{x\}$ have Minkowski dimension 0, but we shall see in Exercise 4.2 that the set

$$E := \left\{\tfrac{1}{n} \,:\, n \in \mathbb{N}\right\} \cup \left\{0\right\}$$

has positive dimension. Hence the Minkowski dimension does not have the **countable stability property**

$$\dim \bigcup_{k=1}^{\infty} E_k = \sup \left\{ \dim E_k \,:\, k \geq 1 \right\}.$$

This is one of the properties we expect from a reasonable concept of dimension. There are two ways out of this problem.

(i) One can use a notion of dimension taking variations of the size in the different sets in a covering into account. This captures finer details of the set and leads to the notion of *Hausdorff dimension*.
(ii) One can enforce the countable stability property by subdividing every set in countably many bounded pieces and taking the maximal dimension of them. The infimum over the numbers such obtained leads to the notion of *packing dimension*.

We follow the first route now, but come back to the second route later in the book. $\diamond$

1.2. The Hausdorff dimension. The Hausdorff dimension and Hausdorff measure were introduced by Felix Hausdorff in 1919. Like the Minkowski dimension, Hausdorff dimension can be based on the notion of a covering of the metric space E. For the definition of the Minkowski dimension we have evaluated coverings crudely by counting the number of sets in the covering. Now we also allow infinite coverings and take the size of the covering sets, measured by their diameter, into account.

Looking back at the example of Exercise 4.2 one can see that the set $E = \{1/n \,:\, n \geq 1\} \cup \{0\}$ can be covered much more effectively, if we decrease the size of the balls as we move from right to left. In this example there is a big difference between evaluations of the covering which take into account that we use small sets in the covering, and the evaluation based on just counting the number of sets used to cover.

A very useful evaluation is the **α-value** of a covering. For every $\alpha \geq 0$ and covering $E_1, E_2, \ldots$ we say that the **α-value** of the covering is

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

The terminology of the α-values of a covering allows to formulate a concept of dimension, which is sensitive to the effect that the fine features of this set occur in different scales at different places.

Definition 4.5. *For every $\alpha \geq 0$ the* **α-Hausdorff content** *of a metric space E is defined as*

$$\mathcal{H}^{\alpha}_{\infty}(E) = \inf \left\{ \sum_{i=1}^{\infty} |E_i|^{\alpha} \,:\, E_1, E_2, \ldots \text{ is a covering of } E \right\},$$

informally speaking the α-value of the most efficient covering. If $0 \le \alpha \le \beta$, and $\mathcal{H}^\alpha_\infty(E) = 0$, then also $\mathcal{H}^\beta_\infty(E) = 0$. Thus we can define

$$\dim E = \inf\Big\{\alpha \ge 0 \,:\, \mathcal{H}^\alpha_\infty(E) = 0\Big\} = \sup\Big\{\alpha \ge 0 \,:\, \mathcal{H}^\alpha_\infty(E) > 0\Big\},$$

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

REMARK 4.6. The Hausdorff dimension may, of course, be infinite. But it is easy to see that subsets of $\mathbb{R}^d$ have Hausdorff dimension no larger than d. Moreover, in Exercise 4.3 we show that for every bounded metric space, the Hausdorff dimension is bounded from above by the lower Minkowski dimension. Finally, in Exercise 4.4 we check that Hausdorff dimension has the countable stability property. ◇

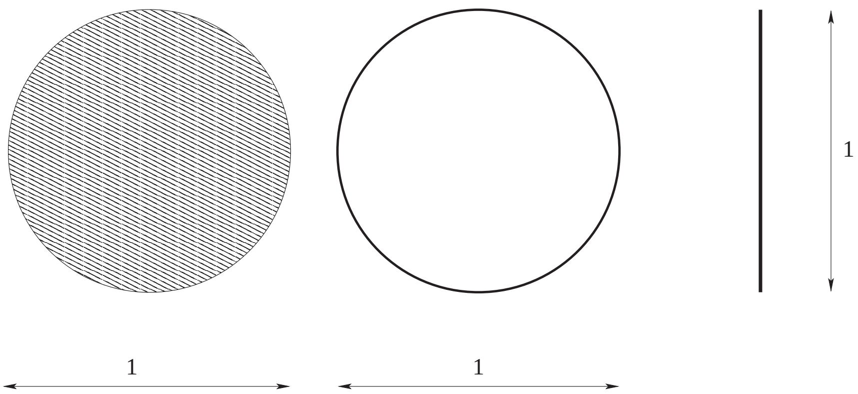

FIGURE 2. The ball, sphere and line segment pictured here all have 1-Hausdorff content equal to one.

The concept of the α-Hausdorff content plays an important part in the definition of the Hausdorff dimension. However, it does not help distinguish the size of sets of the same dimension. For example, the three sets sketched in Figure 2 all have the same 1-Hausdorff content: the ball and the sphere on the left can be covered by a ball of diameter one, so that their 1-Hausdorff content is at most one, but the line segment on the right also does not permit a more effective covering and its 1-Hausdorff content is also 1. Therefore, one considers a refined concept, the *Hausdorff measure.* Here the idea is to consider only coverings by *small* sets.

DEFINITION 4.7. *Let X be a metric space and $E \subset X$. For every $\alpha \ge 0$ and $\delta > 0$ define*

$$\mathcal{H}^\alpha_\delta(E) = \inf\Big\{\sum_{i=1}^\infty |E_i|^\alpha \,:\, E_1, E_2, E_3, \dots \text{ cover } E, \text{ and } |E_i| \le \delta\Big\},$$

i.e. we are considering coverings of E by sets of diameter no more than δ. Then

$$\mathcal{H}^\alpha(E) = \sup_{\delta>0} \mathcal{H}^\alpha_\delta(E) = \lim_{\delta \downarrow 0} \mathcal{H}^\alpha_\delta(E)$$

is the α-**Hausdorff measure** *of the set E.* ◇

Remark 4.8. The α-Hausdorff measure has two obvious properties which, together with $\mathcal{H}^\alpha(\emptyset) = 0$, make it an *outer measure*. These are *countable subadditivity*,

$$\mathcal{H}^\alpha\Big(\bigcup_{i=1}^\infty E_i\Big) \le \sum_{i=1}^\infty \mathcal{H}^\alpha(E_i)\,, \qquad \text{for any sequence } E_1, E_2, E_3, \ldots \subset X\,,$$

and *monotonicity*,

$$\mathcal{H}^\alpha(E) \le \mathcal{H}^\alpha(D), \qquad \text{if } E \subset D \subset X\,.$$

$\diamond$

One can express the Hausdorff dimension in terms of the Hausdorff measure.

Proposition 4.9. *For every metric space E we have*

$$\begin{aligned} \dim E &= \inf\{\alpha : \mathcal{H}^\alpha(E) = 0\} = \inf\{\alpha : \mathcal{H}^\alpha(E) < \infty\} \\ &= \sup\{\alpha : \mathcal{H}^\alpha(E) > 0\} = \sup\{\alpha : \mathcal{H}^\alpha(E) = \infty\}\,. \end{aligned}$$

Proof. We focus on the first equality, all the other arguments are similar. Suppose $\dim E > \alpha$. Then, for all $\beta \le \alpha$, $c := \mathcal{H}^\beta_\infty(E) > 0$, and clearly we have $\mathcal{H}^\beta_\delta(E) \ge c > 0$ for all $\delta > 0$. Hence, $\mathcal{H}^\beta(E) \ge c > 0$ and this implies $\mathcal{H}^\beta(E) > 0$ for all $\beta \le \alpha$. We infer that $\inf\{\beta : \mathcal{H}^\beta(E) = 0\} \ge \alpha$.

Conversely, if $\dim E < \alpha$, then $\mathcal{H}^\alpha_\infty(E) = 0$ and hence, for every $\delta > 0$, there exists a covering by sets $E_1, E_2, \ldots$ with $\sum_{k=1}^\infty |E_k|^\alpha < \delta$. These sets have diameter less than $\delta^{1/\alpha}$, hence $\mathcal{H}^\alpha_{\delta^{1/\alpha}}(E) < \delta$ and letting $\delta \downarrow 0$ yields $\mathcal{H}^\alpha(E) = 0$. This proves $\inf\{\beta : \mathcal{H}^\beta(E) = 0\} \le \alpha$. ■

Remark 4.10. As Lipschitz maps increase the diameter of sets by at most a constant, the image of any set $A \subset E$ under a Lipschitz map has at most the Hausdorff dimension of A. This observation is particularly useful for projections. $\diamond$

A natural generalisation of the last remark arises when we look at the effect of Hölder continuous maps on the Hausdorff dimension.

Definition 4.11. *A function $f\colon (E_1, \rho_1) \to (E_2, \rho_2)$ between metric spaces is called α-Hölder continuous if there exists a (global) constant $C > 0$ such that*

$$\rho_2\big(f(x), f(y)\big) \le C\, \rho_1\big(x, y\big)^\alpha \qquad \textit{for all } x, y \in E_1\,.$$

The constant C is sometimes called the **Hölder constant**. $\diamond$

Remark 4.12. Hölder continuous maps allow some control on the Hausdorff measure of images: We show in Exercise 4.6 that, if $f\colon (E_1, \rho_1) \to (E_2, \rho_2)$ is surjective and α-Hölder continuous with constant C, then for any $\beta \ge 0$,

$$\mathcal{H}^\beta(E_2) \le C^\beta\, \mathcal{H}^{\alpha\beta}(E_1),$$

and therefore $\dim(E_2) \le \frac{1}{\alpha} \dim(E_1)$. $\diamond$

1.3. Upper bounds on the Hausdorff dimension. We now give general upper bounds for the dimension of graph and range of functions, which are based on Hölder continuity.

Definition 4.13. *For a function* $f\colon A \to \mathbb{R}^d$*, for* $A \subset [0,\infty)$*, we define the* **graph** *to be*

$$\mathsf{Graph}_f = \big\{(t, f(t)) \,:\, t \in A\big\} \subset \mathbb{R}^{d+1},$$

and the **range** *or* **path** *to be*

$$\mathsf{Range}_f = f(A) = \big\{f(t) \,:\, t \in A\big\} \subset \mathbb{R}^d. \qquad \diamond$$

Proposition 4.14. *Suppose* $f\colon [0,1] \to \mathbb{R}^d$ *is an* α*-Hölder continuous function. Then*

- **(a)** $\dim(\mathsf{Graph}_f) \le 1 + (1-\alpha)\left(d \wedge \frac{1}{\alpha}\right)$,
- **(b)** *and, for any* $A \subset [0,1]$*, we have* $\dim f(A) \le \frac{\dim A}{\alpha}$.

Proof of (a). Since f is α-Hölder continuous, there exists a constant C such that, if $s,t \in [0,1]$ with $|t-s| \le \varepsilon$, then $|f(t)-f(s)| \le C\varepsilon^\alpha$. Cover $[0,1]$ by no more than $\lceil 1/\varepsilon \rceil$ intervals of length ε. The image of each such interval is then contained in a ball of diameter $C\varepsilon^\alpha$. One can now

- *either* cover each such ball by no more than a constant multiple of $\varepsilon^{d\alpha-d}$ balls of diameter ε,
- *or* use the fact that subintervals of length $(\varepsilon/C)^{1/\alpha}$ in the domain are mapped into balls of diameter ε to cover the image inside the ball by a constant multiple of $\varepsilon^{1-1/\alpha}$ balls of radius ε.

In both cases, look at the cover of the graph consisting of the product of intervals and corresponding balls in $[0,1] \times \mathbb{R}^d$ of diameter ε. The first construction needs a constant multiple of $\varepsilon^{d\alpha-d-1}$ product sets, the second uses $\varepsilon^{-1/\alpha}$ product sets, all of which have diameter of order ε. These coverings give the required upper bounds. ∎

Proof of (b). Suppose that $\dim(A) < \beta < \infty$. Then there exists a covering $A_1, A_2, A_3, \ldots$ such that $A \subset \bigcup_j A_j$ and $\sum_j |A_j|^\beta < \varepsilon$. Then $f(A_1), f(A_2), \ldots$ is a covering of $f(A)$, and $|f(A_j)| \le C|A_j|^\alpha$, where C is the Hölder constant. Thus,

$$\sum_j |f(A_j)|^{\beta/\alpha} \le C^{\beta/\alpha} \sum_j |A_j|^\beta < C^{\beta/\alpha}\varepsilon \to 0$$

as $\varepsilon \downarrow 0$, and hence $\dim f(A) \le \beta/\alpha$. ∎

Remark 4.15. By countable stability of Hausdorff dimension, the statements of Proposition 4.14 remain true if we assume that $f\colon [0,\infty) \to \mathbb{R}^d$ is *locally* α-Hölder continuous. $\diamond$

We now take a first look at dimensional properties of Brownian motion and harvest the results from our general discussion so far. We have shown in Corollary 1.20 that linear Brownian motion is everywhere locally α-Hölder for any $\alpha < 1/2$, almost surely. This extends obviously to d-dimensional Brownian motion, and this allows us to get an upper bound on the Hausdorff dimension of its range and graph.

Corollary 4.16. *The graph of a d-dimensional Brownian motion satisfies, almost surely,*

$$\dim(\mathsf{Graph}) \leq \begin{cases} 3/2 & \text{if } d = 1, \\ 2 & \text{if } d \geq 2\,. \end{cases}$$

For any fixed set $A \subset [0, \infty)$, almost surely

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

Remark 4.17. The corresponding *lower bounds* for the Hausdorff dimension of Graph and Range are more subtle and will be discussed in Section 4.3, when we have more sophisticated tools at our disposal. Our upper bounds also hold for the Minkowski dimension, see Exercise 4.7, and corresponding lower bounds are easier than in the Hausdorff case and obtainable at this stage, see Exercise 4.10. ◇

Corollary 4.16 does not make any statement about the 2-Hausdorff measure of the range, and any such statement requires more information than the Hölder exponent alone can provide, see for example Exercise 4.9. It is however not difficult to show that

$$\mathcal{H}^2\big(B([0,1])\big) < \infty \qquad \text{almost surely.} \tag{1.1}$$

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

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

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

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

using Brownian scaling. The expectation on the right is finite by Theorem 2.39. Hence the expected 2-value of the nth covering is bounded from above by

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

which implies, by Fatou's lemma, that

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

Hence the liminf is almost surely finite, which proves (1.1).

The next theorem improves upon (1.1) by showing that the 2-dimensional Hausdorff measure of the range of d-dimensional Brownian motion is zero for any $d \geq 2$. The proof is considerably more involved and may be skipped on first reading. It makes use of the fact that we have a 'natural' measure on Range at our disposal, which we can use as a tool to pick a good cover by cubes. The idea of using a natural measure supported by the 'fractal' for comparison purposes will also turn out to be crucial for the lower bounds for Hausdorff dimension, which we discuss in the next section.

THEOREM* 4.18. *Let $\{B(t) \colon t \geq 0\}$ be a Brownian motion in dimension $d \geq 2$. Then*

$$\mathcal{H}^2(\mathsf{Range}) = 0 \qquad \textit{almost surely.}$$

Proof. It is sufficient to show that $\mathcal{H}^2(\mathsf{Range}) = 0$ for $d \geq 3$, as 2-dimensional Brownian motion is the projection of 3-dimensional Brownian motion, and projections cannot increase the Hausdorff measure of a set. Moreover it suffices to prove $\mathcal{H}^2(\mathsf{Range} \cap \mathsf{Cube}) = 0$ almost surely, for any half-open cube $\mathsf{Cube} \subset \mathbb{R}^d$ of sidelength one at positive distance from the starting point of the Brownian motion. Without loss of generality we may assume that this cube is the unit cube $\mathsf{Cube} = [0,1)^d$, and our Brownian motion is started at some $x \notin \mathsf{Cube}$.

So let $d \geq 3$, and recall the definition of the (locally finite) occupation measure μ, defined by

$$\mu(A) = \int_0^\infty \mathbb{1}_A(B(s))\, ds\,, \qquad \text{for } A \subset \mathbb{R}^d \text{ Borel.}$$

Let $\mathfrak{D}_k$ be the collection of all cubes $\prod_{i=1}^d [n_i 2^{-k}, (n_i+1)2^{-k})$ where $n_1, \ldots, n_d \in \{0, \ldots, 2^k - 1\}$. We fix a threshold $m \in \mathbb{N}$ and let $M > m$. We call $D \in \mathfrak{D}_k$ with $k \geq m$ a *big* cube if

$$\mu(D) \geq \tfrac{1}{\varepsilon}\, 2^{-2k}\,.$$

The collection $\mathfrak{E}(M)$ consists of all maximal big cubes $D \in \mathfrak{D}_k$, $m \leq k < M$, i.e. all those which are not contained in another big cube, together with all cubes $D \in \mathfrak{D}_M$ which are not contained in a big cube, but intersect Range. Obviously $\mathfrak{E}(M)$ is a cover of $\mathsf{Range} \cap \mathsf{Cube}$ by sets of diameter smaller than $\sqrt{d} 2^{-m}$.

To find the expected 2-value of this cover, first look at a cube $D \in \mathfrak{D}_M$. We denote by $D = D_M \subset D_{M-1} \subset \cdots \subset D_m$ with $D_k \in \mathfrak{D}_k$ the ascending sequence of cubes containing D. Let D_k^* be the cube with the same centre as D_k and $3/2$ its sidelength, see Figure 3.

Let $\tau(D)$ be the first hitting time of the cube D and $\tau_k = \inf\{t > \tau(D) : B(t) \notin D_k^*\}$ be the first exit time from D_k^* for $M > k \geq m$. For the cubes $\mathsf{Cube} = [0,1)^d$ and $\mathsf{Child} = [0, \frac{1}{2})^d$ we also define the expanded cubes Cube^* and Child^* and the stopping time $\tau = \inf\{t > 0 : B(t) \notin \mathsf{Cube}^*\}$. Let

$$q := \sup_{y \in \mathsf{Child}^*} \mathbb{P}_y\Big\{ \int_0^\tau \mathbb{1}_{\mathsf{Cube}}(B(s))\, ds \,\leq \tfrac{1}{\varepsilon} \Big\} < 1\,.$$

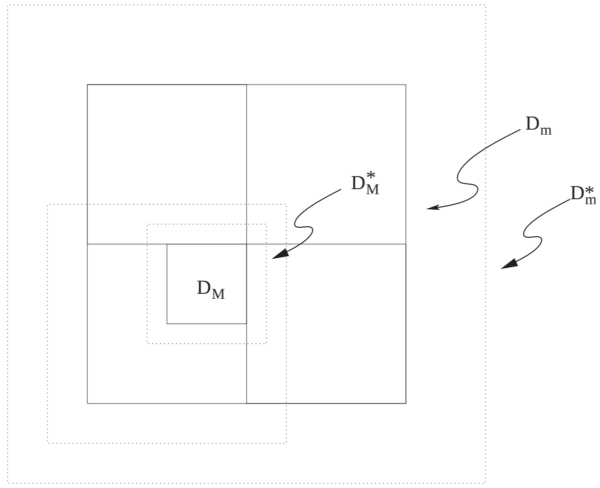

FIGURE 3. Nested systems of cubes, cubes D_k^* indicated by dashed, D_k by solid boundaries.

By the strong Markov property applied to the stopping times $\tau_M < \ldots < \tau_{m+1}$ and Brownian scaling,

$$\begin{aligned}\mathbb{P}_x\big\{\mu(D_k) \le \tfrac{1}{\varepsilon}\, 2^{-2k} \text{ for all } M > k \ge m \,\big|\, \tau(D) < \infty\big\} \\ \le \mathbb{P}_x\Big\{ \int_{\tau_{k+1}}^{\tau_k} \mathbb{1}_{D_k}(B(s))\, ds \le \tfrac{1}{\varepsilon}\, 2^{-2k} \text{ for all } M > k \ge m \,\Big|\, \tau(D) < \infty\Big\} \\ \le \prod_{k=m}^{M-1} \sup_{y\in D_{k+1}^*} \mathbb{P}_y\Big\{ 2^{2k} \int_0^{\tilde\tau_k} \mathbb{1}_{D_k}(B(s))\, ds \le \tfrac{1}{\varepsilon}\Big\} \le q^{M-m},\end{aligned}$$

where $\tilde\tau_k$ is the first exit time of the Brownian motion from D_k^* and the last inequality follows from Brownian scaling. Recall from Theorem 3.17 that $\mathbb{P}_x\{\tau(D) < \infty\} \le c2^{-M(d-2)}$, for a constant $c > 0$ depending only on the dimension d and the fixed distance of x from the unit cube. Hence the probability that any given cube $D \in \mathfrak{D}_M$ is in our cover is

$$\mathbb{P}_x\big\{\mu(D_k) \le \tfrac{1}{\varepsilon}\, 2^{-2k} \text{ for all } M > k \ge m,\ \tau(D) < \infty\big\} \le c2^{-M(d-2)} q^{M-m}.$$

Hence the expected 2-value from the cubes in $\mathfrak{E}(M) \cap \mathfrak{D}_M$ is

$$d2^{dM}2^{-2M}\mathbb{P}_x\big\{\mu(D_k) \le \tfrac{1}{\varepsilon}\, 2^{-2k} \text{ for all } M > k \ge m,\ \tau(D) < \infty\big\} \le cd\, q^{M-m}. \tag{1.2}$$

The 2-value from the cubes in $\mathfrak{E}(M) \cap \bigcup_{k=M+1}^{m} \mathfrak{D}_k$ is bounded by

$$\sum_{k=m}^{M-1} d2^{-2k} \sum_{D\in\mathfrak{E}(M)\cap\mathfrak{D}_k} \mathbb{1}\{\mu(D) \ge 2^{-2k}\tfrac{1}{\varepsilon}\} \le d\varepsilon \sum_{k=m}^{M-1} \sum_{D\in\mathfrak{E}(M)\cap\mathfrak{D}_k} \mu(D) \le d\varepsilon\, \mu(\mathsf{Cube}). \tag{1.3}$$

As $\mathbb{E}\mu(\mathsf{Cube}) < \infty$ by Theorem 3.26, we infer from (1.2) and (1.3) that the expected 2-value of our cover converges to zero for $\varepsilon \downarrow 0$ and a suitable choice $M = M(\varepsilon)$. Hence a subsequence converges to zero almost surely, and, as m was arbitrary, this ensures that $\mathcal{H}^2(\mathsf{Range}) = 0$ almost surely. ∎

2. The mass distribution principle

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 mass distribution principle is a way around this problem, which is based on the existence of a nontrivial 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. For the purpose of this method we call a measure μ on the Borel sets of a metric space E a **mass distribution** on E, if

$$0 < \mu(E) < \infty .$$

The intuition here is that a positive and finite mass is spread over the space E.

THEOREM 4.19 (Mass distribution principle). *Suppose E is a metric space and $\alpha \geq 0$. If there is a mass distribution μ on E and constants $C > 0$ and $\delta > 0$ such that*

$$\mu(V) \leq C|V|^{\alpha} ,$$

for all closed sets V with diameter $|V| \leq \delta$, then

$$\mathcal{H}^{\alpha}(E) \geq \frac{\mu(E)}{C} > 0,$$

and hence $\dim E \geq \alpha$.

Proof. Suppose that $U_1, U_2, \ldots$ is a cover of E by arbitrary sets with $|U_i| \leq \delta$. Let V_i be the closure of U_i and note that $|U_i| = |V_i|$. We have

$$0 < \mu(E) \leq \mu\Big(\bigcup_{i=1}^{\infty} U_i\Big) \leq \mu\Big(\bigcup_{i=1}^{\infty} V_i\Big) \leq \sum_{i=1}^{\infty} \mu(V_i) \leq C \sum_{i=1}^{\infty} |U_i|^{\alpha} .$$

Passing to the infimum over all such covers, and letting $\delta \downarrow 0$ gives the statement. ∎

We now apply this technique to find the Hausdorff dimension of the zero set of a linear Brownian motion. Recall that this is an uncountable set with no isolated points.

At first it is not clear what measure on Zero would be suitable to apply the mass distribution principle. Here Lévy's theorem, see Theorem 2.31, comes to our rescue: Recall the definition of the maximum process $\{M(t) \colon t \geq 0\}$ associated with a Brownian motion from Chapter 2.3.

DEFINITION 4.20. *Let $\{B(t) \colon t \geq 0\}$ be a linear Brownian motion and $\{M(t) \colon t \geq 0\}$ the associated maximum process. A time $t \geq 0$ is a* **record time** *for the Brownian motion if $M(t) = B(t)$ and the set of all record times for the Brownian motion is denoted by* Rec. ◇

Note that the record times are the zeros of the process $\{Y(t) \colon t \geq 0\}$ given by

$$Y(t) = M(t) - B(t).$$

By Theorem 2.31 this process is a reflected Brownian motion, and hence its zero set and the zero set of $\{B(t) : t \geq 0\}$ have the same distribution. A natural measure on Rec is given by the distribution function $\{M(t) : t \geq 0\}$, which allows us to get a lower bound for the Hausdorff dimension of Rec via the mass distribution principle.

LEMMA 4.21. *Almost surely,* $\dim(\mathsf{Zero} \cap [0,1]) = \dim(\mathsf{Rec} \cap [0,1]) \geq 1/2$.

Proof. The first equality follows from Theorem 2.31, so that we can focus in this proof on the record set. Since $t \mapsto M(t)$ is an increasing and continuous function, we can regard it as a distribution function of a positive measure μ, with $\mu(a,b] = M(b) - M(a)$. This measure is obviously supported on the (closed) set Rec of record times. We know that, with probability one, the Brownian motion is locally Hölder continuous with any exponent $\alpha < 1/2$. Thus there exists a (random) constant C_α, such that, almost surely,

$$M(b) - M(a) \leq \max_{0 \leq h \leq b-a} B(a+h) - B(a) \leq C_\alpha (b-a)^\alpha,$$

for all $a, b \in [0,1]$. By the mass distribution principle, we get that, almost surely,

$$\dim(\mathsf{Rec} \cap [0,1]) \geq \alpha.$$

Letting $\alpha \uparrow \frac{1}{2}$ finishes the proof. ∎

To get an upper bound on the Hausdorff dimension of Zero we use a covering consisting of intervals. Define the collection $\mathfrak{D}_k$ of intervals $[j2^{-k}, (j+1)2^{-k})$ for $j = 0, \ldots, 2^k - 1$, and let $Z(I) = 1$ if there exists $t \in I$ with $B(t) = 0$. To estimate the dimension of the zero set we need an estimate for the probability that $Z(I) = 1$, i.e. for the probability that a given interval contains a zero of Brownian motion.

LEMMA 4.22. *There is an absolute constant C such that, for any $a, \varepsilon > 0$,*

$$\mathbb{P}\big\{ \text{ there exists } t \in (a, a+\varepsilon) \text{ with } B(t) = 0 \big\} \leq C \sqrt{\tfrac{\varepsilon}{a+\varepsilon}}.$$

Proof. Consider the event $A = \{|B(a+\varepsilon)| \leq \sqrt{\varepsilon}\}$. By the scaling property of Brownian motion, we can give the upper bound

$$\mathbb{P}(A) = \mathbb{P}\Big\{ |B(1)| \leq \sqrt{\tfrac{\varepsilon}{a+\varepsilon}} \Big\} \leq 2\sqrt{\tfrac{\varepsilon}{a+\varepsilon}}.$$

Knowing that Brownian motion has a zero in $(a, a+\varepsilon)$ makes the event A very likely. Indeed, applying the strong Markov property at the stopping time $T = \inf\{t \geq a : B(t) = 0\}$, we have

$$\mathbb{P}(A) \geq \mathbb{P}\big(A \cap \{0 \in B[a, a+\varepsilon]\}\big) \geq \mathbb{P}\{T \leq a+\varepsilon\} \min_{a \leq t \leq a+\varepsilon} \mathbb{P}\{B(a+\varepsilon) \leq \sqrt{\varepsilon} \,|\, B(t) = 0\}.$$

Clearly the minimum is achieved at $t = a$ and, using the scaling property of Brownian motion, we have $\mathbb{P}\{B(a+\varepsilon) \leq \sqrt{\varepsilon} \,|\, B(a) = 0\} = \mathbb{P}\{|B(1)| \leq 1\} =: c > 0$. Hence,

$$\mathbb{P}\{T \leq a+\varepsilon\} \leq \tfrac{2}{c} \sqrt{\tfrac{\varepsilon}{a+\varepsilon}},$$

and this completes the proof. ∎

Remark 4.23. This is only very crude information about the position of the zeros of a linear Brownian motion. Much more precise information is available, for example in the form of the arcsine law for the last sign-change, which we prove in the next section, and which (after a simple scaling) yields the precise value of the probability in Lemma 4.22. $\diamond$

We have thus shown that, for any $\varepsilon > 0$ and sufficiently large integer k, we have

$$\mathbb{E}[Z(I)] \le c_1\, 2^{-k/2}, \qquad \text{for all } I \in \mathfrak{D}_k \text{ with } I \subset (\varepsilon, 1-\varepsilon)\,,$$

for some constant $c_1 > 0$. Hence the covering of the set $\{t \in (\varepsilon, 1-\varepsilon) \,:\, B(t) = 0\}$ by all $I \in \mathfrak{D}_k$ with $I \cap (\varepsilon, 1-\varepsilon) \neq \emptyset$ and $Z(I) = 1$ has an expected $\frac{1}{2}$-value of

$$\mathbb{E}\Big[\sum_{\substack{I\in\mathfrak{D}_k\\ I\cap(\varepsilon,1-\varepsilon)\neq\emptyset}} \mathbb{1}_{\{Z(I)=1\}}\, 2^{-k/2}\Big] = \sum_{\substack{I\in\mathfrak{D}_k\\ I\cap(\varepsilon,1-\varepsilon)\neq\emptyset}} \mathbb{E}[Z(I)]\, 2^{-k/2} \le c_1\, 2^k\, 2^{-k/2}\, 2^{-k/2} \le c_1.$$

We thus get, from Fatou's lemma,

$$\mathbb{E}\Big[\liminf_{k\to\infty} \sum_{\substack{I\in\mathfrak{D}_k\\ I\cap(\varepsilon,1-\varepsilon)\neq\emptyset}} \mathbb{1}_{\{Z(I)=1\}}\, 2^{-k/2}\Big] \le \liminf_{k\to\infty} \mathbb{E}\Big[\sum_{\substack{I\in\mathfrak{D}_k\\ I\cap(\varepsilon,1-\varepsilon)\neq\emptyset}} \mathbb{1}_{\{Z(I)=1\}}\, 2^{-k/2}\Big] \le c_1.$$

Hence the liminf is almost surely finite, which means that there exists a family of coverings with maximal diameter going to zero and bounded $\frac{1}{2}$-value. This implies that, almost surely,

$$\mathcal{H}^{\frac{1}{2}}\big\{t \in (\varepsilon, 1-\varepsilon) \,:\, B(t) = 0\big\} < \infty,$$

and, in particular, that $\dim(\mathsf{Zero} \cap (\varepsilon, 1-\varepsilon)) \le \frac{1}{2}$. As $\varepsilon > 0$ was arbitrary, we obtain the same bound for the full zero set. Combining this estimate with Lemma 4.21 we have verified the following result.

Theorem 4.24. *Let $\{B(t)\colon 0 \le t \le 1\}$ be a linear Brownian motion. Then, with probability one, we have*

$$\dim\big(\mathsf{Zero} \cap [0,1]\big) = \dim\big(\mathsf{Rec} \cap [0,1]\big) = \tfrac{1}{2}\,.$$

Remark 4.25. As in the case of the Brownian path, the Hausdorff measure is itself *not* a nontrivial measure on the zero set, see Exercise 4.13. In Chapter 6 we shall construct such a measure, the local time at zero. Until then, Lévy's identity will remain the crucial tool. $\diamond$

3. The energy method

The energy method is a technique to find a lower bound for the Hausdorff dimension, which is particularly interesting in applications to random fractals. It replaces the condition on the mass of all closed sets in the mass distribution principle by finiteness of an energy.

Definition 4.26. *Suppose μ is a mass distribution on a metric space (E, ρ) and $\alpha \geq 0$. 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*

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

The simple idea of the energy method is the following: Mass distributions 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 4.27 (Energy method). *Let $\alpha \geq 0$ and μ be a mass distribution on a metric space E. Then, for every $\varepsilon > 0$, we have*

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

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

Remark 4.28. 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 $\dim E \geq \alpha$ almost surely, it suffices to show that $\mathbb{E} I_\alpha(\mu) < \infty$ for a (random) measure on E. $\diamond$

Proof. Suppose that $\{A_n : n = 1, 2, \ldots\}$ is a pairwise disjoint covering of E by sets of diameter $< \varepsilon$. Then

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

Moreover, using the Cauchy-Schwarz inequality,

$$\begin{aligned}\mu(E) &\leq \sum_{n=1}^\infty \mu(A_n) = \sum_{n=1}^\infty |A_n|^{\frac{\alpha}{2}} \frac{\mu(A_n)}{|A_n|^{\frac{\alpha}{2}}} \\ &\leq \sum_{n=1}^\infty |A_n|^\alpha \sum_{n=1}^\infty \frac{\mu(A_n)^2}{|A_n|^\alpha} \leq \mathcal{H}^\alpha_\varepsilon(E) \iint_{\rho(x,y)<\varepsilon} \frac{d\mu(x)\, d\mu(y)}{\rho(x,y)^\alpha}.\end{aligned}$$

Dividing both sides by the integral gives the stated inequality. If $\mathbb{E} I_\alpha(\mu) < \infty$ the integral converges to zero, so that $\mathcal{H}^\alpha_\varepsilon(E)$ diverges to infinity. ∎

We now apply the energy method to resolve questions left open in the first section of this chapter, namely the lower bounds for the Hausdorff dimension of the graph and range of Brownian motion.

The nowhere differentiability of linear Brownian motion established in the first chapter suggests that its graph may have dimension greater than one. For dimensions $d \geq 2$, it is interesting to look at the range of Brownian motion. We have seen that planar Brownian motion is neighbourhood recurrent, that is, it visits every neighbourhood in the plane infinitely often. In this sense, the range of planar Brownian motion is comparable to the plane itself and one can ask whether this is also true in the sense of dimension.

THEOREM 4.29 (Taylor 1953). *Let $\{B(t)\colon t \geq 0\}$ be d-dimensional Brownian motion.*

(a) *If $d = 1$, then* $\dim \mathsf{Graph} = 3/2$ *almost surely.*

(b) *If $d \geq 2$, then* $\dim \mathsf{Range} = \dim \mathsf{Graph} = 2$ *almost surely.*

Recall that we already know the upper bounds from Corollaries 4.16 and 4.16. We now look at lower bounds for the range of Brownian motion in $d \geq 2$.

Proof of Theorem 4.29(b). A natural measure on Range is the occupation measure μ_B defined by $\mu_B(A) = \mathcal{L}(B^{-1}(A) \cap [0,1])$, for all Borel sets $A \subset \mathbb{R}^d$, or, equivalently,

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

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

$$\mathbb{E} \iint \frac{d\mu_B(x)\, d\mu_B(y)}{|x-y|^\alpha} = \mathbb{E} \int_0^1 \int_0^1 \frac{ds\, dt}{|B(t) - B(s)|^\alpha} < \infty. \tag{3.1}$$

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 (3.1) and using Fubini's theorem we get

$$\mathbb{E} I_\alpha(\mu_B) = c \int_0^1 \int_0^1 \frac{ds\, dt}{|t-s|^{\alpha/2}} \leq 2c \int_0^1 \frac{du}{u^{\alpha/2}} < \infty. \tag{3.2}$$

Therefore $I_\alpha(\mu_B) < \infty$ and hence $\dim \mathsf{Range} > \alpha$, almost surely. The lower bound on the range follows by letting $\alpha \uparrow 2$. We also obtain a lower bound for the dimension of the graph: As the graph of a function can be projected onto the path, the dimension of the graph is at least the dimension of the path by Remark 4.10. Hence, if $d \geq 2$, almost surely $\dim \mathsf{Graph} \geq 2$. ∎

Now let us turn to linear Brownian motion and prove the first half of Taylor's theorem.

Proof of Theorem 4.29(a). Again we use the energy method for a sharp lower bound. Recall that we have shown in Corollary 4.16 that $\dim \mathsf{Graph} \leq 3/2$. Let $\alpha < 3/2$ and define a measure μ on the graph by

$$\mu(A) = \mathcal{L}_1(\{0 \leq t \leq 1 : (t, B(t)) \in A\}) \text{ for } A \subset [0,\infty) \times \mathbb{R} \text{ Borel.}$$

Changing variables, the α-energy of μ can be written as

$$\iint \frac{d\mu(x)\, d\mu(y)}{|\, x-y|^\alpha} = \int_0^1 \int_0^1 \frac{ds\, dt}{(|\, t-s|^2 + |\, B(t)-B(s)|^2)^{\alpha/2}}.$$

Bounding the integrand, taking expectations, and applying Fubini we get that

$$\mathbb{E} I_\alpha(\mu) \le 2 \int_0^1 \mathbb{E}\left((t^2 + B(t)^2)^{-\alpha/2}\right) dt. \tag{3.3}$$

Let $\mathfrak{p}(z) = \sqrt{2\pi}^{-1} \exp(-z^2/2)$ denote the standard normal density. By scaling, the expectation above can be written as

$$2 \int_0^{+\infty} (t^2 + tz^2)^{-\alpha/2} \mathfrak{p}(z)\, dz. \tag{3.4}$$

Comparing the size of the summands in the integration suggests separating $z \le \sqrt{t}$ from $z > \sqrt{t}$. Then we can bound (3.4) above by twice

$$\int_0^{\sqrt{t}} (t^2)^{-\alpha/2} dz + \int_{\sqrt{t}}^\infty (tz^2)^{-\alpha/2} \mathfrak{p}(z)\, dz = t^{\frac{1}{2}-\alpha} + t^{-\alpha/2} \int_{\sqrt{t}}^\infty z^{-\alpha} \mathfrak{p}(z)\, dz.$$

Furthermore, we separate the last integral at 1. We get

$$\int_{\sqrt{t}}^\infty z^{-\alpha} \mathfrak{p}(z)\, dz \le 1 + \int_{\sqrt{t}}^1 z^{-\alpha}\, dz.$$

The latter integral is of order $t^{(1-\alpha)/2}$. Substituting these results into (3.3), we see that the expected energy is finite when $\alpha < 3/2$. The claim now follows from the energy method. ∎

4. Frostman's lemma and capacity

In this section we provide a converse to the mass distribution principle, i.e. starting from a lower bound on the Hausdorff measure we construct a mass distribution on a set. This is often useful, for example if one wants to relate the Hausdorff dimension of a set and its image under some transformation.

THEOREM 4.30 (Frostman's lemma). *If $A \subset \mathbb{R}^d$ is a closed set such that $\mathcal{H}^\alpha(A) > 0$, then there exists a Borel probability measure μ supported on A and a constant $C > 0$ such that $\mu(D) \le C|D|^\alpha$ for all Borel sets D.*

We now give a proof of Frostman's lemma, which is based on a tree representation of Euclidean space and a famous result from graph theory, the *max-flow min-cut theorem.* The proof given here is based on the representation of compact subsets of $\mathbb{R}^d$ by trees, an idea that we will encounter again in Chapter 9.

Definition 4.31. *A* **tree** $T = (V, E)$ *is a connected graph described by a finite or countable set* V *of* **vertices***, which includes a distinguished vertex* ϱ *designated as the root, and a set* $E \subset V \times V$ *of ordered* **edges***, such that*

- *for every vertex* $v \in V$ *the set* $\{w \in V : (w, v) \in E\}$ *consists of exactly one element* $\overline{v}$*, the* **parent***, except for the* **root** $\varrho \in V$*, which has no parent;*
- *for every vertex* v *there is a unique self-avoiding path from the root to* v *and the number of edges in this path is the* **order** *or* **generation** $|v|$ *of the vertex* $v \in V$*;*
- *for every* $v \in V$*, the set of* **offspring** *or* **children** *of* $\{w \in V : (v, w) \in E\}$ *is finite.* $\diamond$

Notation 4.32. Suppose $T = (V, E)$ is a tree. For any $v, w \in V$ we denote by $v \wedge w$ the element on the intersection of the paths from the root to v, respectively w with maximal order, i.e. the last common ancestor of v and w. The order $|e|$ of an edge $e = (u, v)$ is the order of its end-vertex v. Every infinite self-avoiding path started in the root is called a **ray**. The set of rays is denoted ∂T, the **boundary** of T. For any two rays ξ and η we define $\xi \wedge \eta$ the vertex in the intersection of the rays, which maximises the order. Note that $|\xi \wedge \eta|$ is the number of edges that two rays ξ and η have in common. The distance between two rays ξ and η is defined to be $|\xi - \eta| := 2^{-|\xi \wedge \eta|}$, and this definition makes the boundary ∂T a metric space. $\diamond$

Remark 4.33. The boundary ∂T of a tree is an interesting fractal in its own right. Its Hausdorff dimension is $\log_2(\operatorname{br} T)$ where $\operatorname{br} T$ is a suitably defined average offspring number. This, together with other interesting probabilistic aspects of trees, is discussed in depth in [**LP05**]. $\diamond$

Definition 4.34. *Suppose* capacities *are assigned to the edges of a tree* T*, i.e. there is a mapping* $C : E \to [0, \infty)$*. A* **flow** *of strength* $c > 0$ *through a tree with capacities* C *is a mapping* $\theta : E \to [0, c]$ *such that*

- *for the root we have* $\sum_{\overline{w} = \varrho} \theta(\varrho, w) = c$*, for every other vertex* $v \neq \varrho$ *we have*
$$\theta(\overline{v}, v) = \sum_{w : \overline{w} = v} \theta(v, w),$$
i.e. the flow into and out of each vertex other than the root is conserved.
- $\theta(e) \leq C(e)$*, i.e. the flow through the edge* e *is bounded by its capacity.*

A set Π *of edges is called a* **cutset** *if every ray includes an edge from* Π*.* $\diamond$

The key to the mass distribution principle is a famous result of graph theory, the *max-flow min-cut theorem* of Ford and Fulkerson [**FF56**], which we prove in Section 4 of Appendix II.

THEOREM 4.35 (Max-flow min-cut theorem).

$$\max\Big\{\text{strength}\,(\theta)\ :\ \theta \textit{ a flow with capacities } C\Big\} = \inf\Big\{\sum_{e\in\Pi} C(e)\ :\ \Pi \textit{ a cutset }\Big\}.$$

Proof of Frostman's lemma. We may assume $A \subset [0,1]^d$. Any compact cube in $\mathbb{R}^d$ of sidelength s can be split into 2^d nonoverlapping compact cubes of side length $s/2$. We first create a tree with a root that we associate with the cube $[0,1]^d$. Every vertex in the tree has 2^d edges emanating from it, each leading to a vertex that is associated with one of the 2^d subcubes with half the sidelength of the original cube. We then erase the edges ending in vertices associated with subcubes that do not intersect A. In this way we construct a tree $T = (V,E)$ such that the rays in ∂T correspond to sequences of nested compact cubes, see Figure 4.

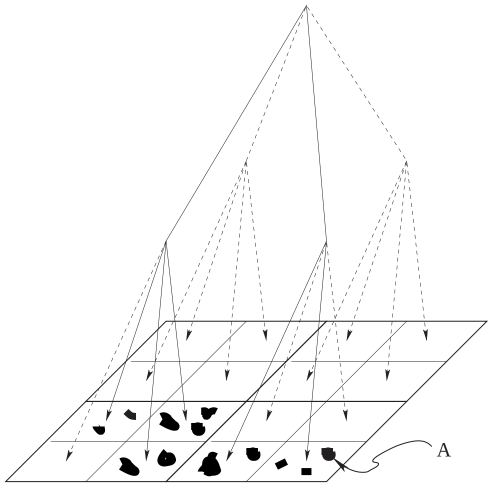

FIGURE 4. The first two stages in the construction of the tree associated with the shaded set $A \subset [0,1]^2$. Dotted edges in the tree are erased.

There is a canonical mapping $\Phi\colon \partial T \to A$, which maps sequences of nested cubes to their intersection. Note that if $x \in A$, then there is an infinite path emanating from the root, all of whose vertices are associated with cubes that contain x and thus intersect A. Hence Φ is surjective.

For any edge e at level n define the capacity $C(e) = 2^{-n\alpha}$. We now associate to every cutset Π a covering of A, consisting of those cubes associated with the initial vertices of the edges in the cutset. To see that the resulting collection of cubes is indeed a covering, let ξ be a ray. As Π is a cutset, it contains one of the edges in this ray, and the cube associated with the initial vertex of this edge contains the point $\Phi(\xi)$. Hence we indeed cover the entire set $\Phi(\partial T) = A$. This implies that

$$\inf\Big\{\sum_{e\in\Pi} C(e)\ :\ \Pi \text{ a cutset }\Big\} \geq \inf\Big\{\sum_j |A_j|^\alpha\ :\ A \subset \bigcup_j A_j\Big\},$$

and as $\mathcal{H}^\alpha_\infty(A) > 0$ this is bounded from zero. Thus, by the max-flow min-cut theorem, there exists a flow $\theta\colon E \to [0,\infty)$ of positive strength such that $\theta(e) \leq C(e)$ for all edges $e \in E$.

We now show how to define a suitable measure on the space of infinite paths. Given an edge $e \in E$ we associate a set $T(e) \subset \partial T$ consisting of all rays containing the edge e. Define

$$\widetilde{\nu}\big(T(e)\big) = \theta(e).$$

It is easily checked that the collection $\mathcal{C}(\partial T)$ of subsets $T(v) \subset \partial T$ for all $v \in T$ is a semi-algebra on ∂T. Recall that this means that if $A, B \in \mathcal{C}(\partial T)$, then $A \cap B \in \mathcal{C}(\partial T)$ and A^c is a finite disjoint union of sets in $\mathcal{C}(\partial T)$. Because the flow through any vertex is preserved, $\widetilde{\nu}$ is countably additive. Thus, using a measure extension theorem such as, for example [**Du95**, A.1(1.3)], we can extend $\widetilde{\nu}$ to a measure ν on the σ-algebra generated by $\mathcal{C}(\partial T)$. We can now define a Borel measure $\mu = \nu \circ \Phi^{-1}$ on A, which satisfies $\mu(C) = \theta(e)$, where C is the cube associated with the initial vertex of the edge e. Suppose now that D is a Borel subset of $\mathbb{R}^d$ and n is the integer such that $2^{-n} < |D \cap [0,1]^d| \le 2^{-(n-1)}$. Then $D \cap [0,1]^d$ can be covered with 3^d of the cubes in the above construction having side length 2^{-n}. Using the bound, we have

$$\mu(D) \le d^{\frac{\alpha}{2}}\, 3^d 2^{-n\alpha} \le d^{\frac{\alpha}{2}}\, 3^d |D|^\alpha,$$

so we have a finite measure μ satisfying the requirement of the lemma. Normalising μ to get a probability measure completes the proof. ■

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

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

Theorem 4.27 states that a set of positive α-capacity has dimension at least α. We now show that, in this formulation the method is sharp. Our proof of this fact relies on Frostman's lemma and hence refers to closed subsets of Euclidean space.

THEOREM 4.36. *For any closed set $A \subset \mathbb{R}^d$,*

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

Proof. It only remains to show $\le$, and for this purpose it suffices to show that if $\dim A > \alpha$, then there exists a Borel probability measure μ on A such that

$$I_\alpha(\mu) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{d\mu(x)\, d\mu(y)}{|\,x-y|^\alpha} < \infty.$$

By our assumption for some sufficiently small $\beta > \alpha$ we have $\mathcal{H}^\beta(A) > 0$. By Frostman's lemma, there exists a nonzero Borel probability measure μ on A and a constant C such that $\mu(D) \le C|D|^\beta$ for all Borel sets D. By restricting μ to a smaller set if necessary, we can make the support of μ have diameter less than one. Fix $x \in A$, and for $k \ge 1$ let $S_k(x) = \{y : 2^{-k} < |\,x-y| \le 2^{1-k}\}$. Since μ has no atoms, we have

$$\int_{\mathbb{R}^d} \frac{d\mu(y)}{|\,x-y|^\alpha} = \sum_{k=1}^\infty \int_{S_k(x)} \frac{d\mu(y)}{|\,x-y|^\alpha} \le \sum_{k=1}^\infty \mu(S_k(x)) 2^{k\alpha},$$

where the equality follows from the monotone convergence theorem and the inequality holds by the definition of the S_k. Also,

$$\sum_{k=1}^{\infty} \mu(S_k(x))2^{k\alpha} \le C \sum_{k=1}^{\infty} |2^{2-k}|^{\beta} 2^{k\alpha} = C' \sum_{k=1}^{\infty} 2^{k(\alpha-\beta)},$$

where $C' = 2^{2\beta}C$. Since $\beta > \alpha$, we have

$$I_\alpha(\mu) \le C' \sum_{k=1}^{\infty} 2^{k(\alpha-\beta)} < \infty,$$

which proves the theorem. ■

In Corollary 4.16 we have seen that the image of a set $A \subset [0,\infty)$ under Brownian motion has at most twice the Hausdorff dimension of A. Naturally, the question arises whether this is a sharp estimate. The following result of McKean shows that, if $d \ge 2$, this is sharp for *any* set A, while in $d = 1$ it is sharp as long as $\dim A \le \frac{1}{2}$.

THEOREM 4.37 (McKean 1955). *Let $A \subset [0,\infty)$ be a closed subset and $\{B(t) : t \ge 0\}$ a d-dimensional Brownian motion. Then, almost surely,*

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

Proof. The upper bound was verified in Corollary 4.16. For the lower bound let $\alpha < \dim(A) \wedge (d/2)$. By Theorem 4.36, there exists a Borel probability measure μ on A such that $I_\alpha(\mu) < \infty$. Denote by μ_B the measure defined by

$$\mu_B(D) = \mu(\{t \ge 0 : B(t) \in D\})$$

for all Borel sets $D \subset \mathbb{R}^d$. Then

$$\mathbb{E}[I_{2\alpha}(\mu_B)] = \mathbb{E}\left[\iint \frac{d\mu_B(x)\, d\mu_B(y)}{|x-y|^{2\alpha}}\right] = \mathbb{E}\left[\int_0^\infty \int_0^\infty \frac{d\mu(t)\, d\mu(s)}{|B(t)-B(s)|^{2\alpha}}\right],$$

where the second equality can be verified by a change of variables. Note that the denominator on the right hand side has the same distribution as $|t-s|^\alpha |Z|^{2\alpha}$, where Z is a d-dimensional standard normal random variable. Since $2\alpha < d$, we have that

$$\mathbb{E}[|Z|^{-2\alpha}] = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} |y|^{-2\alpha} e^{-|y|^2/2}\, dy < \infty.$$

Hence, using Fubini's theorem,

$$\mathbb{E}[I_{2\alpha}(\mu_B)] = \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{E}[|Z|^{-2\alpha}] \frac{d\mu(t)\, d\mu(s)}{|t-s|^\alpha} \le \mathbb{E}[|Z|^{-2\alpha}] I_\alpha(\mu) < \infty.$$

Thus, $\mathbb{E}[I_{2\alpha}(\mu_B)] < \infty$, and hence $I_{2\alpha}(\mu_B) < \infty$ almost surely. Moreover, μ_B is supported on $B(A)$ because μ is supported on A. It follows from Theorem 4.27 that $\dim B(A) \ge 2\alpha$ almost surely. By letting $\alpha \uparrow \dim(A) \wedge d/2$, we see that $\dim(B(A)) \ge 2\dim(A) \wedge d$ almost surely. This completes the proof of Theorem 4.37. ■

Remark 4.38. We have indeed shown that, if $\mathrm{Cap}_\alpha(A) > 0$, then $\mathrm{Cap}_{2\alpha}(B(A)) > 0$ almost surely. The converse of this statement is also true and will be discussed later, see Theorem 9.36. $\diamond$

Remark 4.39. Later in the book, we shall be able to significantly improve McKean's theorem and show that for Brownian motion in dimension $d \geq 2$, almost surely, for any $A \subset [0, \infty)$, we have $\dim B(A) = 2 \dim(A)$. This result is Kaufman's theorem, see Theorem 9.28. Note the difference between the results of McKean and Kaufman: In Theorem 4.37, the null probability set depends on A, while Kaufman's theorem has a much stronger claim: it states dimension doubling simultaneously for *all* sets. This allows us to plug in random sets A, which may depend completely arbitrarily on the Brownian motion. For Kaufman's theorem, $d \geq 2$ is a necessary condition: we have seen that the zero set of one dimensional Brownian motion has dimension $1/2$, while its image is a single point. $\diamond$

Exercises

Exercise 4.1 (∗). Show that for the ternary Cantor set C, we have $\dim_M C = \frac{\log 2}{\log 3}$.

Exercise 4.2 (∗). Let $E := \{1/n : n \in \mathbb{N}\} \cup \{0\}$. Then $\dim_M E = \frac{1}{2}$.

Exercise 4.3 (∗). Show that, for every bounded metric space, the Hausdorff dimension is bounded from above by the lower Minkowski dimension.

Exercise 4.4 (∗). Show that Hausdorff dimension has the countable stability property.

Exercise 4.5. Show that, for the ternary Cantor set C we have $\dim C = \frac{\log 2}{\log 3}$.

Exercise 4.6 (∗). If $f\colon (E_1, \rho_1) \to (E_2, \rho_2)$ is surjective and α-Hölder continuous with constant C, then for any $\beta \geq 0$,

$$\mathcal{H}^\beta(E_2) \leq C^\beta \, \mathcal{H}^{\alpha\beta}(E_1),$$

and therefore $\dim(E_2) \leq \frac{1}{\alpha} \dim(E_1)$.

Exercise 4.7. Suppose $f\colon [0,1] \to \mathbb{R}^d$ is an α-Hölder continuous function. Then

(a) $\overline{\dim}_{\mathrm{M}}(\mathsf{Graph}_f) \leq 1 + (1-\alpha)\left(d \wedge \frac{1}{\alpha}\right)$,

(b) and, for any $A \subset [0,1]$, we have $\overline{\dim}_{\mathrm{M}} f(A) \leq \frac{\overline{\dim}_{\mathrm{M}} A}{\alpha}$.

Exercise 4.8 (∗). For any integer $d \geq 1$ and $0 < \alpha < d$ construct a compact set $A \subset \mathbb{R}^d$ such that $\dim A = \alpha$.

Exercise 4.9. Construct a function $f\colon [0,1] \to \mathbb{R}^d$ which is α-Hölder continuous for any $\alpha < \beta$, but has $\mathcal{H}^\beta(f[0,1]) = \infty$.

Exercise 4.10. A function $f\colon [0,1] \to \mathbb{R}$ is called **reverse β-Hölder** for some $0 < \beta < 1$ if there exists a constant $C > 0$ such that for any interval $[t,s]$, there is a subinterval $[t_1, s_1] \subset [t,s]$, such that $|f(t_1) - f(s_1)| \geq C|t-s|^\beta$. Let $f\colon [0,1] \to \mathbb{R}$ be reverse β-Hölder. Then $\dim_{\mathrm{M}}(\mathsf{Graph}_f) \geq 2 - \beta$.

Exercise 4.11. Show that for $\{B(t) : 0 \leq t \leq 1\}$ we have $\dim_M \mathsf{Graph} = \frac{3}{2}$ if $d = 1$, and $\dim_{\mathrm{M}} \mathsf{Graph} = \dim_M B[0,1] = 2$ if $d \geq 2$, almost surely.

Exercise 4.12. Show that $\dim_{\mathrm{M}} \{0 \leq t \leq 1 : B(t) = 0\} = \frac{1}{2}$, almost surely.

Exercise 4.13 (∗). Show that $\mathcal{H}^{1/2}(\mathsf{Zero}) = 0$, almost surely.

Notes and Comments

Felix Hausdorff introduced the Hausdorff measure in his seminal paper [**Ha19**]. Credit should also be given to Carathéodory [**Ca14**] who introduced a general construction in which Hausdorff measure can be naturally embedded. The Hausdorff measure indeed defines a measure on the Borel sets, proofs can be found in [**Ma95**] and [**Ro99**]. If $X = \mathbb{R}^d$ and $\alpha = d$ the Hausdorff measure $\mathcal{H}^\alpha$ is a constant multiple of Lebesgue measure $\mathcal{L}_d$, moreover if α is an integer and X an embedded α-submanifold, then $\mathcal{H}^\alpha$ is the surface measure. This idea can also be used to develop vector analysis on sets with much less smoothness than a differentiable manifold. For more about Hausdorff dimension and geometric questions related to it we strongly recommend Mattila [**Ma95**]. The classic text of Rogers [**Ro99**], which first appeared in 1970, is a thorough discussion of Hausdorff measures. Falconer [**Fa97a, Fa97b**] covers a range of applications and current developments, but with more focus on deterministic fractals.

The results on the Hausdorff dimension of graph and range of a Brownian motion are due to S.J. Taylor [**Ta53, Ta55**] and independently to Lévy [**Le51**] though the latter paper does not contain full proofs. Taylor also proved in [**Ta55**] that the dimension of the zero set of a Brownian motion in dimension one is $1/2$. Stronger results show that, almost surely, the Hausdorff dimension of *all* nontrivial level sets is $1/2$. For this and much finer results see [**Pe81**]. A classical survey, which inspired a lot of activity in the area of Hausdorff dimension and stochastic processes is [**Ta86**] and a modern survey is [**Xi04**].

The energy method and Frostman's lemma all stem from Otto Frostman's famous 1935 thesis [**Fr35**], which lays the foundations of modern potential theory. The elegant quantitative proof of the energy method given here is due to Oded Schramm. Frostman's lemma was generalised to complete, separable metric spaces by Howroyd [**Ho95**] using a functional-analytic approach. The main difficulty arising in the proof is that, if $\mathcal{H}^\alpha(E) = \infty$, one has to find a subset $A \subset E$ with $0 < \mathcal{H}^\alpha(A) < \infty$, which is tricky to do in abstract metric spaces. Frostman's original proof uses, in a way, the same idea as the proof presented here, though the transfer to the tree setup is not done explicitly. Probability using trees became fashionable in the 1990s and indeed, this is the right way to look at many problems of Hausdorff dimension and fractal geometry. Survey articles are [**Pe95**] and [**Ly96**], more information can be found in [**Pe99**] and [**LP05**].

McKean's theorem is due to H.P. McKean jr. [**McK55**]. Its surprising extension by Kaufman is not as hard as one might think considering the wide applicability of the result. The original source is [**Ka69**], we discuss the result in depth in Chapter 9.

The concept of 'reverse Hölder' mappings only partially extends from Minkowski to Hausdorff dimension. If $f\colon [0,1] \to \mathbb{R}$ is both β-Hölder and reverse β-Hölder for some $0 < \beta < 1$, it satisfies $\dim(\mathsf{Graph}_f) > 1$, see Przytycki and Urbański [**PU89**]. For example, the Weierstrass nowhere differentiable function, defined by $W(t) = \sum_{n=0}^\infty a^n \cos(b^n t)$, for $ab > 1$, $0 < a < 1$ is β-Hölder and reverse β-Hölder for some $0 < \beta < 1$. The Hausdorff dimension of its graph is, however, not rigorously known in general.

There is a natural refinement of the notions of Hausdorff dimension and Hausdorff measure, which is based on evaluating sets by applying an arbitrary 'gauge' function φ to the diameter, rather than taking a power. Measuring sets using a gauge function not only allows much finer results, it also turns out that the natural measures on graph and range of Brownian paths, which we have encountered in this chapter, turn out to be Hausdorff measures for suitable gauge functions. Results in this direction are [**CT62, Ra63a, Ta64**] and we include elements of this discussion in Chapter 6, where the zero set of Brownian motion is considered.

CHAPTER 5

Brownian motion and random walk

In this chapter we discuss some aspects of the relation between random walk and Brownian motion. The first two sections aim to demonstrate the nature of this relation by examples, which are of interest in their own right. These are *first* the law of the iterated logarithm, which is easier to prove for Brownian motion and can be extended to random walks by an embedding argument, and *second* a proof that Brownian motion does not have points of increase, which is based on a combinatorial argument for a class of random walks, and then extended to Brownian motion. We then discuss the Skorokhod embedding problem systematically, and give a proof of the Donsker invariance principle based on the Skorokhod embedding. We give a variety of applications of Donsker's theorem, including the arcsine laws.

1. The law of the iterated logarithm

For a standard linear Brownian motion $\{B(t) : t \geq 0\}$, although at any given time t and for any open set $U \subset \mathbb{R}$ the probability of the event $\{B(t) \in U\}$ is positive, over a long time Brownian motion cannot grow arbitrarily fast. We have seen in Corollary 1.11 that, for any small $\varepsilon > 0$, almost surely, there exists $t_0 > 0$ such that $|B(t)| \leq \varepsilon t$ for all $t \geq t_0$, whereas Proposition 1.23 ensures that for every large k, almost surely, there exist arbitrarily large times t such that $|B(t)| \geq k\sqrt{t}$. It is therefore natural to ask for the asymptotic *smallest upper envelope* of the Brownian motion, i.e. for a function $\psi\colon (1,\infty) \to \mathbb{R}$ such that

$$\limsup_{t\to\infty} \frac{B(t)}{\psi(t)} = 1.$$

The law of the iterated logarithm (whose name comes from the answer to this question but is by now firmly established for this type of upper-envelope results) provides such a 'gauge' function, which determines the almost-sure *asymptotic growth* of a Brownian motion.

A similar problem arises for arbitrary random walks $\{S_n : n \geq 0\}$, where we ask for a sequence $(a_n : n \geq 0)$ such that

$$\limsup_{n\to\infty} \frac{S_n}{a_n} = 1.$$

These two questions are closely related, and we start with an answer to the first one.

Theorem 5.1 (Law of the Iterated Logarithm for Brownian motion). *Suppose $\{B(t) : t \geq 0\}$ is a standard linear Brownian motion. Then, almost surely,*

$$\limsup_{t\to\infty} \frac{B(t)}{\sqrt{2t \log\log(t)}} = 1.$$

REMARK 5.2. By symmetry it follows that, almost surely,

$$\liminf_{t\to\infty} \frac{B(t)}{\sqrt{2t\log\log(t)}} = -1.$$

Hence, for any $\varepsilon > 0$, there exists t_0 such that $|B(t)| \le (1+\varepsilon)\sqrt{2t\log\log(t)}$ for any $t \ge t_0$, while there exist arbitrarily large times t with $|B(t)| \ge (1-\varepsilon)\sqrt{2t\log\log(t)}$. ◇

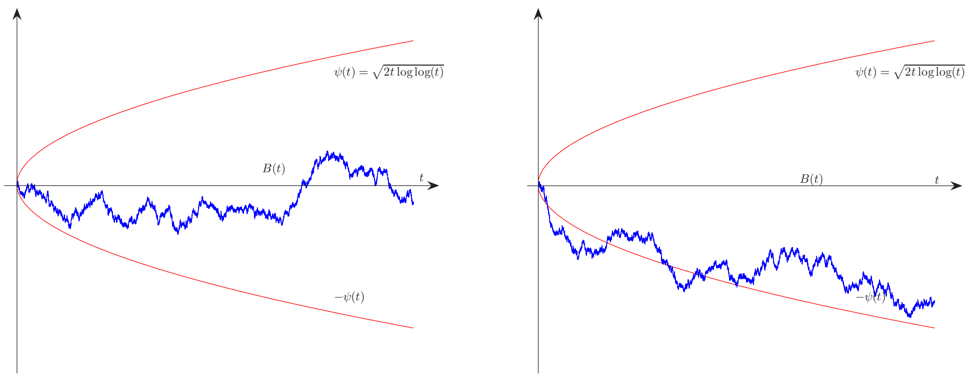

FIGURE 1. The picture on the left shows the asymptotic upper envelope $\psi(t) = \sqrt{2t\log\log(t)}$ and a typical Brownian path indicating that scales where the path comes near to the envelope are very sparse. The picture on the right shows Brownian motion at such a scale. Due to the special nature of this scale the Brownian path (which is implicitly conditioned on the event of ending up near the upper envelope) has features untypical of Brownian paths. See the 'Notes and Comments' section for more details.

Proof. The main idea is to scale by a geometric sequence. Let $\psi(t) = \sqrt{2t\log\log(t)}$. We first prove the upper bound. Fix $\varepsilon > 0$ and $q > 1$. Let

$$A_n = \left\{ \max_{0\le t\le q^n} B(t) \ge (1+\varepsilon)\psi(q^n) \right\}.$$

By Theorem 2.18 the maximum of Brownian motion up to a fixed time t has the same distribution as $|B(t)|$. Therefore

$$\mathbb{P}(A_n) = \mathbb{P}\left\{ \frac{|B(q^n)|}{\sqrt{q^n}} \ge (1+\varepsilon)\,\frac{\psi(q^n)}{\sqrt{q^n}} \right\}.$$

We can use the tail estimate $\mathbb{P}\{Z > x\} \le e^{-x^2/2}$ for a standard normally distributed Z and $x > 1$, see Lemma II.3.1, to conclude that, for large n,

$$\mathbb{P}(A_n) \le \exp\left(-(1+\varepsilon)^2 \log\log q^n\right) = \frac{1}{(n\log q)^{(1+\varepsilon)^2}}.$$

This is summable in n and hence, by the Borel-Cantelli lemma, we get that only finitely many of these events occur. For large t write $q^{n-1} \le t < q^n$. We have

$$\frac{B(t)}{\psi(t)} = \frac{B(t)}{\psi(q^n)}\,\frac{\psi(q^n)}{q^n}\,\frac{t}{\psi(t)}\,\frac{q^n}{t} \le (1+\varepsilon)q,$$

since $\psi(t)/t$ is decreasing in t. Thus

$$\limsup_{t\to\infty} \frac{B(t)}{\psi(t)} \leq (1+\varepsilon)q, \text{ almost surely.}$$

Since this holds for any $\varepsilon > 0$ and $q > 1$ we have proved that $\limsup B(t)/\psi(t) \leq 1$.

For the lower bound, fix $q > 1$. In order to use the Borel-Cantelli lemma in the other direction, we need to create a sequence of *independent* events. Let

$$D_n = \big\{B(q^n) - B(q^{n-1}) \geq \psi(q^n - q^{n-1})\big\}\,.$$

We now use Lemma II.3.1 to see that there is a constant $c > 0$ such that, for large x,

$$\mathbb{P}\{Z > x\} \geq \frac{ce^{-x^2/2}}{x}.$$

Using this estimate we get, for some further constant $\tilde{c} > 0$,

$$\mathbb{P}(D_n) = \mathbb{P}\Big\{Z \geq \tfrac{\psi(q^n - q^{n-1})}{\sqrt{q^n-q^{n-1}}}\Big\} \geq c\frac{e^{-\log\log(q^n-q^{n-1})}}{\sqrt{2\log\log(q^n - q^{n-1})}} \geq \frac{ce^{-\log(n\log q)}}{\sqrt{2\log(n\log q)}} > \frac{\tilde{c}}{n\log n},$$

and therefore $\sum_n \mathbb{P}(D_n) = \infty$. Thus for infinitely many n

$$B(q^n) \geq B(q^{n-1}) + \psi(q^n - q^{n-1}) \geq -2\psi(q^{n-1}) + \psi(q^n - q^{n-1}),$$

where the second inequality follows from applying the previously proved upper bound to $-B(q^{n-1})$. From the above we get that, for infinitely many n,

$$\frac{B(q^n)}{\psi(q^n)} \geq \frac{-2\psi(q^{n-1}) + \psi(q^n - q^{n-1})}{\psi(q^n)} \geq \frac{-2}{\sqrt{q}} + \frac{q^n - q^{n-1}}{q^n} = 1 - \frac{2}{\sqrt{q}} - \frac{1}{q}. \tag{1.1}$$

Indeed, to obtain the second inequality first note that

$$\frac{\psi(q^{n-1})}{\psi(q^n)} = \frac{\psi(q^{n-1})}{\sqrt{q^{n-1}}}\,\frac{\sqrt{q^n}}{\psi(q^n)}\,\frac{1}{\sqrt{q}} \leq \frac{1}{\sqrt{q}},$$

since $\psi(t)/\sqrt{t}$ is increasing in t for large t. For the second term we just use the fact that $\psi(t)/t$ is decreasing in t. Now (1.1) implies that

$$\limsup_{t\to\infty} \frac{B(t)}{\psi(t)} \geq -\frac{2}{\sqrt{q}} + 1 - \frac{1}{q} \text{ almost surely,}$$

and letting $q \uparrow \infty$ concludes the proof of the lower bound. ∎

Corollary 5.3. *Suppose $\{B(t) : t \geq 0\}$ is a standard Brownian motion. Then, almost surely,*

$$\limsup_{h\downarrow 0} \frac{|B(h)|}{\sqrt{2h\log\log(1/h)}} = 1.$$

Proof. By Theorem 1.9 the process $\{X(t) : t \geq 0\}$ defined by $X(t) = tB(1/t)$ for $t > 0$ is a standard Brownian motion. Hence, using Theorem 5.1, we get

$$\limsup_{h\downarrow 0} \frac{|B(h)|}{\sqrt{2h\log\log(1/h)}} = \limsup_{t\uparrow 0} \frac{|X(t)|}{\sqrt{2t\log\log t}} = 1\,. \qquad ∎$$

The law of the iterated logarithm is a result which is easier to prove for Brownian motion than for random walks, as scaling arguments can be used to good effect in the proof. We now use an ad hoc argument to obtain a law of the iterated logarithm for simple random walks, i.e. the random walk with increments taking the values ± 1 with equal probability, from Theorem 5.1. A version for more general walks will follow with analogous arguments from the embedding techniques of Section 3, see Theorem 5.17.

THEOREM 5.4 (Law of the Iterated Logarithm for simple random walk). *Let $\{S_n : n \geq 0\}$ be a simple random walk. Then, almost surely,*

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2n\log\log n}} = 1.$$

We now start the technical work to transfer the result from Brownian motion to simple random walk. The next result shows that the limsup does not change if we only look along a sufficiently dense sequence of random times. We abbreviate $\psi(t) = \sqrt{2t\log\log(t)}$.

LEMMA 5.5. *If $\{T_n : n \geq 1\}$ is a sequence of random times (not necessarily stopping times) satisfying $T_n \to \infty$ and $T_{n+1}/T_n \to 1$ almost surely, then*

$$\limsup_{n\to\infty} \frac{B(T_n)}{\psi(T_n)} = 1 \text{ almost surely.}$$

Furthermore, if $T_n/n \to a$ almost surely, then

$$\limsup_{n\to\infty} \frac{B(T_n)}{\psi(an)} = 1 \text{ almost surely.}$$

Proof. The upper bound follows from the upper bound for continuous time without any conditions on $\{T_n : n \geq 1\}$. For the lower bound some restrictions are needed, which prevent us from choosing, for example, $T_0 = 0$ and $T_n = \inf\{t > T_{n-1} + 1 \colon B(t) < \frac{1}{n}\}$. Our conditions $T_{n+1}/T_n \to 1$ and $T_n \to \infty$ make sure that the times are sufficiently dense to rule out this effect. Define, for fixed $q > 4$,

$$D_k = \left\{B(q^k) - B(q^{k-1}) \geq \psi(q^k - q^{k-1})\right\},$$

$$\Omega_k = \left\{\min_{q^k \leq t \leq q^{k+1}} B(t) - B(q^k) \geq -\sqrt{q^k}\right\} \text{ and } D_k^* = D_k \cap \Omega_k.$$

Note that D_k and Ω_k are independent events. From Brownian scaling and Lemma II.3.1 it is easy to see that, for a suitable constant $c > 0$,

$$\mathbb{P}(D_k) = \mathbb{P}\Big\{B(1) \geq \tfrac{\psi(q^k - q^{k-1})}{\sqrt{q^k - q^{k-1}}}\Big\} \geq \frac{c}{k\log k}.$$

Moreover, by scaling, $\mathbb{P}(\Omega_k) =: c_q > 0$, and c_q that does not depend on k. As $\mathbb{P}(D_k^*) = c_q\mathbb{P}(D_k)$ the sum $\sum_k \mathbb{P}(D_{2k}^*)$ is infinite. As the events $\{D_{2k}^* : k \geq 1\}$ are independent, by the Borel-Cantelli lemma, for infinitely many (even) k,

$$\min_{q^k \leq t \leq q^{k+1}} B(t) \geq B(q^{k-1}) + \psi(q^k - q^{k-1}) - \sqrt{q^k}.$$

By Remark 5.2, for all sufficiently large k, we have $B(q^{k-1}) \geq -2\psi(q^{k-1})$ and, by easy asymptotics, $\psi(q^k - q^{k-1}) \geq \psi(q^k)(1 - \frac{1}{q})$. Hence, for infinitely many k,

$$\min_{q^k \leq t \leq q^{k+1}} B(t) \geq \psi(q^k - q^{k-1}) - 2\psi(q^{k-1}) - \sqrt{q^k} \geq \psi(q^k)\left(1 - \tfrac{1}{q} - \tfrac{2}{\sqrt{q}}\right) - \sqrt{q^k}\,,$$

with the right hand side being positive by our choice of q. Now define $n(k) = \min\{n : T_n > q^k\}$. Since the ratios T_{n+1}/T_n tend to 1, it follows that for any fixed $\varepsilon > 0$, we have $q^k \leq T_{n(k)} < q^k\,(1+\varepsilon)$ for all large k. Thus, for infinitely many k,

$$\frac{B(T_{n(k)})}{\psi(T_{n(k)})} \geq \frac{\psi(q^k)}{\psi(q^k(1+\varepsilon))}\left(1 - \tfrac{1}{q} - \tfrac{2}{\sqrt{q}}\right) - \frac{\sqrt{q^k}}{\psi(q^k)}\,.$$

But since $\sqrt{q^k}/\psi(q^k) \to 0$ and $\psi(q^k)/\psi(q^k(1+\varepsilon)) \to 1/\sqrt{1+\varepsilon}$, we conclude that

$$\limsup_{n\to\infty} \frac{B(T_n)}{\psi(T_n)} \geq \frac{1}{1+\varepsilon}\left(1 - \tfrac{1}{q} - \tfrac{2}{\sqrt{q}}\right),$$

and since the left hand side does not depend on q and $\varepsilon > 0$ we can let $q \uparrow \infty$ and $\varepsilon \downarrow 0$ to arrive at the desired conclusion. For the last part, note that if $T_n/n \to a$ then $\psi(T_n)/\psi(an) \to 1$. ∎

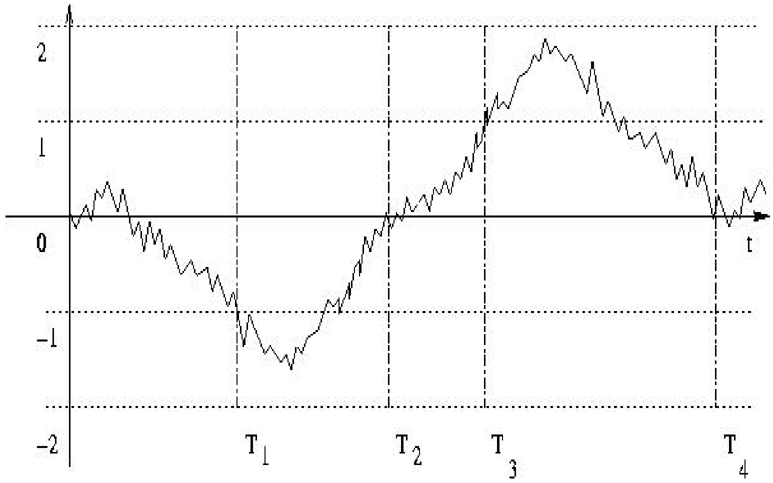

FIGURE 2. Embedding simple random walk into Brownian motion

Proof of Theorem 5.4. To prove the law of the iterated logarithm for simple random walk, we let $T_0 = 0$ and, for $n \geq 1$,

$$T_n = \min\{t > T_{n-1} : |B(t) - B(T_{n-1})| = 1\}.$$

The times T_n are stopping times for Brownian motion and, hence, by the strong Markov property, the waiting times $T_n - T_{n-1}$ are independent and identically distributed random variables. Obviously, $\mathbb{P}\{B(T_n) - B(T_{n-1}) = 1\} = \mathbb{P}\{B(T_n) - B(T_{n-1}) = -1\} = \frac{1}{2}$, and therefore $\{B(T_n) : n \geq 0\}$ is a simple random walk. By Theorem 2.45, we have $\mathbb{E}[T_n - T_{n-1}] = 1$, and hence the law of large numbers ensures that T_n/n converges almost surely to 1, and the theorem follows from Lemma 5.5. ∎

Remark 5.6. The technique we have used to get Theorem 5.4 from Theorem 5.1 was based on finding an increasing sequence of stopping times $\{T_n : n \geq 0\}$ for the Brownian motion, such that $S_n = B(T_n)$ defines a simple random walk, while we keep some control over the size of T_n. This 'embedding technique' will be extended substantially in Section 3. $\diamond$

2. Points of increase for random walk and Brownian motion

A point $t \in (0, \infty)$ is a local point of increase for the function $f\colon (0,\infty) \to \mathbb{R}$ if for some open interval (a, b) containing t we have $f(s) \leq f(t)$ for all $s \in (a, t)$ and $f(t) \leq f(s)$ for all $s \in (t, b)$. In this section we show that Brownian motion almost surely has no local points of increase. Our proof uses a combinatorial argument to derive a quantitative result for simple random walks, and then uses this result to study the case of Brownian motion. A crucial tool in the proof is an inequality of Harris [**Ha60**], which is of some independent interest.

Theorem 5.7 (Harris' inequality). *Suppose that $X = (X_1, \ldots, X_d)$ is a random variable with values in $\mathbb{R}^d$ and independent coordinates. Let $f, g\colon \mathbb{R}^d \to \mathbb{R}$ be measurable functions, which are nondecreasing in each coordinate. Then,*

$$\mathbb{E}\big[f(X)g(X)\big] \geq \mathbb{E}[f(X)]\,\mathbb{E}[g(X)]\,, \tag{2.1}$$

provided the above expectations are well-defined.

Proof. One can argue, using the monotone convergence theorem, that it suffices to prove the result when f and g are bounded. We assume f and g are bounded and proceed by induction on the dimension d. Suppose first that $d = 1$. Note that

$$(f(x) - f(y))(g(x) - g(y)) \geq 0\,, \qquad \text{for all } x, y \in \mathbb{R}.$$

Therefore, for Y an independent random variable with the same distribution as X,

$$\begin{aligned} 0 &\leq \mathbb{E}\big[(f(X) - f(Y))(g(X) - g(Y))\big] \\ &= 2\mathbb{E}\big[f(X)g(X)\big] - 2\mathbb{E}\big[f(X)\big]\,\mathbb{E}\big[g(Y)\big]\,, \end{aligned}$$

and (2.1) follows easily. Now, suppose (2.1) holds for $d - 1$. Define

$$f_1(x_1) = \mathbb{E}\big[f(x_1, X_2, \ldots, X_d)\big]\,,$$

and define g_1 similarly. Note that $f_1(x_1)$ and $g_1(x_1)$ are non-decreasing functions of x_1. Since f and g are bounded, we may apply Fubini's theorem to write the left hand side of (2.1) as

$$\int_{\mathbb{R}} \mathbb{E}\big[f(x_1, X_2, \ldots, X_d)\, g(x_1, X_2, \ldots, X_d)\big]\, d\mu_1(x_1)\,, \tag{2.2}$$

where μ_1 denotes the law of X_1. The expectation in the integral is at least $f_1(x_1)g_1(x_1)$ by the induction hypothesis. Thus, using the result for the $d = 1$ case, we can bound (2.2) from below by $\mathbb{E}[f_1(X_1)]\,\mathbb{E}[g_1(X_2)]$, which equals the right hand side of (2.1), completing the proof. ∎

For the rest of this section, let $X_1, X_2, \ldots$ be independent random variables with

$$\mathbb{P}\{X_i = 1\} = \mathbb{P}\{X_i = -1\} = \tfrac{1}{2},$$

and let $S_k = \sum_{i=1}^k X_i$ be their partial sums. Denote

$$p_n = \mathbb{P}\{S_i \geq 0 \text{ for all } 1 \leq i \leq n\}. \tag{2.3}$$

Then $\{S_n$ is a maximum among $S_0, S_1, \ldots S_n\}$ is precisely the event that the reversed random walk given by $S'_k = X_n + \ldots + X_{n-k+1}$ is nonnegative for all $k = 1, \ldots, n$. Hence this event also has probability p_n. The following lemma gives the order of magnitude of p_n, the proof will be given as Exercise 5.4.

LEMMA 5.8. *There are positive constants C_1 and C_2 such that*

$$\frac{C_1}{\sqrt{n}} \leq \mathbb{P}\{S_i \geq 0 \text{ for all } 1 \leq i \leq n\} \leq \frac{C_2}{\sqrt{n}} \text{ for all } n \geq 1.$$

The next lemma expresses, in terms of the p_n defined in (2.3), the probability that S_j stays between 0 and S_n for j between 0 and n.

LEMMA 5.9. *We have* $p_n^2 \leq \mathbb{P}\big\{0 \leq S_j \leq S_n \text{ for all } 1 \leq j \leq n\big\} \leq p_{\lfloor n/2 \rfloor}^2$.

Proof. The two events

$$\begin{aligned} A &= \{0 \leq S_j \text{ for all } j \leq \lfloor n/2 \rfloor\} \text{ and} \\ B &= \{S_j \leq S_n \text{ for } j \geq \lfloor n/2 \rfloor\} \end{aligned}$$

are independent, since A depends only on $X_1, \ldots, X_{\lfloor n/2 \rfloor}$ and B depends only on the remaining $X_{\lfloor n/2 \rfloor + 1}, \ldots, X_n$. Therefore,

$$\mathbb{P}\{0 \leq S_j \leq S_n\} \leq \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) \leq p_{\lfloor n/2 \rfloor}^2,$$

which proves the upper bound.

For the lower bound, we let $f(x_1, \ldots, x_n) = 1$ if all the partial sums $x_1 + \ldots + x_k$ for $k = 1, \ldots, n$ are nonnegative, and $f(x_1, \ldots, x_n) = 0$ otherwise. Also, define $g(x_1, \ldots, x_n) = f(x_n, \ldots, x_1)$. Then f and g are nondecreasing in each component. By Harris' inequality, for $X = (X_1, \ldots, X_n)$,

$$\mathbb{E}\big[f(X)g(X)\big] \geq \mathbb{E}[f(X)]\,\mathbb{E}[g(X)] = p_n^2.$$

Also,

$$\begin{aligned} \mathbb{E}\big[f(X)g(X)\big] &= \mathbb{P}\{X_1 + \ldots + X_j \geq 0 \text{ and } X_{j+1} + \ldots + X_n \geq 0 \text{ for all } j\} \\ &= \mathbb{P}\big\{0 \leq S_j \leq S_n \text{ for all } 1 \leq j \leq n\big\}, \end{aligned}$$

which proves the lower bound. ∎

DEFINITION 5.10.

(a) *A sequence $s_0, s_1, \ldots, s_n$ of reals has a (global)* **point of increase** *at $k \in \{0, \ldots, n\}$, if $s_i \leq s_k$ for $i = 0, 1, \ldots, k-1$ and $s_k \leq s_j$ for $j = k+1, \ldots, n$.*

(b) *A real-valued function f has a* **global point of increase in the interval** (a, b) *if there is a point $t \in (a, b)$ such that $f(s) \leq f(t)$ for all $s \in (a, t)$ and $f(t) \leq f(s)$ for all $s \in (t, b)$. t is a* **local point of increase** *if it is a global point of increase in some interval.* ◇

Theorem 5.11. *Let $S_0, S_1, \ldots, S_n$ be a simple random walk. Then*

$$\mathbb{P}\{S_0, \ldots, S_n \text{ has a point of increase}\} \le \frac{C}{\log n} \ ,$$

for all $n > 1$, where C does not depend on n.

The key to Theorem 5.11 is the following upper bound, which holds for more general random walks. It will be proved as Exercise 5.5.

Lemma 5.12. *For any random walk $\{S_j : j \ge 0\}$ on the line,*

$$\mathbb{P}\{S_0, \ldots, S_n \text{ has a point of increase}\} \le 2\, \frac{\sum_{k=0}^{n} p_k p_{n-k}}{\sum_{k=0}^{\lfloor n/2 \rfloor} p_k^2} \ . \tag{2.4}$$

Remark 5.13. Equation (2.4) is easy to interpret: The expected number of points of increase by time $\lfloor n/2 \rfloor$ is the numerator in (2.4), and given that there is at least one such point, the expected number is bounded below by the denominator; hence twice the ratio of these expectations bounds the required probability. ◇

Proof of Theorem 5.11. To bound the numerator in (2.4), we can use symmetry to deduce from Lemma 5.8 that

$$\begin{aligned}\sum_{k=0}^{n} p_k p_{n-k} \le 2 + 2 \sum_{k=1}^{\lfloor n/2 \rfloor} p_k p_{n-k} &\le 2 + 2C_2^2 \sum_{k=1}^{\lfloor n/2 \rfloor} k^{-1/2} (n-k)^{-1/2} \\ &\le \ 2 + 4C_2^2 n^{-1/2} \sum_{k=1}^{\lfloor n/2 \rfloor} k^{-1/2} \,,\end{aligned}$$

which is bounded above because the last sum is $O(n^{1/2})$. Since Lemma 5.8 implies that the denominator in (2.4) is at least $C_1^2 \log \lfloor n/2 \rfloor$, this completes the proof. ■

We now see how we can use embedding ideas to pass from the result about *simple* random walks to the result about Brownian motion.

Theorem 5.14. *Brownian motion almost surely has no local points of increase.*

Proof. To deduce this, it suffices to apply Theorem 5.11 to a *simple* random walk on the integers. Indeed, it clearly suffices to show that the Brownian motion $\{B(t) : t \ge 0\}$ almost surely has no global points of increase in a fixed time interval (a, b) with rational endpoints. Sampling the Brownian motion when it visits a lattice yields a simple random walk; by refining the lattice, we may make this walk as long as we wish, which will complete the proof.

More precisely, for any vertical spacing $h > 0$ define τ_0 to be the first $t \ge a$ such that $B(t)$ is an integral multiple of h, and for $i \ge 0$ let τ_{i+1} be the minimal $t \ge \tau_i$ such that $|B(t) - B(\tau_i)| = h$. Define $N_b = \max\{k \in \mathbb{Z} : \tau_k \le b\}$. For integers i satisfying $0 \le i \le N_b$, define

$$S_i = \frac{B(\tau_i) - B(\tau_0)}{h}.$$

Then $\{S_i : i = 1, \dots, N_b\}$ is a finite portion of a simple random walk. If the Brownian motion has a (global) point of increase t_0 in (a, b) at t, and if k is an integer such that $\tau_{k-1} \le t_0 \le \tau_k$, then this random walk has points of increase at $k-1$ and k. If $t_0 \in (a+\varepsilon, b-\varepsilon)$, for some $\varepsilon > 0$, such a k is guaranteed to exist if $|B(a+\varepsilon) - B(a)| > h$ and $|B(b-\varepsilon) - B(b)| > h$. Therefore, for all n,

$$\begin{aligned}
&\mathbb{P}\big\{\{B(t)\colon t \ge 0\} \text{ has a global point of increase in } (a,b) \text{ situated in } (a+\varepsilon, b-\varepsilon)\big\} \\
&\le \mathbb{P}\{N_b \le n\} + \mathbb{P}\big\{|B(a+\varepsilon) - B(a)| \le h\big\} + \mathbb{P}\big\{|B(b-\varepsilon) - B(b)| \le h\big\} \\
&\quad + \sum_{m=n+1}^{\infty} \mathbb{P}\big\{S_0, \dots, S_m \text{ has a point of increase and } N_b = m\big\}.
\end{aligned} \tag{2.5}$$

Note that $N_b \le n$ implies $|B(b) - B(a)| \le (n+1)h$, so

$$\mathbb{P}\{N_b \le n\} \le \mathbb{P}\big\{|B(b) - B(a)| \le (n+1)h\big\} = \mathbb{P}\Big\{|Z| \le \tfrac{(n+1)h}{\sqrt{b-a}}\Big\},$$

where Z has a standard normal distribution. Since $S_0, \dots, S_m$, conditioned on $N_b = m$ is a finite portion of a simple random walk, it follows from Theorem 5.11 that for some constant C, we have

$$\begin{aligned}
&\sum_{m=n+1}^{\infty} \mathbb{P}\big\{S_0, \dots, S_m \text{ has a point of increase, and } N_b = m\big\} \\
&\qquad \le \sum_{m=n+1}^{\infty} \mathbb{P}\{N_b = m\} \frac{C}{\log m} \le \frac{C}{\log(n+1)}.
\end{aligned}$$

Thus, the probability in (2.5) can be made arbitrarily small by first taking n large and then picking $h > 0$ sufficiently small. Finally, let $\varepsilon \downarrow 0$ to complete the proof. ∎

3. The Skorokhod embedding problem

In the proof of Theorem 5.4 we have made use of the fact that there exists a stopping time T for linear Brownian motion with the property that $\mathbb{E}[T] < \infty$ and the law of $B(T)$ is the uniform distribution on $\{-1, 1\}$. To use the same method for random walks $\{S_n\colon n \in \mathbb{N}\}$ with general increments, it would be necessary to find, for a given random variable X representing an increment, a stopping time T with $\mathbb{E}[T] < \infty$, such that $B(T)$ has the law of X.

This problem is called the *Skorokhod embedding problem.* By Wald's lemmas, Theorem 2.40 and Theorem 2.44, for any integrable stopping time T, we have

$$\mathbb{E}\big[B(T)\big] = 0 \qquad \text{and} \qquad \mathbb{E}\big[B(T)^2\big] = \mathbb{E}[T] < \infty\,,$$

so that the Skorokhod embedding problem can only be solved for random variables X with mean zero and finite second moment. However, these are the only restrictions, as the following result shows.

Theorem 5.15 (Skorokhod embedding theorem). *Suppose that* $\{B(t) : t \geq 0\}$ *is a standard Brownian motion and that* X *is a real valued random variable with* $E[X] = 0$ *and* $E[X^2] < \infty$. *Then there exists a stopping time* T*, with respect to the natural filtration* $(\mathcal{F}(t) : t \geq 0)$ *of the Brownian motion, such that* $B(T)$ *has the law of* X *and* $\mathbb{E}[T] = E[X^2]$.

Example 5.16. Assume that X may take two values $a < b$. In order that $E[X] = 0$ we must have $a < 0 < b$ and $P\{X = a\} = b/(b-a)$ and $P\{X = b\} = -a/(b-a)$. We have seen in Theorem 2.45 that, for the stopping time $T = \inf\{t : B(t) \notin (a,b)\}$ the random variable $B(T)$ has the same distribution as X, and that $\mathbb{E}[T] = -ab$ is finite. $\diamond$

Note that the Skorokhod embedding theorem allows us to use the arguments developed for the proof of the law of the iterated logarithm for simple random walks, Theorem 5.4, and obtain a much more general result.

Theorem 5.17 (Hartman-Wintner law of the iterated logarithm). *Let* $\{S_n : n \in \mathbb{N}\}$ *be a random walk with increments* $S_n - S_{n-1}$ *of zero mean and finite variance* σ^2. *Then*

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2\sigma^2\, n \log\log n}} = 1.$$

We now present two proofs of the Skorokhod embedding theorem, which actually represent different constructions of the required stopping times. Both approaches, Dubins' embedding, and the Azéma-Yor embedding are very elegant and have their own merits.

3.1. The Dubins' embedding theorem. The first one, due to Dubins [**Du68**], is particularly simple and based on the notion of binary splitting martingales. We say that a martingale $\{X_n : n \in \mathbb{N}\}$ is **binary splitting** if, whenever for some $x_0, \ldots, x_n \in \mathbb{R}$ the event

$$A(x_0, \ldots, x_n) := \{X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n\}$$

has positive probability, the random variable X_{n+1} conditioned on $A(x_0, \ldots, x_n)$ is supported on at most two values.

Lemma 5.18. *Let* X *be a random variable with* $E[X^2] < \infty$. *Then there is binary splitting martingale* $\{X_n : n \in \mathbb{N}\}$ *such that* $X_n \to X$ *almost surely and in* L^2.

Proof. We define the martingale $\{X_n : n \in \mathbb{N}\}$ and the associated filtration $(\mathcal{G}_n : n \in \mathbb{N})$ recursively. Let $\mathcal{G}_0$ be the trivial σ-algebra and $X_0 = EX$. Define the random variable ξ_0 by

$$\xi_0 = \begin{cases} 1\,, & \text{if } X \geq X_0\,, \\ -1\,, & \text{if } X < X_0\,. \end{cases}$$

For any $n > 0$, let $\mathcal{G}_n = \sigma\{\xi_0, \ldots, \xi_{n-1}\}$ and $X_n = E[X \mid \mathcal{G}_n]$. Also define the random variable ξ_n by

$$\xi_n = \begin{cases} 1\,, & \text{if } X \geq X_n\,, \\ -1\,, & \text{if } X < X_n\,. \end{cases}$$

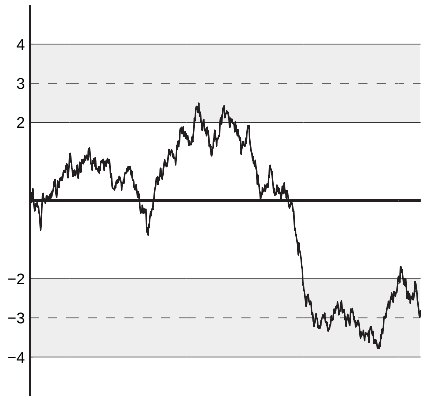

FIGURE 3. Dubins' embedding for the uniform distribution on $\{-4,-2,0,2,4\}$: First go until you hit $\{-3,3\}$, in this picture you hit -3. Given that, continue until you hit either -2 or -4, in this picture you hit -2. Hence $B(T)=-2$ for this sample.

Note that $\mathcal{G}_n$ is generated by a partition $\mathcal{P}_n$ into 2^n sets, each of which has the form $A(x_0,\dots,x_n)$. As each element of $\mathcal{P}_n$ is a union of two elements of $\mathcal{P}_{n+1}$, the martingale $\{X_n : n\in\mathbb{N}\}$ is binary splitting. Also we have, for example as in Appendix II.(3.1), that

$$E\big[X^2\big]=E\big[(X-X_n)^2\big]+E\big[X_n^2\big]\geq E\big[X_n^2\big]\,.$$

Hence $\{X_n : n\in\mathbb{N}\}$ is bounded in L^2 and, from the convergence theorem for L^2-bounded martingales and Lévy's upward theorem, see Theorems II.4.12 and II.4.9, we get

$$X_n\to X_\infty:=E\big[X\,\big|\,\mathcal{G}_\infty\big]\qquad\text{almost surely and in } L^2,$$

where $\mathcal{G}_\infty=\sigma\big(\bigcup_{i=0}^\infty\mathcal{G}_i\big)$. To conclude the proof we have to show that $X=X_\infty$ almost surely.

We claim that, almost surely,

$$\lim_{n\uparrow\infty}\xi_n\,(X-X_{n+1})=|X-X_\infty|\,. \tag{3.1}$$

Indeed, if $X(\omega)=X_\infty(\omega)$ this is trivial. If $X(\omega)<X_\infty(\omega)$ then for some large enough N we have $X(\omega)<X_n(\omega)$ for any $n>N$, hence $\xi_n=-1$ and (3.1) holds. Similarly, if $X(\omega)>X_\infty(\omega)$ then $\xi_n=1$ for $n>N$ and so (3.1) holds.

Using that ξ_n is $\mathcal{G}_{n+1}$-measurable, we find that

$$E\big[\xi_n\,(X-X_{n+1})\big]=E\big[\xi_n\,E[X-X_{n+1}\mid\mathcal{G}_{n+1}]\big]=0\,.$$

Recall that if $Y_n\to Y$ almost surely, and $\{Y_n\colon n=0,1,\cdots\}$ is L^2-bounded, then $EY_n\to EY$ (see, for example, the discussion of uniform integrability in Appendix II.3). Hence, as the left hand side of (3.1) is L^2-bounded, we conclude that $E|X-X_\infty|=0$. ∎

Proof of Theorem 5.15. From Lemma 5.18 we take a binary splitting martingale $\{X_n : n \in \mathbb{N}\}$ such that $X_n \to X$ almost surely and in L^2. Recall from the example preceding this proof that if X is supported on a set two elements $\{-a, b\}$ for some $a, b > 0$ then $T = \inf\{t : B(t) \in \{-a, b\}\}$ is the required stopping time. Hence, as X_n conditioned on $A(x_0, \ldots, x_{n-1})$ is supported on at most two values it is clear we can find a sequence of stopping times $T_0 \le T_1 \le \ldots$ such that $B(T_n)$ is distributed as X_n and $\mathbb{E}T_n = E[X_n^2]$. As T_n is an increasing sequence, we have $T_n \uparrow T$ almost surely for some stopping time T. Also, by the monotone convergence theorem

$$\mathbb{E}T = \lim_{n\uparrow\infty} \mathbb{E}T_n = \lim_{n\uparrow\infty} E\big[X_n^2\big] = E\big[X^2\big]\,.$$

As $B(T_n)$ converges in distribution to X by our construction, and converges almost surely to $B(T)$ by continuity of the Brownian sample paths, we conclude that $B(T)$ is distributed as X. ■

3.2. The Azéma-Yor embedding theorem. In this section we discuss a second solution to the Skorokhod embedding problem with a more explicit construction of the stopping times.

Theorem* 5.19 (Azéma-Yor embedding theorem). *Suppose that X is a real valued random variable with $E[X] = 0$ and $E[X^2] < \infty$. Let*

$$\Psi(x) = E\big[X \,\big|\, X \ge x\big] \qquad \text{if } P\{X \ge x\} > 0\,,$$

and $\Psi(x) = 0$ otherwise. For a Brownian motion $\{B(t) : t \ge 0\}$ let $\{M(t) : t \ge 0\}$ be the maximum process and define a stopping time τ by

$$\tau = \inf\{t \ge 0 \,:\, M(t) \ge \Psi(B(t))\}.$$

Then $\mathbb{E}[\tau] = E[X^2]$ and $B(\tau)$ has the same law as X.

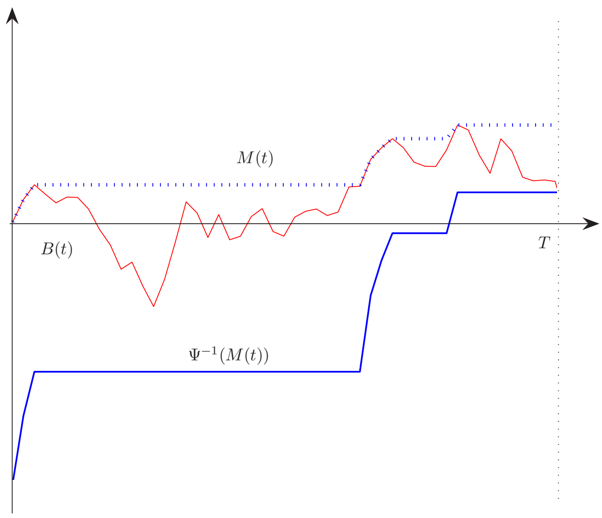

Figure 4. The Azéma-Yor embedding: the path is stopped when the Brownian motion hits the level $\Psi^{-1}(M(t))$, where $\Psi^{-1}(x) = \sup\{b\colon \Psi(b) \le x\}$.

We proceed in three steps. In the first step we formulate an embedding for random variables taking only finitely many values.

Lemma 5.20. *Suppose the random variable X takes only finitely many values*

$$x_1 < x_2 < \cdots < x_n.$$

Define $y_1 < y_2 < \cdots < y_{n-1}$ by $y_i = \Psi(x_{i+1})$, and define stopping times $T_0 = 0$ and

$$T_i = \inf\big\{t \geq T_{i-1} \,:\, B(t) \notin (x_i, y_i)\big\} \qquad \text{for } i \leq n-1.$$

Then $T = T_{n-1}$ satisfies $\mathbb{E}[T] = E[X^2]$ and $B(T)$ has the same law as X.

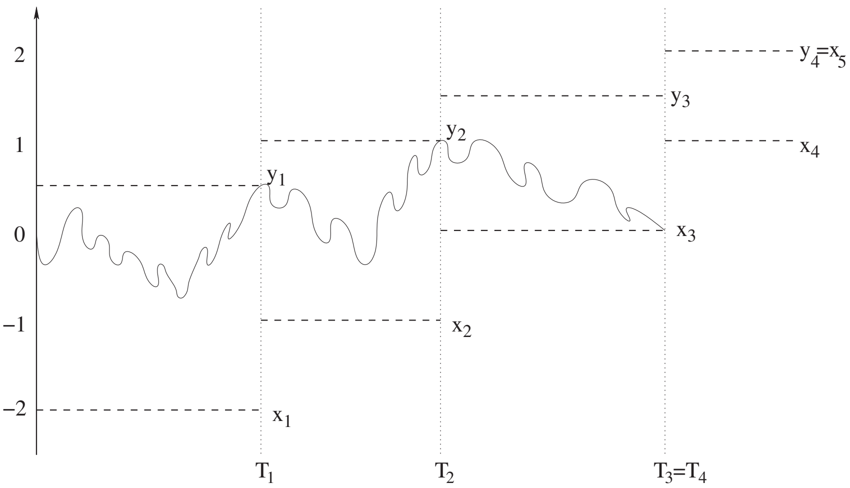

Figure 5. The Azéma-Yor embedding for the uniform distribution on the set $\{-2, -1, 0, 1, 2\}$. The drawn path samples the value $B(T) = 0$ with $T = T_4$.

Proof. First observe that $y_i \geq x_{i+1}$ and equality holds if and only if $i = n-1$. We have $\mathbb{E}[T_{n-1}] < \infty$, by Theorem 2.45, and $\mathbb{E}[T_{n-1}] = \mathbb{E}[B(T_{n-1})^2]$, from Theorem 2.44. For $i = 1, \ldots, n-1$ define random variables

$$Y_i = \begin{cases} E[X \,|\, X \geq x_{i+1}] & \text{if } X \geq x_{i+1}, \\ X & \text{if } X \leq x_i. \end{cases}$$

Note that Y_1 has expectation zero and takes on the two values x_1, y_1. For $i \geq 2$, given $Y_{i-1} = y_{i-1}$, the random variable Y_i takes the values x_i, y_i and has expectation y_{i-1}. Given $Y_{i-1} = x_j$, $j \leq i-1$ we have $Y_i = x_j$. Note that $Y_{n-1} = X$. We now argue that

$$(B(T_1), \ldots, B(T_{n-1})) \stackrel{d}{=} (Y_1, \ldots, Y_{n-1}).$$

Clearly, $B(T_1)$ can take only the values x_1, y_1 and has expectation zero, hence the law of $B(T_1)$ agrees with the law of Y_1. For $i \geq 2$, given $B(T_{i-1}) = y_{i-1}$, the random variable $B(T_i)$ takes the values x_i, y_i and has expectation y_{i-1}. Given $B(T_{i-1}) = x_j$ where $j \leq i-1$, we have $B(T_i) = x_j$. Hence the two tuples have the same law and, in particular, $B(T_{n-1})$ has the same law as X. ■

In the second step, we show that the stopping time we have constructed in Lemma 5.20 agrees with the stopping time τ in the Azéma-Yor embedding.

LEMMA 5.21. *The stopping time T constructed in Lemma 5.20 and the stopping time τ in Theorem 5.19 are equal.*

Proof. Suppose that $B(T_{n-1}) = x_i$, and hence $\Psi(B(T_{n-1})) = y_{i-1}$. If $i \leq n-1$, then i is minimal with the property that $B(T_i) = \cdots = B(T_{n-1})$, and thus $B(T_{i-1}) \neq B(T_i)$. Hence $M(T_{n-1}) \geq y_{i-1}$. If $i = n$ we also have $M(T_{n-1}) = x_n \geq y_{i-1}$, which implies in any case that $\tau \leq T$. Conversely, if $T_{i-1} \leq t < T_i$ then $B(t) \in (x_i, y_i)$ and this implies $M(t) < y_i \leq \Psi(B(t))$. Hence $\tau \geq T$, and altogether we have seen that $T = \tau$. ■

This completes the proof of Theorem 5.19 for random variables taking finitely many values. The general case follows from a limiting procedure, which is left as Exercise 5.9.

4. The Donsker invariance principle

Let $\{X_n : n \geq 0\}$ be a sequence of independent and identically distributed random variables and assume that they are normalized, so that $\mathbb{E}[X_n] = 0$ and $\mathrm{Var}(X_n) = 1$. This assumption is no loss of generality for X_n with finite variance, since we can always consider the normalization

$$\frac{X_n - \mathbb{E}[X_n]}{\sqrt{\mathrm{Var}(X_n)}}.$$

We look at the *random walk* generated by the sequence

$$S_n = \sum_{k=1}^{n} X_k \,,$$

and interpolate linearly between the integer points, i.e.

$$S(t) = S_{[t]} + (t - [t])(S_{[t]+1} - S_{[t]}) \,.$$

This defines a random function $S \in \mathcal{C}[0,\infty)$. We now define a sequence $\{S_n^* : n \geq 1\}$ of random functions in $\mathcal{C}[0,1]$ by

$$S_n^*(t) = \frac{S(nt)}{\sqrt{n}} \text{ for all } t \in [0,1].$$

THEOREM 5.22 (Donsker's Invariance Principle). *On the space $\mathcal{C}[0,1]$ of continuous functions on the unit interval with the metric induced by the sup-norm, the sequence $\{S_n^* : n \geq 1\}$ converges in distribution to a standard Brownian motion $\{B(t) : t \in [0,1]\}$.*

REMARK 5.23. Donsker's invariance principle is also called the *functional central limit theorem.* The name *invariance principle* comes from the fact that the limit in Theorem 5.22 does not depend on the choice of the exact distribution of the normalised random variables X_n. ◇

The idea of the proof is to construct the random variables $X_1, X_2, X_3, \ldots$ on the same probability space as the Brownian motion in such a way that $\{S_n^* : n \geq 1\}$ is with high probability close to a scaling of this Brownian motion.

LEMMA 5.24. *Suppose that $\{B(t)\colon t \geq 0\}$ a linear Brownian motion. Then, for any random variable X with mean zero and variance one, there exists a sequence of stopping times*

$$0 = T_0 \leq T_1 \leq T_2 \leq T_3 \leq \dots$$

with respect to the Brownian motion, such that

(a) *the sequence $\{B(T_n)\colon n \geq 0\}$ has the distribution of the random walk with increments given by the law of X,*

(b) *the sequence of functions $\{S_n^*\colon n \geq 0\}$ constructed from this random walk satisfies*

$$\lim_{n\to\infty} \mathbb{P}\Big\{ \sup_{0\leq t\leq 1} \Big| \frac{B(nt)}{\sqrt{n}} - S_n^*(t) \Big| > \varepsilon \Big\} = 0\,.$$

Proof. Using Skorokhod embedding, we define T_1 to be a stopping time with $\mathbb{E}[T_1] = 1$ such that $B(T_1) = X$ in distribution. By the strong Markov property,

$$\{B_2(t)\colon t \geq 0\} = \{B(T_1 + t) - B(T_1)\colon t \geq 0\}$$

is a Brownian motion and independent of $\mathcal{F}^+(T_1)$ and, in particular, of $(T_1, B(T_1))$. Hence we can define a stopping time T_2' for the Brownian motion $\{B_2(t)\colon t \geq 0\}$ such that $\mathbb{E}[T_2'] = 1$ such that $B(T_2') = X$ in distribution. Then $T_2 = T_1 + T_2'$ is a stopping time for the original Brownian motion with $\mathbb{E}[T_2] = 2$, such that $B(T_2)$ is the second value in a random walk with increments given by the law of X. We can proceed inductively to get a sequence $0 = T_0 \leq T_1 \leq T_2 \leq T_3 < \dots$ such that $S_n = B(T_n)$ is the embedded random walk, and $\mathbb{E}[T_n] = n$.

Abbreviate $W_n(t) = \frac{B(nt)}{\sqrt{n}}$ and let A_n be the event that there exists $t \in [0, 1)$ such that $|S_n^*(t) - W_n(t)| > \varepsilon$. We have to show that $\mathbb{P}(A_n) \to 0$. Let $k = k(t)$ be the unique integer with $(k-1)/n \leq t < k/n$. Because S_n^* is linear on such an interval we have

$$\begin{aligned} A_n \subset \big\{ &\text{ there exists } t \in [0,1) \text{ such that } \big|S_k/\sqrt{n} - W_n(t)\big| > \varepsilon \big\} \\ &\cup \big\{ \text{ there exists } t \in [0,1) \text{ such that } \big|S_{k-1}/\sqrt{n} - W_n(t)\big| > \varepsilon \big\}. \end{aligned}$$

As $S_k = B(T_k) = \sqrt{n} W_n(T_k/n)$, we obtain

$$\begin{aligned} A_n \subset\; A_n^* :=&\big\{ \text{ there exists } t \in [0,1) \text{ such that } \big|W_n\big(T_k/n\big) - W_n(t)\big| > \varepsilon \big\} \\ &\cup \big\{ \text{ there exists } t \in [0,1) \text{ such that } \big|W_n\big(T_{k-1}/n\big) - W_n(t)\big| > \varepsilon \big\}. \end{aligned}$$

For given $0 < \delta < 1$ the event A_n^* implies that either

$$\big\{ \text{ there exist } s, t \in [0,2] \text{ such that } |s-t| < \delta\,,\ |W_n(s) - W_n(t)| > \varepsilon \big\} \tag{4.1}$$

or

$$\big\{ \text{ there exists } t \in [0,1) \text{ such that } |T_k/n - t| \vee |T_{k-1}/n - t| \geq \delta \big\}\,. \tag{4.2}$$

Note that the probability of (4.1) does not depend on n. Choosing $\delta > 0$ small, we can make this probability as small as we wish, since Brownian motion is uniformly continuous on $[0, 2]$. It remains to show that for arbitrary, fixed $\delta > 0$, the probability of (4.2) converges to zero as $n \to \infty$. To prove this we use that

$$\lim_{n\to\infty} \frac{T_n}{n} = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} (T_k - T_{k-1}) = 1 \text{ almost surely.}$$

This is Kolmogorov's law of large numbers for the sequence $\{T_k - T_{k-1}\}$ of independent identically distributed random variables with mean 1. Observe that for every sequence $\{a_n\}$ of reals one has

$$\lim_{n\to\infty} \frac{a_n}{n} = 1 \Rightarrow \lim_{n\to\infty} \sup_{0\le k\le n} |a_k - k|/n = 0\,.$$

This is a matter of plain (deterministic) arithmetic and eaqsily checked. Hence we have,

$$\lim_{n\to\infty} \mathbb{P}\Big\{ \sup_{0\le k\le n} \frac{|T_k - k|}{n} \ge \delta \Big\} = 0\,. \tag{4.3}$$

Now recall that $t \in [(k-1)/n, k/n)$ and let $n > 2/\delta$. Then

$$\begin{aligned}
&\mathbb{P}\big\{ \text{ there exist } t \in [0,1] \text{ such that } |T_k/n - t| \vee |T_{k-1}/n - t| \ge \delta \big\} \\
\le\ & \mathbb{P}\Big\{ \sup_{1\le k\le n} \frac{(T_k - (k-1)) \vee (k - T_{k-1})}{n} \ge \delta \Big\} \\
\le\ & \mathbb{P}\Big\{ \sup_{1\le k\le n} \frac{T_k - k}{n} \ge \delta/2 \Big\} + \mathbb{P}\Big\{ \sup_{1\le k\le n} \frac{(k-1) - T_{k-1}}{n} \ge \delta/2 \Big\}\,,
\end{aligned}$$

and by (4.3) both summands converge to 0. ∎

Proof of the Donsker invariance principle. Choose the sequence of stopping times as in Lemma 5.24 and recall from the scaling property of Brownian motion that the random functions $\{W_n(t) \colon 0 \le t \le 1\}$ given by $W_n(t) = B(nt)/\sqrt{n}$ are standard Brownian motions. Suppose that $K \subset \mathcal{C}[0,1]$ is closed and define

$$K[\varepsilon] = \{f \in \mathcal{C}[0,1] \colon \|f - g\|_{\sup} \le \varepsilon \text{ for some } g \in K\}.$$

Then $\mathbb{P}\{S_n^* \in K\} \le \mathbb{P}\{W_n \in K[\varepsilon]\} + \mathbb{P}\{\|S_n^* - W_n\|_{\sup} > \varepsilon\}$. As $n \to \infty$, the second term goes to 0, whereas the first term does not depnd on n and is equal to $\mathbb{P}\{B \in K[\varepsilon]\}$ for a Brownian motion B. As K is closed we have

$$\lim_{\varepsilon\downarrow 0} \mathbb{P}\{B \in K[\varepsilon]\} = \mathbb{P}\Big\{B \in \bigcap_{\varepsilon>0} K[\varepsilon]\Big\} = \mathbb{P}\{B \in K\}.$$

Putting these facts together, we obtain $\limsup_{n\to\infty} \mathbb{P}\{S_n^* \in K\} \le \mathbb{P}\{B \in K\}$, which is condition (ii) in the Portmanteau theorem, Theorem II.1.6. Hence Donsker's invariance principle is proved. ∎

Below and in the following section we harvest a range of results for random walks, which we can transfer from Brownian motion by means of Donsker's invariance principle. Readers unfamiliar with the nature of convergence in distribution are recommended to look at the appendix, Chapter II.1.

Theorem 5.25. *Suppose that $\{X_k : k \ge 1\}$ is a sequence of independent, identically distributed random variables with $\mathbb{E}[X_1] = 0$ and $0 < \mathbb{E}[X_1^2] = \sigma^2 < \infty$. Let $\{S_n : n \ge 0\}$ be the associated random walk and*

$$M_n = \max\{S_k : 0 \le k \le n\}$$

its maximal value up to time n. Then, for all $x \geq 0$,

$$\lim_{n\to\infty} \mathbb{P}\{M_n \geq x\sqrt{n}\} = \frac{2}{\sqrt{2\pi\sigma^2}} \int_x^\infty e^{-y^2/2\sigma^2}\, dy\,.$$

Proof. By scaling we can assume that $\sigma^2 = 1$. Suppose now that $g\colon \mathbb{R} \to \mathbb{R}$ is a continuous bounded function. Define a function $G\colon \mathcal{C}[0,1] \to \mathbb{R}$ by

$$G(f) = g\Big(\max_{x\in[0,1]} f(x)\Big)\,,$$

and note that G is continuous and bounded. Then, by definition,

$$\mathbb{E}\Big[G(S_n^*)\Big] = \mathbb{E}\Big[g\Big(\max_{0\leq t\leq 1} \frac{S(tn)}{\sqrt{n}}\Big)\Big] = \mathbb{E}\Big[g\Big(\frac{\max_{0\leq k\leq n} S_k}{\sqrt{n}}\Big)\Big]\,,$$

and

$$\mathbb{E}\Big[G(B)\Big] = \mathbb{E}\Big[g\Big(\max_{0\leq t\leq 1} B(t)\Big)\Big].$$

Hence, by Donsker's invariance principle,

$$\lim_{n\to\infty} \mathbb{E}\Big[g\Big(\frac{M_n}{\sqrt{n}}\Big)\Big] = \mathbb{E}\Big[g\Big(\max_{0\leq t\leq 1} B(t)\Big)\Big].$$

From the Portmanteau theorem, Theorem II.1.6, and the reflection principle, Theorem 2.18, we infer

$$\lim_{n\to\infty} \mathbb{P}\{M_n \geq x\sqrt{n}\} = \mathbb{P}\{\max_{0\leq t\leq 1} B(t) \geq x\} = 2\mathbb{P}\{B(1) \geq x\}\,,$$

and the latter probability is the given integral. ■

5. The arcsine laws

We now discuss the two famous arcsine laws for Brownian motion and also for random walks. Their name comes from the **arcsine distribution**, which is the distribution on $(0,1)$ which has the density

$$\frac{1}{\pi\sqrt{x(1-x)}} \qquad \text{for } x \in (0,1).$$

The cumulative distribution function of an arcsine distributed random variable X is therefore given by

$$\mathbb{P}\{X \leq x\} = \frac{2}{\pi} \arcsin(\sqrt{x}) \qquad \text{for } x \in (0,1)\,.$$

The *first arcsine law* describes the law of the last passage over level zero by a Brownian motion or random walk running for finite time. In the case of a Brownian motion we shall find this law by a smart calculation, and then Donsker's invariance principle will allow us to transfer the result to random walks. Observe that the following result is surprising: the rightmost zero of Brownian motion in the interval $(0,1)$ is most likely to be near zero or one, see Figure 6.

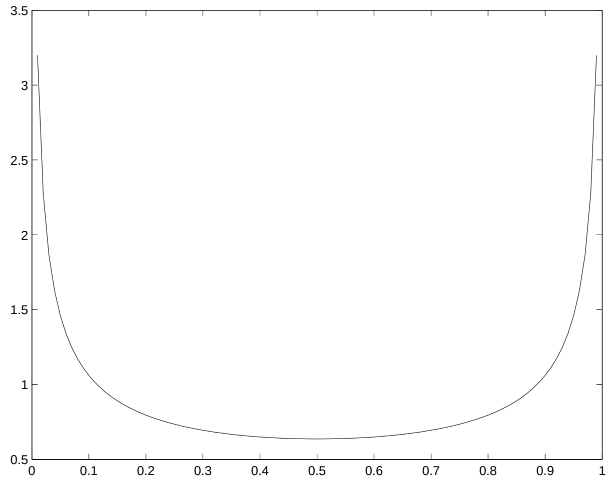

FIGURE 6. The density of the arcsine distribution is concentrated near the boundary values 0 and 1.

THEOREM 5.26 (First arcsine law for Brownian motion). *Let $\{B(t) : t \geq 0\}$ be a standard linear Brownian motion. Then,*

(a) *the random variable $L = \sup\big\{t \in [0,1] : B(t) = 0\big\}$, the last zero of Brownian motion in $[0,1]$, is arcsine distributed, and*
(b) *the random variable $M \in [0,1]$, which is uniquely determined by $B(M) = \max_{s\in[0,1]} B(s)$, is arcsine distributed.*

Proof. The discussion of local extrema in Chapter 2.1 has shown that there is indeed a unique maximum, and hence M is well-defined. Moreover Theorem 2.31 shows that M, which is the last zero of the process $\{M(t) - B(t) : t \geq 0\}$ has the same law as L. Hence it suffices to prove part (b).

Recall that $\{M(t) : 0 \leq t \leq 1\}$ is defined by $M(t) = \max_{0\leq s\leq t} B(s)$. For $s \in [0,1]$,

$$\begin{aligned}
\mathbb{P}\{M \leq s\} &= \mathbb{P}\big\{\max_{0\leq u\leq s} B(u) > \max_{s\leq v\leq 1} B(v)\big\} \\
&= \mathbb{P}\big\{\max_{0\leq u\leq s} B(u) - B(s) > \max_{s\leq v\leq 1} B(v) - B(s)\big\} \\
&= \mathbb{P}\big\{M_1(s) > M_2(1-s)\big\},
\end{aligned}$$

where $\{M_1(t) : 0 \leq t \leq 1\}$ is the maximum process of the Brownian motion $\{B_1(t) : t \geq 0\}$, which is given by $B_1(t) = B(s-t) - B(s)$, and $\{M_2(t) : 0 \leq t \leq 1\}$ is the maximum process of the independent Brownian motion $\{B_2(t) : 0 \leq s \leq 1\}$, which is given by $B_2(t) = B(s+t) - B(s)$. Since, by Theorem 2.18, for any fixed t, the random variable $M(t)$ has the same law as $|B(t)|$, we have

$$\mathbb{P}\big\{M_1(s) > M_2(1-s)\big\} = \mathbb{P}\big\{|B_1(s)| > |B_2(1-s)|\big\}.$$

Using the scaling invariance of Brownian motion we can express this in terms of a pair of two independent standard normal random variables Z_1 and Z_2 , by

$$\mathbb{P}\big\{|B_1(s)| > |B_2(1-s)|\big\} = \mathbb{P}\big\{\sqrt{s}\,|Z_1| > \sqrt{1-s}\,|Z_2|\big\} = \mathbb{P}\Big\{\frac{|Z_2|}{\sqrt{Z_1^2+Z_2^2}} < \sqrt{s}\Big\}.$$

In polar coordinates, $(Z_1, Z_2) = (R\cos\theta, R\sin\theta)$ pointwise. The fact that the random variable θ is uniformly distributed on $[0, 2\pi]$ follows from Lemma II.3.3 in the appendix. So the last quantity becomes

$$\begin{aligned}\mathbb{P}\Big\{\frac{|Z_2|}{\sqrt{Z_1^2+Z_2^2}} < \sqrt{s}\Big\} &= \mathbb{P}\big\{|\sin(\theta)| < \sqrt{s}\big\} = 4\mathbb{P}\big\{\theta < \arcsin(\sqrt{s})\big\} \\ &= 4\left(\frac{\arcsin(\sqrt{s})}{2\pi}\right) = \frac{2}{\pi}\arcsin(\sqrt{s}).\end{aligned}$$

It follows by differentiating that M has density $(\pi\sqrt{s(1-s)})^{-1}$ for $s \in (0,1)$. ■

For random walks the first arc-sine law takes the form of a limit theorem, as the length of the walk tends to infinity.

Proposition 5.27 (Arcsine law for the last sign-change). *Suppose that $\{X_k : k \geq 1\}$ is a sequence of independent, identically distributed random variables with $\mathbb{E}[X_1] = 0$ and $0 < \mathbb{E}[X_1^2] = \sigma^2 < \infty$. Let $\{S_n : n \geq 0\}$ be the associated random walk and*

$$N_n = \max\{1 \leq k \leq n : S_k S_{k-1} \leq 0\}$$

the last time the random walk changes its sign before time n. Then, for all $x \in (0,1)$,

$$\lim_{n\to\infty} \mathbb{P}\{N_n \leq xn\} = \frac{2}{\pi}\arcsin(\sqrt{x})\,.$$

Proof. The strategy of proof is to use Theorem 5.26, and apply Donsker's invariance principle to extend the result to random walks. As N_n is unchanged under scaling of the random walk we may assume that $\sigma^2 = 1$. Define a bounded function g on $\mathcal{C}[0,1]$ by

$$g(f) = \max\{t \leq 1 : f(t) = 0\}.$$

It is clear that $g(S_n^*)$ differs from N_n/n by a term, which is bounded by $1/n$ and therefore vanishes asymptotically. Hence Donsker's invariance principle would imply convergence of N_n/n in distribution to $g(B) = \sup\{t \leq 1 : B(t) = 0\}$ — if g was continuous. g is *not* continuous, but we show that g is continuous on the set $\mathcal{C}$ of all $f \in \mathcal{C}[0,1]$ such that f takes positive and negative values in every neighbourhood of every zero and $f(1) \neq 0$. As, by Theorem 2.25, Brownian motion is almost surely in $\mathcal{C}$, we get from property (v) in the Portmanteau theorem, Theorem II.1.6, and by Donsker's invariance principle, that, for every continuous bounded $h\colon \mathbb{R} \to \mathbb{R}$,

$$\lim_{n\to\infty} \mathbb{E}\Big[h\Big(\frac{N_n}{n}\Big)\Big] = \lim_{n\to\infty} \mathbb{E}\big[h \circ g(S_n^*)\big] = \mathbb{E}\big[h \circ g(B)\big] = \mathbb{E}\big[h(\sup\{t \leq 1 : B(t) = 0\})\big],$$

which completes the proof subject to the claim. To see that g is continuous on $\mathcal{C}$, let $\varepsilon > 0$ be given and $f \in \mathcal{C}$. Let

$$\delta_0 = \min_{t \in [g(f)+\varepsilon, 1]} |f(t)|\,,$$

and choose δ_1 such that

$$(-\delta_1, \delta_1) \subset f(g(f) - \varepsilon, g(f) + \varepsilon)\,.$$

Let $0 < \delta < \delta_0 \wedge \delta_1$. If now $\|h - f\|_\infty < \delta$, then h has no zero in $(g(f) + \varepsilon, 1]$, but has a zero in $(g(f) - \varepsilon, g(f) + \varepsilon)$, because there are $s, t \in (g(f) - \varepsilon, g(f) + \varepsilon)$ with $h(t) < 0$ and $h(s) > 0$. Thus $|g(h) - g(f)| < \varepsilon$. This shows that g is continuous on $\mathcal{C}$. ■

There is a second arcsine law for Brownian motion, which describes the law of the random variable $\mathcal{L}\big\{t \in [0,1] \,:\, B(t) > 0\big\}$, the time spent by Brownian motion above the x-axis. This statement is much harder to derive directly for Brownian motion, though we will do this using more sophisticated tools in Chapter 8. At this stage we can use random walks to derive the result for Brownian motion.

THEOREM 5.28 (Second arcsine law for Brownian motion). *Let $\{B(t) \,:\, t \geq 0\}$ be a standard linear Brownian motion. Then, $\mathcal{L}\big\{t \in [0,1] \,:\, B(t) > 0\big\}$, is arcsine distributed.*

The idea is to prove a direct relationship between the first maximum and the number of positive terms for a *simple* random walk by a combinatorial argument, and then transfer this to Brownian motion using Donsker's theorem.

LEMMA 5.29 (Richard's lemma). *Let $\{S_k \,:\, k = 1, \ldots, n\}$ be a simple, symmetric random walk on the integers. Then*

$$\#\big\{k \in \{1, \ldots, n\} \,:\, S_k > 0\big\} \stackrel{d}{=} \min\big\{k \in \{0, \ldots, n\} \,:\, S_k = \max_{0 \leq j \leq n} S_j\big\}. \tag{5.1}$$

Proof. Let $X_k = S_k - S_{k-1}$ for each $k \in \{1, \ldots, n\}$. We rearrange the tuple $(X_1, \ldots, X_n)$ by

- placing first in *decreasing* order of k the terms X_k for which $S_k > 0$,
- and then in *increasing* order of k the X_k for which $S_k \leq 0$.

Denote the new tuple by

$$(\widetilde{X}_1, \ldots, \widetilde{X}_n)$$

and let $\widetilde{S}_k$ be the associated k^{th} partial sum. We first show that

$$(X_1, \ldots, X_n) \stackrel{d}{=} (\widetilde{X}_1, \ldots, \widetilde{X}_n)\,.$$

Indeed, suppose first that all partial sums are nonpositive, then trivially the conditional distributions are the same. Next condition on the fact that k is the position of the last positive S_k. Note that under this condition the tuples $(X_1, \ldots, X_k)$ and $(X_{k+1}, \ldots, X_n)$ are still independent. Moreover, the $(X_1, \ldots, X_k)$ are independent and identically distributed random variables conditioned on the total sum being one, and therefore they are exchangeable. Hence the conditional law of of the vector $(X_k, X_1, \ldots, X_{k-1})$ is the same. Repeating this argument now for $(X_1, \ldots, X_{k-1})$ we see after finitely many steps that the two tuples have the same law.

Hence $\{\widetilde{S}_k \,:\, k = 1, \ldots, n\}$ is a random walk and we now check by induction on n that

$$\#\big\{k \in \{1, \ldots, n\} \,:\, S_k > 0\} = \min\{k \in \{0, \ldots, n\} \,:\, \widetilde{S}_k = \max_{0 \leq j \leq n} \widetilde{S}_j\big\}.$$

Indeed, this holds trivially for $n = 1$. When X_{n+1} is appended there are two possibilities:

- if $S_{n+1} > 0$, then $\widetilde{X}_1 = X_{n+1}$ and the position of the leftmost maximum in $\{\widetilde{S}_k : k = 0, \ldots, n\}$ is shifted by one position to the right.
- if $S_{n+1} \leq 0$, then $\widetilde{X}_{n+1} = X_{n+1}$ and the position of the leftmost maximum in $\{\widetilde{S}_k : k = 0, \ldots, n\}$ remains the same.

This completes the induction step and proves the lemma. ■

Proof of Theorem 5.28. Starting point is (5.1). First look at the right hand side of the equation, which divided by n can be written as $g(S_n^*)$ for the function $g\colon \mathcal{C}[0,1] \to [0,1]$ defined by

$$g(f) = \inf \big\{ t \in [0,1] \,:\, f(t) = \sup_{s \in [0,1]} f(s) \big\}\,.$$

The function g is continuous in every $f \in \mathcal{C}[0,1]$ which has a unique maximum, hence almost everywhere with respect to the Wiener measure. Hence, by combining Donsker's theorem and the Portmanteau theorem, the right hand side in (5.1) divided by n converges to the distribution of $g(B)$, which by Theorem 5.26 is the arcsine distribution.

Similarly, the left hand side of (5.1) can be approximated by $h(S_n^*)$ for the function $h\colon \mathcal{C}[0,1] \to [0,1]$ defined by

$$h(f) = \mathcal{L}\{t \in [0,1] \,:\, f(t) > 0\big\}\,.$$

The approximation error is bounded by

$$\tfrac{1}{n}\,\#\big\{k \in \{1, \ldots, n\} \,:\, S_k = 0\big\},$$

which converges to zero. The function h is obviously continuous in every $f \in \mathcal{C}[0,1]$ with the property that

$$\lim_{\varepsilon \downarrow 0} \mathcal{L}\{t \in [0,1] \,:\, 0 \leq f(t) \leq \varepsilon\big\} = 0,$$

which again is equivalent to $\mathcal{L}\{t \in [0,1] \,:\, f(t) = 0\big\} = 0$, a property which Brownian motion has almost surely. Hence, by combining Donsker's theorem and the Portmanteau theorem again, the left hand side in (5.1) divided by n converges to the distribution of

$$h(B) = \mathcal{L}\big\{t \in [0,1] \,:\, B(t) > 0\big\},$$

and this completes the argument. ■

Remark 5.30. The second arcsine law for general random walks follows from this using, by now, familiar arguments, see Exercise 5.10. ◇

Exercises

Exercise 5.1 ($*$). Suppose $\{B(t)\colon t \geq 0\}$ is a standard linear Brownian motion. Show that

$$\limsup_{n \uparrow 0} \sup_{n \leq t < n+1} \frac{B(t) - B(n)}{\sqrt{2 \log n}} = 1 \qquad \text{almost surely.}$$

Exercise 5.2 ($*$). Derive from Theorem 5.1 that, for a d-dimensional Brownian motion,

$$\limsup_{t \uparrow \infty} \frac{|B(t)|}{\sqrt{2t \log \log t}} = 1 \qquad \text{almost surely.}$$

Exercise 5.3 ($*$). Suppose $\{B(t)\colon t \geq 0\}$ is a linear Brownian motion and τ the first hitting time of level 1. Show that, almost surely,

$$\limsup_{h \downarrow 0} \frac{B(\tau) - B(\tau - h)}{\sqrt{2h \log \log(1/h)}} \leq 1.$$

Exercise 5.4 ($*$). Show that there are positive constants C_1 and C_2 such that

$$\frac{C_1}{\sqrt{n}} \leq \mathbb{P}\{S_i \geq 0 \ \text{ for all } \ 1 \leq i \leq n\} \leq \frac{C_2}{\sqrt{n}} \qquad \text{for all } n \geq 1.$$

Hint. For simple random walk a *reflection principle* holds in quite the same way as for Brownian motion. The key to the proof is to verify that

$$\mathbb{P}\{S_i \geq 0 \ \text{ for all } \ 1 \leq i \leq n\} = \mathbb{P}\{S_n \geq 0\} - \mathbb{P}\{S_n^* \leq -2\}$$

where S_n^* is the random walk reflected at the stopping time $\tau_{-1} = \min\{k\colon S_k = -1\}$.

Exercise 5.5 ($*$). Prove that, for any random walk $\{S_j : j \geq 0\}$ on the line,

$$\mathbb{P}\big\{S_0, \ldots, S_n \ \text{ has a point of increase}\big\} \leq 2\, \frac{\sum_{k=0}^{n} p_k p_{n-k}}{\sum_{k=0}^{\lfloor n/2 \rfloor} p_k^2},$$

where $p_0, \ldots, p_n$ are as in (2.3).

Exercise 5.6. An event $A \subset \mathbb{R}^d$ is an **increasing event** if,

$$(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots x_d) \in A \text{ and } \widetilde{x}_i \geq x_i$$
$$\Longrightarrow \qquad (x_1, \ldots, x_{i-1}, \widetilde{x}_i, x_{i+1}, \ldots x_d) \in A.$$

If A and B are increasing events, show that

$$\mathbb{P}(A \cap B) \geq \mathbb{P}(A)\mathbb{P}(B),$$

i.e. A and B are positively correlated.

Exercise 5.7 (∗). Show that we can obtain a lower bound on the probability that a random walk has a point of increase that differs from the upper bound only by a constant factor. More precisely, for any random walk on the line,

$$\mathbb{P}\big\{S_0, \ldots, S_n \text{ has a point of increase}\big\} \geq \frac{\sum_{k=0}^{n} p_k p_{n-k}}{2\sum_{k=0}^{\lfloor n/2 \rfloor} p_k^2},$$

where $p_0, \ldots, p_n$ are as in (2.3).

Exercise 5.8. Suppose $X_1, \ldots, X_n$ are independent and identically distributed and consider their ordered relabeling given by $X_{(1)} \geq X_{(2)} \geq \ldots \geq X_{(n)}$. Show that

$$\mathbb{E}[X_{(i)}X_{(j)}] \geq \mathbb{E}[X_{(i)}]\mathbb{E}[X_{(j)}],$$

provided these expectations are well-defined.

Exercise 5.9 (∗). Given a centred random variable X, show that there exist centred random variables X_n taking only finitely many values, such that X_n converges to X in law and, for $\Psi_n(x) = E\big[X_n \,\big|\, X_n \geq x\big]$, the embedding stopping times

$$\tau_n = \inf\{t \geq 0 \,:\, M(t) \geq \Psi_n(B(t))\}$$

converge almost surely to τ. Infer that $B(\tau)$ has the same law as X, and $\mathbb{E}[\tau] = \mathbb{E}[X^2]$.

Exercise 5.10. Suppose that $\{X_k \,:\, k \geq 1\}$ is a sequence of independent, identically distributed random variables with $\mathbb{E}[X_1] = 0$, $\mathbb{P}\{X_1 = 0\} = 0$ and $0 < \mathbb{E}[X_1^2] = \sigma^2 < \infty$. Let $\{S_n \,:\, n \geq 0\}$ be the associated random walk and

$$P_n = \#\{1 \leq k \leq n \,:\, S_k > 0\}$$

the number of positive values of the random walk before time n. Then, for all $x \in (0, 1)$,

$$\lim_{n\to\infty} \mathbb{P}\{P_n \leq xn\} = \frac{2}{\pi} \arcsin(\sqrt{x})\,.$$

Notes and Comments

Historically, the law of the iterated logarithm was first proved for simple random walk by Khinchin [**Kh23, Kh24**] and later generalised to other random walks by Kolmogorov [**Ko29**] and Hartman and Wintner [**HW41**]. The original arguments of Kolmogorov, Hartman and Wintner were extremely difficult, and a lot of authors have since provided more accessible proofs, see, for example, de Acosta [**dA83**]. For Brownian motion the law of the iterated logarithm is also due to Khinchin [**Kh33**]. The idea of using embedding arguments to transfer the result from the Brownian motion to the random walk case is due to Strassen [**St64**]. For a survey of laws of the iterated logarithm, see [**Bi86**].

An extension of the law of the iterated logarithm is Strassen's law, which is first proved in [**St64**]. If a standard Brownian motion on the interval $[0,t]$ is rescaled by a factor $1/t$ in time and a factor $\sqrt{2t\log\log(1/t)}$ in space, the set of limit points in $\mathcal{C}[0,1]$ are exactly the functions f with $f(0)=0$ and $\int_0^1 (f'(t))^2\,dt \le 1$. Strassen's law also explains the approximate form of the curve in the right half of Figure 5.1. Any function in this class with $f(1)=1$ satisfies

$$1 \ge \int_0^1 (f'(t))^2\,dt \ge \Big(\int_0^1 f'(t)\,dt\Big)^2 = 1,$$

which implies that $f'(t)$ is constant and thus $f(t)=t$ for all $t\in(0,1)$. Therefore, for large t, the Brownian path conditioned on ending near to its upper envelope resembles a straight line in the sup-norm, as can be seen in Figure 5.1.

The nonincrease phenomenon, which is described in Theorem 5.11, holds for arbitrary symmetric random walks, and can thus be viewed as a combinatorial consequence of fluctuations in random sums. Indeed, our argument shows this — subject to a generalisation of Lemma 5.8. The latter result holds if the increments X_i have a symmetric distribution, or if the increments have mean zero and finite variance, see e.g. Feller [**Fe66**, Section XII.8]. Dvoretzky, Erdős and Kakutani [**DEK61**] were the first to prove that Brownian motion almost surely has no local points of increase. Knight [**Kn81**] and Berman [**Be83**] noted that this follows from properties of the local time of Brownian motion; direct proofs were given by Adelman [**Ad85**] and Burdzy [**Bu90**]. The proof we give is taken from [**Pe96c**].

A higher-dimensional analogue of this question is whether, for Brownian motion in the plane, there exists a line such that the Brownian motion path, projected onto that line, has a global point of increase, or equivalently whether the Brownian motion path admits cut lines. We say a line ℓ is a *cut line* for the Brownian motion if, for some $t_0\in(0,1)$ with $B(t_0)\in\ell$, the points $B(t)$ lies on one side of ℓ for all $t\in[0,t_0)$ and on the other side of ℓ for all $t\in(t_0,1]$. It was proved by Bass and Burdzy [**BB97**] that planar Brownian motion almost surely does *not* have cut lines. Burdzy [**Bu89**], with a correction to the proof in [**Bu95**], however showed that Brownian motion in the plane almost surely does have *cut points*, which are points $B(t)$ such that the Brownian motion path with the point $B(t)$ removed is disconnected. It was conjectured that the Hausdorff dimension of the set of cut points is $3/4$. This conjecture has recently been proved by Lawler, Schramm and Werner [**LSW01**], see also [**La96a**].

For Brownian motion in three dimensions, there almost surely exist cut planes, where we say P is a *cut plane* if for some t, the set $\{B(s)\colon 0 < s < t\}$ lies on one side of the plane and the set $\{B(s)\colon 1 > s > t\}$ on the other side. This result, original due to Pemantle, is also described in Bass and Burdzy [**BB97**]. An argument of Evans, which is closely related to material we discuss in the final section of Chapter 10, shows that the set of times corresponding to cut planes has Hausdorff dimension zero.

Pemantle [**Pe97**] has shown that the range of planar Brownian motion almost surely does not cover any straight line segment. Which curves can and which cannot be covered by a Brownian motion path is, in general, an open question. Also unknown is the minimal Hausdorff dimension of curves contained in the range of planar Brownian motion, though it is known that it contains a curve of Hausdorff dimension 4/3, namely its outer boundary, see [**LSW01**].

Harris' inequality was discovered by Harris [**Ha60**] and is also known as *FKG inequality* in recognition of the work of Fortuin, Kasteleyn and Ginibre [**FKG71**] who extended the original inequality beyond the case of product measures. 'Correlation inequalities' like these play an extremely important role in percolation theory and spatial statistical physics. Exercise 5.8 indicates the important role of this idea in the investigation of order statistics, see Lehmann [**Le66**] and Bickel [**Bi67**] for further discussion and applications.

The Skorokhod embedding problem is a classic, which still leads to some attractive research. The first embedding theorem is due to Skorokhod [**Sk65**]. The Russian original of this work appeared in 1961 and the Dubins embedding, which we have presented is not much younger, see [**Du68**]. Our presentation, based on the idea of binary splitting martingales, follows Neveu [**Ne75**, Ex. II.7, p 34] and we thank Jim Pitman for directing us to this reference. Another classic embedding technique is Root's embedding, see [**Ro69**]. The Azéma-Yor embedding was first described in [**AY79**], but we follow Meilijson [**Me83**] in the proof. One of the attractive features of the Azéma-Yor embedding is that, among all stopping times T with $\mathbb{E}T < \infty$ which represent a given random variable X, it maximizes the $\max_{0\le t\le T} B(t)$. Generalisation of the embedding problem to more general classes of probability laws require different forms of minimality for the embedding stopping time, or more general processes in which one embeds. A survey of current developments is [**Ob04**].

The idea of an invariance principle that allows to transfer limit theorems from special cases to general random walks can be traced to Erdős and Kac [**EK46, EK47**]. The first general result of this nature is due to Donsker [**Do51**] following an idea of Doob [**Do49**]. Besides the embedding technique carried out in our proof, there is also a popular alternative proof, which goes back to Prohorov [**Pr56**]. Suppose that a subsequence of $\{S_n^* : n \ge 1\}$ converges in distribution to a limit X. This limit is a continuous random function, which is easily seen to have stationary, independent increments, which have expectation zero and variance equal to their length. By a general result this implies that X is a Brownian motion. So Brownian motion is the only possible limit point of the sequence $\{S_n^* : n \ge 1\}$. The difficult part of this proof is now to show that every subsequence of $\{S_n^* : n \ge 1\}$ has a convergent subsubsequence, the *tightness property*. Many interesting applications and extensions of Donsker's theorem can be found in [**Bi68**].

An important class of extensions of Donsker's theorem are the strong approximation theorems which were provided by Skorokhod [**Sk65**] and Strassen [**St64**]. In these results the Brownian motion and the random walk are constructed on the same probability space in such a way that they are close almost surely. An optimal result in this direction is the famous paper of Komlós, Major and Tusnády [**KMT75**]. For an exposition of their work and applications, see [**CR81**].

The arcsine laws for Brownian motion were first proved by Lévy in [**Le39, Le48**]. The proof of the first law, which we give here, follows Kallenberg [**Ka02**]. This law can also be proved by a direct calculation, which however is slightly longer, see for example [**Du95**]. Our proof of the second arcsine law goes back to an idea of Baxter [**Ba62**].

CHAPTER 6

Brownian local time

In this chapter we focus on linear Brownian motion and address the question how to measure the amount of time spent by a Brownian path at a given level. As we already know from Theorem 3.25 that the occupation times up to time t are absolutely continuous measures, their densities are a viable measure for the time spent at level a during the time interval $[0,t]$. We shall show that these densities make up a continuous random field $\{L^a(t)\colon a \in \mathbb{R},\, t \geq 0\}$, which is called the Brownian local time. Nontrivial information about the distribution of this process is contained in a theorem of Lévy (studying it as function of time) and the Ray-Knight theorem (studying it as function of the level). We finally show how to interpret local time as a family of Hausdorff measures.

1. The local time at zero

How can we measure the amount of time spent by a standard linear Brownian motion $\{B(t)\colon t \geq 0\}$ at zero? We have already seen that, almost surely, the zero set $\{s \in [0,t)\colon B(s) = 0\}$ is a set of Hausdorff dimension 1/2. Moreover, by Exercise 4.13, the 1/2-dimensional Hausdorff measure of the zero set is zero, so Hausdorff measure as defined so far does not give a nontrivial answer.

We approach this problem by counting the number of downcrossings of a nested sequence of intervals decreasing to zero. More precisely, for a linear Brownian motion $\{B(t)\colon t \geq 0\}$ with arbitrary starting point, given $a < b$, we define stopping times $\tau_0 = 0$ and, for $j \geq 1$,

$$\sigma_j = \inf\big\{t > \tau_{j-1}\colon B(t) = b\big\}, \qquad \tau_j = \inf\big\{t > \sigma_j\colon B(t) = a\big\}. \tag{1.1}$$

We call the random functions

$$B^{(j)}\colon [0, \tau_j - \sigma_j] \to \mathbb{R}, \qquad B^{(j)}(s) = B(\sigma_j + s)$$

the jth downcrossing of $[a,b]$. For every $t > 0$ we denote by

$$D(a,b,t) = \max\big\{j \in \mathbb{N} \,:\, \tau_j \leq t\big\}$$

the number of *downcrossings* of the interval $[a,b]$ before time t. Note that $D(a,b,t)$ is almost surely finite by the absolute continuity of Brownian motion on the compact interval $[0,t]$.

Theorem 6.1 (Downcrossing representation of the local time at zero). *There exists a nontrivial stochastic process $\{L(t)\colon t \geq 0\}$ called the* **local time at zero** *such that for any sequences $a_n \uparrow 0$ and $b_n \downarrow 0$ with $a_n < b_n$, almost surely,*

$$\lim_{n\to\infty} 2(b_n - a_n)\, D(a_n, b_n, t) = L(t) \text{ for every } t > 0.$$

Moreover, this process is almost surely locally γ-Hölder continuous for any $\gamma < 1/2$.

Remark 6.2. One might wonder about the meaning of the normalisation factor 2 in our theorem, which is omitted in some treatments (e.g. [**KS88**]). An intuitive answer is that the time spent near zero should be approximated by the number of downcrossings *plus* the number of upcrossings of a small interval centred in zero. The factor thus compensates for the number of upcrossings, which is (up to an error of at most one) the same as the number of downcrossings. ⋄

The key ingredient of the proof of Theorem 6.1 is the following fact.

Lemma 6.3. *Suppose that $a < m < b$ and let $\{B(t) : 0 \le t \le T\}$ be a linear Brownian motion stopped at the time T when it first hits a given level above b. Let*

- *D be the number of downcrossings of the interval $[a, b]$,*
- *D_{l} be the number of downcrossings of the interval $\big[a, m\big]$,*
- *D_{u} be the number of downcrossings of the interval $\big[m, b\big]$.*

There exist two independent sequences $X_0, X_1, \ldots$ and $Y_0, Y_1, \ldots$ of independent nonnegative random variables, which are also independent of D, such that for $j \ge 1$ the random variables X_j are geometric with mean $(b-a)/(m-a)$ and the random variables Y_j are geometric with mean $(b-a)/(b-m)$, and

$$D_{\mathrm{l}} = X_0 + \sum_{j=1}^{D} X_j \qquad \text{and} \qquad D_{\mathrm{u}} = Y_0 + \sum_{j=1}^{D} Y_j\,.$$

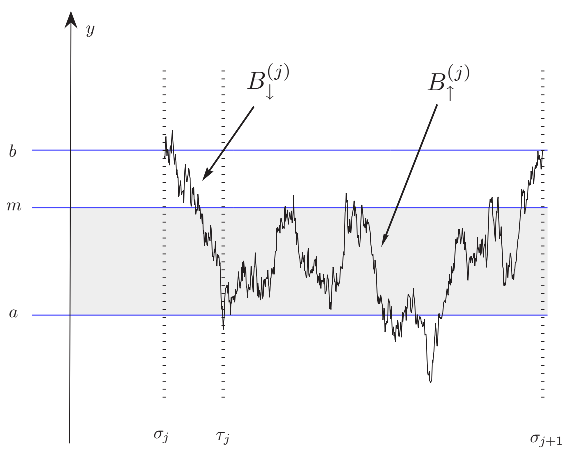

Figure 1. The downcrossing of $[a, b]$ contains one downcrossing of $[a, m]$ and the following upcrossing of $[a, b]$ contains one further downcrossing of $[a, m]$.

Proof. Recall the definition of the stopping times σ_j, τ_j from (1.1). For $j \geq 1$, define the j^{th} downcrossings, resp. upcrossings, of $[a,b]$ by

$$B_\downarrow^{(j)} \colon [0, \tau_j - \sigma_j] \to \mathbb{R}, \qquad B_\downarrow^{(j)}(s) = B(\sigma_j + s)$$
$$B_\uparrow^{(j)} \colon [0, \sigma_{j+1} - \tau_j] \to \mathbb{R}, \qquad B_\uparrow^{(j)}(s) = B(\tau_j + s),$$

By the strong Markov property all these pieces of the Brownian path are independent. Note that D depends only on the family $(B_\downarrow^{(j)} \colon j \geq 1)$ of downcrossings.

First look at D_{l} and denote by X_0 the number of downcrossings of $[a,m]$ before the first downcrossing of $[a,b]$. The j^{th} downcrossing of $[a,b]$ contains exactly one downcrossing of $[a,m]$ and the j^{th} upcrossing of $[a,b]$ contains a random number X_j of downcrossings of $[a,m]$, which, by Theorem 2.45, satisfies

$$\mathbb{P}\big\{X_j = k\big\} = \Big(\frac{m-a}{b-a}\Big)\Big(\frac{b-m}{b-a}\Big)^{k-1} \qquad \text{for every } k \in \{1, \ldots\}.$$

In other words X_j is geometrically distributed with (success) parameter $(m-a)/(b-a)$.

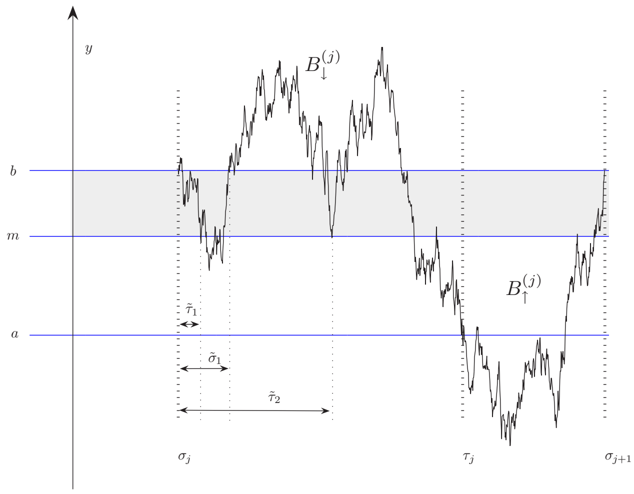

FIGURE 2. The downcrossing of $[a,b]$ contains three downcrossings of $[m,b]$ and the following upcrossing of $[a,b]$ contains no further downcrossings of $[m,b]$.

Second look at D_{u} and denote by Y_0 the number of downcrossings of $[m,b]$ after the last downcrossing of $[a,b]$. No downcrossings of $[m,b]$ can occur during an upcrossing of $[a,b]$. Fix a j and look at the downcrossing $B_\downarrow^{(j)}$ of $[a,b]$. Define stopping times $\tilde\sigma_0 = 0$ and, for $i \geq 1$,

$$\tilde\tau_i = \inf\big\{t > \tilde\sigma_{i-1} \colon B_\downarrow^{(j)}(t) = m\big\}, \qquad \tilde\sigma_i = \inf\big\{t > \tilde\tau_i \colon B_\downarrow^{(j)}(t) = b\big\}.$$

This subdivides the path of $B_\downarrow^{(j)}$ into independent downcrossing periods $[\tilde\sigma_{i-1}, \tilde\tau_i]$, and upcrossing periods $[\tilde\tau_i, \tilde\sigma_i]$ of $[m,b]$. By our assumption the upper hitting boundary is above b and therefore can only be hit during the downcrossing periods, while the lifetime of $B_\downarrow^{(j)}$ expires when the lower boundary a is hit, which can only occur during an upcrossing period. The probability of this event equals $(b-m)/(b-a)$ by Theorem 2.45.

Hence the number of downcrossings of $[m, b]$ during the j^{th} downcrossing of $[a, b]$ is a geometric random variable Y_j with (success) parameter $(b-m)/(b-a)$, which completes the proof. ∎

For the proof of Theorem 6.1 we first prove the convergence for the case when the Brownian motion is stopped at the time $T = T_b$ when it first reaches some level $b > b_1$. This has the advantage that there cannot be any uncompleted upcrossings.

LEMMA 6.4. *For any two sequences $a_n \uparrow 0$ and $b_n \downarrow 0$ with $a_n < b_n$, the discrete time stochastic process*

$$\Big\{2(b_n - a_n)\, D(a_n, b_n, T) \colon n \in \mathbb{N}\Big\}$$

is a submartingale with respect to the natural filtration $(\mathcal{F}_n \colon n \in \mathbb{N})$.

Proof. We may assume that, for each n, we have

$$\textit{either } (1)\ a_n = a_{n+1} \qquad \textit{or } (2)\ b_n = b_{n+1}\,,$$

which is no loss of generality, as we may replace a step where both a_n and b_n are changed by two steps, where only one is changed at the time. The original sequence is then a subsequence of the modified one and inherits the submartingale property.

Now fix n and *first* assume that we are in case (1) $a_n = a_{n+1}$. By Lemma 6.3 for D_{l}, the total number $D(a_n, b_{n+1}, T)$ of downcrossings of $[a_n, b_{n+1}]$ given $\mathcal{F}_n$ is the sum of $D(a_n, b_n, T)$ independent geometric random variables with parameter $(b_{n+1}-a_n)/(b_n-a_n)$ plus a nonnegative contribution. Hence,

$$\mathbb{E}\big[(b_{n+1} - a_n)\, D(a_n, b_{n+1}, T) \,\big|\, \mathcal{F}_n\big] \geq (b_n - a_n)\, D(a_n, b_n, T),$$

which is the submartingale property (for the nth step).

Second assume that we are in case (1) $b_n = b_{n+1}$. Then Lemma 6.3 for D_{u} shows that the number of downcrossings of $[a_{n+1}, b_n]$ given $\mathcal{F}_n$ is the sum of $D(a_n, b_n, T)$ independent geometric random variables with parameter $(b_n - a_{n+1})/(b_n - a_n)$ plus a nonnegative contribution. Hence

$$\mathbb{E}\big[(b_n - a_{n+1})\, D(a_{n+1}, b_n, T) \,\big|\, \mathcal{F}_n\big] \geq (b_n - a_n)\, D(a_n, b_n, T),$$

and together with the first case this establishes that $\{2(b_n - a_n)\, D(a_n, b_n, T) \colon n \in \mathbb{N}\}$ is a submartingale with respect to its natural filtration. ∎

LEMMA 6.5. *For any two sequences $a_n \uparrow 0$ and $b_n \downarrow 0$ with $a_n < b_n$ the limit*

$$L(T_b) := \lim_{n\to\infty} 2(b_n - a_n)\, D(a_n, b_n, T_b)$$

exists almost surely. It does not depend on the choice of sequences.

Proof. Observe that $D(a_n, b_n, T_b)$ is a geometric random variable with parameter $(b_n - a_n)/(b - a_n)$. Recall that the variance of a geometric random variable with parameter p is $(1-p)/p^2$, and so its second moment is bounded by $2/p^2$. Hence

$$\mathbb{E}\big[4(b_n - a_n)^2\, D(a_n, b_n, T_b)^2\,\big] \leq 8\,(b - a_n)^2,$$

and thus the submartingale in Lemma 6.4 is L^2-bounded. By the submartingale convergence theorem, see Theorem II.4.5, the limit

$$\lim_{n\uparrow\infty} 2(b_n - a_n)\, D(a_n, b_n, T_b)$$

exists almost surely, and by Theorem II.4.12 also in L^2 ensuring that the limit is nontrivial. Finally, note that the limit does not depend on the choice of the sequence $a_n \uparrow 0$ and $b_n \downarrow 0$ because if it did, then given two sequences with different limits we could construct a sequence of intervals alternating between the two sequences, which would not converge. ■

LEMMA 6.6. *For any fixed time $t > 0$, almost surely, the limit*

$$L(t) := \lim_{n\to\infty} 2(b_n - a_n)\, D(a_n, b_n, t) \qquad \textit{exists.}$$

Proof. We define an auxiliary Brownian motion $\{B_t(s) \colon s \geq 0\}$ by $B_t(s) = B(t+s)$. For any integer $b > b_1$ we denote by $D_t(a_n, b_n, T_b)$ the number of downcrossings of the interval $[a_n, b_n]$ by the auxiliary Brownian motion before it hits b. Then, almost surely,

$$L_t(T_b) := \lim_{n\uparrow\infty} 2(b_n - a_n)\, D_t(a_n, b_n, T_b),$$

exists by the previous lemma. Given $t > 0$ we fix a Brownian path such that this limit exists for all integers $b > b_1$. Pick b so large that $T_b > t$. Define

$$L(t) := L(T_b) - L_t(T_b)\,.$$

To show that this is the required limit, observe that

$$D(a_n, b_n, T_b) - D_t(a_n, b_n, T_b) - 1 \leq D(a_n, b_n, t) \leq D(a_n, b_n, T_b) - D_t(a_n, b_n, T_b),$$

where the correction -1 on the left hand side arises from the possibility that t interrupts a downcrossing. Multiplying by $2(b_n - a_n)$ and taking a limit gives $L(T_b) - L_t(T_b)$ for both bounds, proving convergence. ■

We now have to study the dependence of $L(t)$ on the time t in more detail. To simplify the notation we write

$$I_n(s,t) = 2(b_n - a_n)\big(D(a_n, b_n, t) - D(a_n, b_n, s)\big) \qquad \text{for all } 0 \leq s < t\,.$$

The following lemma contains a probability estimate, which is sufficient to get the convergence of the downcrossing numbers jointly for all times and to establish Hölder continuity.

LEMMA 6.7. *Let $\gamma < 1/2$ and $0 < \varepsilon < (1-2\gamma)/3$. Then, for all $t \geq 0$ and $0 < h < 1$, we have*

$$\mathbb{P}\big\{L(t+h) - L(t) > h^\gamma\big\} \leq 2\, \exp\{-\tfrac{1}{2}\, h^{-\varepsilon}\}\,.$$

Proof. As, by Fatou's lemma,

$$\mathbb{P}\big\{L(t+h) - L(t) > h^\gamma\big\} = \mathbb{P}\big\{\liminf_{n\to\infty} I_n(t, t+h) > h^\gamma\big\} \leq \liminf_{n\to\infty} \mathbb{P}\big\{I_n(t, t+h) > h^\gamma\big\}$$

we can focus on estimating $\mathbb{P}\{I_n(t, t+h) > h^\gamma\}$ for fixed large n. By the Markov property it suffices to estimate $\mathbb{P}_x\{I_n(0,h) > h^\gamma\}$ uniformly for all $x \in \mathbb{R}$. This probability is clearly maximal when $x = b_n$, so we may assume this. Let $T_h = \inf\{s > 0\colon B(s) = b_n + h^{(1-\varepsilon)/2}\}$ and observe that

$$\big\{I_n(0,h) > h^\gamma\big\} \subset \big\{I_n(0,T_h) > h^\gamma\big\} \cup \big\{T_h < h\big\}.$$

The number of downcrossings of $[a_n, b_n]$ during the period before T_h is geometrically distributed with mean $(b_n - a_n)^{-1} h^{(1-\varepsilon)/2} + 1$ and thus

$$\begin{aligned}\mathbb{P}_{b_n}\big\{I_n(0,T_h) > h^\gamma\big\} &= \Big(\frac{h^{(1-\varepsilon)/2}}{b_n - a_n + h^{(1-\varepsilon)/2}}\Big)^{\lceil \frac{1}{2(b_n-a_n)}\, h^\gamma \rceil - 1} \\ &\stackrel{n\to\infty}{\longrightarrow} \exp\big\{ -\tfrac12\, h^{\gamma - \frac12 + \frac{\varepsilon}{2}}\big\} \le \exp\big\{ -\tfrac12\, h^{-\varepsilon}\big\}.\end{aligned}$$

With $\{W(s)\colon s \ge 0\}$ denoting a standard linear Brownian motion,

$$\mathbb{P}_{b_n}\big\{T_h < h\big\} = \mathbb{P}\Big\{\max_{0\le s\le h} W(s) \ge h^{(1-\varepsilon)/2}\Big\} \le \sqrt{\tfrac{2}{\pi h^{-\varepsilon}}}\, \exp\big\{ -\tfrac12\, h^{-\varepsilon}\big\}$$

where we have used Remark 2.19 in the last step. The result follows by adding the last two displayed formulas. ∎

Lemma 6.8. *Almost surely,*

$$L(t) := \lim_{n\to\infty} 2(b_n - a_n)\, D(a_n, b_n, t)$$

exists for every $t \ge 0$.

Proof. It suffices to prove the simultaneous convergence for all $0 \le t \le 1$. We define a countable set of gridpoints

$$\mathcal{G} = \bigcup_{m\in\mathbb{N}} \mathcal{G}_m \cup \{1\}, \qquad \text{for } \mathcal{G}_m = \big\{\tfrac{k}{m}\colon k \in \{0,\ldots,m-1\}\big\}$$

and show that the stated convergence holds on the set

$$\bigcup_{t\in\mathcal{G}} \big\{L(t) = \lim_{n\to\infty} 2(b_n - a_n)\, D(a_n, b_n, t) \text{ exists }\big\} \cup \bigcup_{m>M} \bigcup_{t\in\mathcal{G}_m} \big\{L(t + \tfrac1m) - L(t) \le (1/m)^\gamma\big\}.$$

which, by choosing M suitably, has probability arbitrarily close to one by the previous two lemmas. Given any $t \in [0,1)$ and a large m we find $t_1, t_2 \in \mathcal{G}_m$ with $t_2 - t_1 = \frac1m$ and $t \in [t_1, t_2]$. We obviously have

$$2(b_n - a_n)\, D(a_n, b_n, t_1) \le 2(b_n - a_n)\, D(a_n, b_n, t) \le 2(b_n - a_n)\, D(a_n, b_n, t_2).$$

Both bounds converge on our set, and the difference of the limits is $L(t_2) - L(t_1)$, which is bounded by $m^{-\gamma}$ and thus can be made arbitrarily small by choosing a large m. ∎

Lemma 6.9. *For $\gamma < \frac12$, almost surely, the process $\{L(t)\colon t \ge 0\}$ is locally γ-Hölder continuous.*

Proof. It suffices to look at $0 \le t < 1$. We use the notation of the proof of the previous and show that γ-Hölder continuity holds on the set constructed there. Indeed, whenever $0 \le s < t < 1$ and $t - s < 1/M$ we pick $m \ge M$ such that

$$\tfrac{1}{m+1} \le t - s < \tfrac1m\,.$$

We take $t_1 < s$ with $t_1 \in \mathcal{G}_m$ and $s - t_1 < 1/m$, and $t_2 > t$ with $t_2 \in \mathcal{G}_m$ and $t_2 - t < 1/m$.

Note that $t_2 - t_1 \le 2/m$ by construction and hence,

$$L(t) - L(s) \le L(t_2) - L(t_1) \le 2(1/m)^\gamma \le 2\big(\tfrac{m+1}{m}\big)^\gamma\,(t-s)^\gamma.$$

The result follows as the fraction on the right is bounded by 2. ∎

This completes the proof of the downcrossing representation, Theorem 6.1. It is easy to see from this representation that, almost surely, the local time at zero increases only on the zero set of the Brownian motion, see Exercise 6.1.

Observe that the increasing process $\{L(t) : t \ge 0\}$ is *not* a Markov process. Heuristically, the size of the increment $L(t+h) - L(t)$ depends on the position of the first zero of the Brownian motion after time t, which is strongly dependent on the position of the last zero before time t. The last zero however is the position of the last point of increase of the local time process before time t, and therefore the path $\{L(s) : 0 \le s \le t\}$ contains relevant information beyond its endpoint. Nevertheless, we can describe the law of the local time process, thanks to the following famous theorem of Paul Lévy, which describes the law of the local time at zero in terms of the maximum process of Brownian motion. It opens the door to finer results on the local time at zero, like those presented in Section 4 of this chapter.

THEOREM 6.10 (Lévy). *The local time at zero $\{L(t)\colon t \ge 0\}$ and the maximum process $\{M(t)\colon t \ge 0\}$ of a standard linear Brownian motion have the same distribution.*

The proof uses the simple random walk embedded in the Brownian motion, a technique which we will exploit extensively in the next section. Define stopping times $\tau_0 := \tau_0^{(n)} := 0$ and

$$\tau_k := \tau_k^{(n)} := \inf\big\{t > \tau_{k-1}\colon |B(t) - B(\tau_{k-1})| = 2^{-n}\big\}, \qquad \text{for } k \ge 1\,.$$

The nth embedded random walk $\{X_k^{(n)}\colon k = 1, 2, \ldots\}$ is defined by

$$X_k := X_k^{(n)} := 2^n\, B\big(\tau_k^{(n)}\big)\,.$$

The length of the embedded random walk is

$$N := N^{(n)}(t) := \max\{k \in \mathbb{N}\colon \tau_k \le t\}\,,$$

which is easily seen to be independent of the actual walk.

LEMMA 6.11. *For every $t > 0$, almost surely, $\lim_{n\to\infty} 2^{-2n} N^{(n)}(t) = t$.*

Proof. First note that $\{\xi_k^{(n)}\colon k = 1, 2, \ldots\}$ defined by

$$\xi_k := \xi_k^{(n)} := \tau_k^{(n)} - \tau_{k-1}^{(n)}$$

is a sequence of independent random variables, for each n. By Theorem 2.45 the mean of ξ_k is 2^{-2n} and its variance is, by Brownian scaling, equal to $c2^{-4n}$ for some constant $c > 0$. (See, for example, Exercise 2.13 for instructions how to find the constant.) Define

$$S^{(n)}(t) = \sum_{k=1}^{\lceil 2^{2n} t\rceil} \xi_k^{(n)}\,.$$

Then $\mathbb{E}S^{(n)}(t) = \lceil 2^{2n}t\rceil 2^{-2n} \to t$ and $\mathrm{Var}\big(S^{(n)}(t)\big) = c2^{-4n}\lceil 2^{2n}t\rceil$, hence

$$\mathbb{E}\sum_{n=1}^{\infty}\big(S^{(n)}(t) - \mathbb{E}S^{(n)}(t)\big)^2 < \infty\,.$$

We infer that, almost surely, $\lim_{n\to\infty} S^{(n)}(t) = t$. For fixed $\varepsilon > 0$, we pick n_0 large so that

$$S^{(n)}(t-\varepsilon) \le t \le S^{(n)}(t+\varepsilon) \text{ for all } n \ge n_0.$$

The sum over ξ_k up to $N^{(n)}(t) + 1$ is at least t, by definition, and hence we get $N^{(n)}(t) + 1 \ge \lceil 2^{2n}(t-\varepsilon)\rceil$. Conversely, the sum over ξ_k up to $N^{(n)}(t)$ is at most t and hence $N^{(n)}(t) \le \lceil 2^{2n}(t+\varepsilon)\rceil$. The result follows as $\varepsilon > 0$ was arbitrary. ∎

Lemma 6.12. *Almost surely, for every $t > 0$,*

$$\lim_{n\uparrow\infty} 2^{-n}\#\big\{k \in \{1,\dots,N^{(n)}(t)\} \,:\, |X_{k-1}| = 0, |X_k| = 1\big\} = L(t)$$

Proof. By Theorem 6.1 applied to the sequences $a_n = -2^{-n}$ and $b_n = 0$ we have

$$\lim_{n\uparrow\infty} 2^{-n}\#\big\{k \in \{1,\dots,N^{(n)}(t)\} \,:\, X_{k-1} = 0, X_k = -1\big\} = \tfrac{1}{2}L(t)\,.$$

Applying Theorem 6.1 to the sequences $a_n = 0$ and $b_n = 2^{-n}$ we get

$$\lim_{n\uparrow\infty} 2^{-n}\#\big\{k \in \{1,\dots,N^{(n)}(t)\} \,:\, X_{k-1} = 1, X_k = 0\big\} = \tfrac{1}{2}L(t)\,.$$

As $\#\{k \le N\colon X_{k-1} = 1, X_k = 0\}$ and $\#\{k \le N\colon X_{k-1} = 0, X_k = 1\}$ differ by no more than one, the result follows by adding up the two displayed formulas. ∎

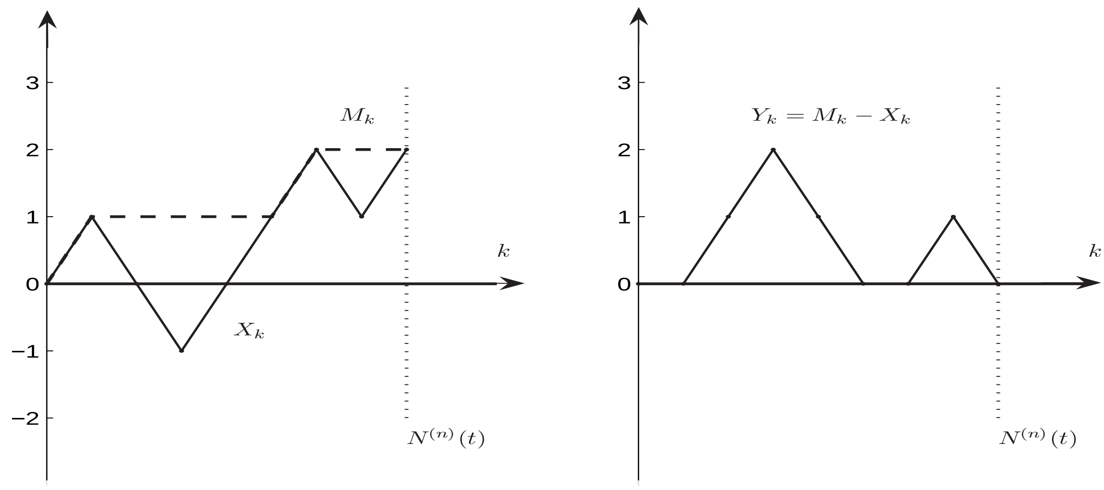

Figure 3. On the left an embedded random walk $\{X_k\colon k \ge 0\}$ together with its maximum process $\{M_k\colon k \ge 0\}$. On the right the associated difference process $\{Y_k\colon k \ge 0\}$ defined by $Y_k = M_k - X_k$.

We define the maximum process $\{M_k^{(n)}: k = 1, 2, \ldots\}$ associated with the embedded random walk by

$$M_k = M_k^{(n)} = \max\big\{X_j^{(n)} \,:\, j \in \{0, \ldots, k\}\big\}.$$

Then the process $\{Y_k^{(n)}: k = 1, 2, \ldots\}$ defined by $Y_k := Y_k^{(n)} := M_k - X_k$ is a Markov chain with statespace $\{0, 1, 2, \ldots\}$ and the following transition mechanism

- if $j \neq 0$ then $\mathbb{P}\{Y_{k+1} = j + 1 \,|\, Y_k = j\} = \frac{1}{2} = \mathbb{P}\{Y_{k+1} = j - 1 \,|\, Y_k = j\}$,
- $\mathbb{P}\{Y_{k+1} = 0 \,|\, Y_k = 0\} = \frac{1}{2} = \mathbb{P}\{Y_{k+1} = 1 \,|\, Y_k = 0\}$.

One can recover the maximum process $\{M_k: k = 1, 2, \ldots\}$ from $\{Y_k: k = 1, 2, \ldots\}$ by counting the number of flat steps

$$M_k = \#\big\{j \in \{1, \ldots, k\}: Y_j = Y_{j-1}\big\}.$$

Hence we obtain, asymptotically, the maximum process of the Brownian motion as a limit of the number of flat steps in $\{Y_k^{(n)}: k = 1, 2, \ldots\}$.

Lemma 6.13. *For any time $t > 0$, almost surely,*

$$M(t) = \lim_{n\uparrow\infty} 2^{-n}\, \#\big\{j \in \{1, \ldots, N^{(n)}(t)\}: Y_j^{(n)} = Y_{j-1}^{(n)}\big\}.$$

Proof. Note that $\#\big\{j \in \{1, \ldots, N\}: Y_j = Y_{j-1}\big\}$ is the maximum of the random walk $\{X_k: k = 1, 2, \ldots, N\}$ over its entire length. This maximum, multiplied by 2^{-n}, differs from $M(t)$ by no more than 2^{-n}, and this completes the argument. ■

Removing the flat steps in the process $\{Y_j^{(n)}: j = 1, 2, \ldots\}$ we obtain a process $\{\tilde{Y}_k^{(n)}: k = 1, 2, \ldots\}$, which has the same law as $\{|X_k|: k = 1, 2, \ldots\}$. By Lemma 6.12 we therefore have the convergence in distribution, as $n \uparrow \infty$,

$$2^{-n}\#\big\{k \in \{1, \ldots, N^{(n)}(t)\} \,:\, \tilde{Y}_{k-1}^{(n)} = 0, \tilde{Y}_k^{(n)} = 1\big\} \Longrightarrow L(t), \tag{1.2}$$

jointly for any finite set of times.

Lemma 6.14. *Almost surely,*

$$\lim_{n\uparrow\infty} 2^{-n}\Big(\#\big\{j \in \{1, \ldots, N^{(n)}(t)\}: Y_{j-1}^{(n)} = Y_j^{(n)}\big\} \\ - \#\big\{k \in \{1, \ldots, N^{(n)}(t)\} \,:\, \tilde{Y}_{k-1}^{(n)} = 0, \tilde{Y}_k^{(n)} = 1\big\}\Big) = 0.$$

Proof. First note that when $\{Y_j: j = 1, 2, \ldots\}$ returns to zero for the ith time, the number of steps before it moves to one is given by a random variable Z_i with distribution

$$\mathbb{P}\{Z_i = k\big\} = 2^{-k-1} \text{ for } k = 0, 1, \ldots.$$

Denoting by Z_0 the number of steps before it moves initially, the random variables $Z_0, Z_1, \ldots$ are independent and independent of the process $\{\tilde{Y}_k^{(n)}: k = 1, 2, \ldots\}$. Let

$$A^{(n)} = \#\big\{j \in \{1, \ldots, N^{(n)}(t)\}: Y_{j-1}^{(n)} = 1, Y_j^{(n)} = 0\big\}$$

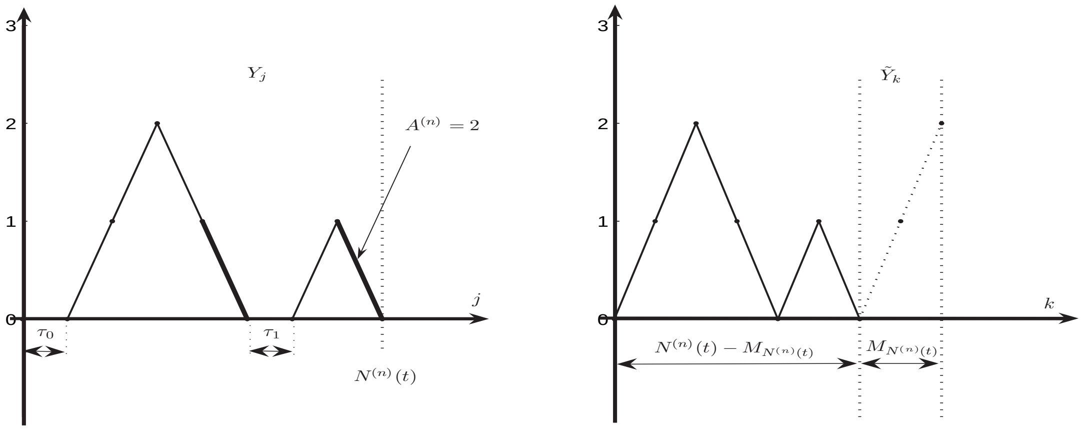

FIGURE 4. On the left a sample of the processes $\{Y_j : 0 \le j \le N^{(n)}(t)\}$. On the right the associated $\{\tilde{Y}_k : 0 \le k \le N^{(n)}(t)\}$, which is obtained by removing the two flat steps and extending the path to its original length.

be the total number of returns to zero before time N. Then, almost surely, as $n \uparrow \infty$,

$$2^{-n}\Big(\#\,\big\{j \in \{1,\dots,N^{(n)}(t)\} : Y_j^{(n)} = Y_{j-1}^{(n)}\big\} - \#\big\{j \in \{1,\dots,N^{(n)}(t)\}\ :\ Y_{j-1}^{(n)} = 0, Y_j^{(n)} = 1\big\}\Big)$$

$$= 2^{-n}\sum_{i=0}^{A^{(n)}}(Z_i - 1) = \big(2^{-n}A^{(n)}\big)\,\frac{1}{A^{(n)}}\sum_{i=0}^{A^{(n)}}\big(Z_i - \mathbb{E}Z_i\big) \longrightarrow 0,$$

because the first factor converges by Lemma 6.12 and the second by the law of large numbers.

To study the effect of the removal of the flat pieces, recall that almost surely the length $N^{(n)}(t)$ of the walk is of order $2^{2n}t$, by Lemma 6.11, and the number of flat pieces is $M_{N^{(n)}(t)}$, which is of order 2^n, by Lemma 6.13. Hence, for all $\varepsilon > 0$, if n is large enough,

$$N^{(n)}(t-\varepsilon) + M_{N^{(n)}(t)} \le N^{(n)}(t).$$

We infer from this that

$$2^{-n}\Big(\#\big\{j \in \{1,\dots,N^{(n)}(t)\}\ :\ \tilde{Y}_{j-1}^{(n)} = 0, \tilde{Y}_j^{(n)} = 1\big\}$$
$$- \#\big\{j \in \{1,\dots,N^{(n)}(t)\}\ :\ Y_{j-1}^{(n)} = 0, Y_j^{(n)} = 1\big\}\Big)$$
$$\le 2^{-n}\#\big\{j \in \{N^{(n)}(t-\varepsilon)+1,\dots,N^{(n)}(t)\}\ :\ \tilde{Y}_{j-1}^{(n)} = 0, \tilde{Y}_j^{(n)} = 1\big\},$$

and the right hand side converges almost surely to a random variable, which has the law of $L(t) - L(t-\varepsilon)$ and hence can be made arbitrarily small by choice of $\varepsilon > 0$. ∎

Proof of Theorem 6.10 Note that both processes in Theorem 6.10 are continuous, so that it suffices to compare their finite dimensional distributions. Equality of these follows directly by combining Lemma 6.13, Equation (1.2) and Lemma 6.14. ∎

THEOREM 6.15 (Occupation time representation of the local time at zero). *For all sequences $a_n \uparrow 0$ and $b_n \downarrow 0$ with $a_n < b_n$, almost surely,*

$$\lim_{n\to\infty} \frac{1}{b_n - a_n} \int_0^t \mathbb{1}\{a_n \le B(s) \le b_n\}\, ds = L(t) \text{ for every } t > 0.$$

The proof is prepared by the following lemma, which we prove as Exercise 6.4.

LEMMA 6.16. *Let $\{W(s) \colon s \ge 0\}$ be a standard linear Brownian motion and τ_1 its first hitting time of level 1. Then $\mathbb{E} \int_0^{\tau_1} \mathbb{1}\{0 \le W(s) \le 1\}\, ds = 1$.*

Proof of Theorem 6.15 Recall the stopping times τ_j defined for $a_n < b_n$ as in (1.1). For the proof of the lower bound note that

$$\int_0^t \mathbb{1}\{a_n \le B(s) \le b_n\}\, ds \ge \sum_{j=1}^{D(a_n,b_n,t)} \int_{\tau_{j-1}}^{\tau_j} \mathbb{1}\{a_n \le B(s) \le b_n\}\, ds\,.$$

By Brownian scaling

$$\int_{\tau_{j-1}}^{\tau_j} \mathbb{1}\{a_n \le B(s) \le b_n\}\, ds = (b_n - a_n)^2 \int_0^\tau \mathbb{1}\{0 \le W(s) \le 1\}\, ds\,,$$

where $\{W(s) \colon s \ge 0\}$ is a standard linear Brownian motion and

$$\tau = \inf\big\{s > 0 \colon W(s) = 0 \text{ and there exists } t < s \text{ with } W(t) = 1\big\}.$$

Hence

$$\frac{1}{b_n - a_n} \sum_{j=1}^{D(a_n,b_n,t)} \int_{\tau_{j-1}}^{\tau_j} \mathbb{1}\{a_n \le B(s) \le b_n\}\, ds$$
$$= (b_n - a_n) D(a_n, b_n, t) \Big[\frac{1}{D(a_n, b_n, t)} \sum_{j=1}^{D(a_n,b_n,t)} \int_0^\tau \mathbb{1}\{0 \le W(s) \le 1\}\, ds\Big]\,.$$

The first factor converges almost surely to $\frac{1}{2} L(t)$, by Theorem 6.1. From the law of large numbers and Lemma 6.16 we get for the second factor,

$$\lim_{n\uparrow\infty} \frac{1}{D(a_n, b_n, t)} \sum_{j=1}^{D(a_n,b_n,t)} \int_0^\tau \mathbb{1}\{0 \le W(s) \le 1\}\, ds = \mathbb{E} \int_0^\tau \mathbb{1}\{0 \le W(s) \le 1\}\, ds = 2\,.$$

This verifies the lower bound. The upper bound can be obtained by including the period $[\tau_j, \tau_{j+1}]$ for $j = D(a_n, b_n, t)$ in the summation and using the same arguments as for the lower bound. This completes the proof of Theorem 6.15. ■

2. A random walk approach to the local time process

Given a level $a \in \mathbb{R}$ the construction of the previous chapter allows us to define the *local time at level* a for a linear Brownian motion $\{B(t)\colon t \geq 0\}$. Indeed, simply let $\{L^a(t)\colon t \geq 0\}$ be the local time at zero of the auxiliary Brownian motion $\{B^a(t)\colon t \geq 0\}$ defined by $B^a(t) = B(t) - a$. Using Theorem 6.15 it is not hard to show that $\{L^a(t)\colon a \in \mathbb{R}\}$ is the density of the occupation measure μ_t introduced in Theorem 3.25.

THEOREM 6.17. *For linear Brownian motion $\{B(t)\colon t \geq 0\}$, almost surely, for any bounded measurable $g\colon \mathbb{R} \to \mathbb{R}$ and $t > 0$,*

$$\int g(a)\, d\mu_t(a) = \int_0^t g(B(s))\, ds = \int_{-\infty}^{\infty} g(a)\, L^a(t)\, da.$$

Proof. First, observe that for the statement it suffices to have $\{L^a(t)\colon t \geq 0\}$ defined for $\mathcal{L}$-almost every a. Second, we may assume that t is fixed. Indeed, it suffices to verify the second equality for a countable family of bounded measurable $g\colon \mathbb{R} \to \mathbb{R}$, for example the indicator functions of rational intervals. Having fixed such a g both sides are continuous in t.

For fixed t, we know from Theorem 3.25 that $\mu_t \ll \mathcal{L}$ almost surely, hence a density f exists by the Radon-Nikodym theorem and may be obtained as

$$f(a) = \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \int_0^t \mathbb{1}\{a - \varepsilon \leq B(s) \leq a + \varepsilon\}\, ds,$$

which equals $L^a(t)$ by Theorem 6.15, almost surely for $\mathcal{L}$-almost every a. ∎

A major result about linear Brownian motion is the continuity of the density $\{L^a(t)\colon a \in \mathbb{R}\}$ of the occupation measures, which we now prove. To explore $L^a(t)$ as a function of the levels a we extend the downcrossing representation to hold *simultaneously* at all levels a. We approach this problem via the random walks embedded in a Brownian motion.

For a Brownian motion $\{B(t)\colon t \geq 0\}$ started in the origin we recall the definition of the embedded random walks $\{X_k^{(n)}\colon k = 1, 2, \ldots\}$ from the previous section: Define stopping times $\tau_0 := \tau_0^{(n)} := 0$ and

$$\tau_k := \tau_k^{(n)} := \inf\big\{t > \tau_{k-1}\colon |B(t) - B(\tau_{k-1})| = 2^{-n}\big\}, \qquad \text{for } k \geq 1,$$

and define the nth embedded random walk by

$$X_k := X_k^{(n)} := 2^n\, B\big(\tau_k^{(n)}\big)\,.$$

For any time $t > 0$ the length of the embedded random walk is

$$N := N^{(n)}(t) := \max\{k \in \mathbb{N}\colon \tau_k \leq t\}\,,$$

which is independent of the nth walk itself. Further, given $a \in \mathbb{R}$, choose $j(a) \in \{0, 1, \ldots\}$ such that $j(a)\, 2^{-n} \leq a < (j(a) + 1)\, 2^{-n}$. Denote the number of downcrossings of $2^n a$ by the nth embedded random walk by

$$D^{(n)}(a, t) := \#\big\{k \in \{0, \ldots, N^{(n)}(t)\}\colon X_k^{(n)} = j(a) + 1,\ X_{k+1}^{(n)} = j(a)\big\}\,.$$

Theorem 6.18 (Trotter's theorem). *Let $\{B(t)\colon t \geq 0\}$ be a linear Brownian motion and let $D^{(n)}(a,t)$ be the number of downcrossings of $2^n a$ by the nth embedded random walk stopped at time $N^{(n)}(t)$. Then, almost surely,*

$$L^a(t) := \lim_{n\to\infty} 2^{-n+1}\, D^{(n)}(a,t) \qquad \text{exists for all } a \in \mathbb{R} \text{ and } t \geq 0.$$

Moreover, for every $\gamma < \frac{1}{2}$, the random field

$$\{L^a(t)\colon a \in \mathbb{R}, t \geq 0\}$$

is almost surely locally γ-Hölder continuous.

Remark 6.19. Note that $\{L^a(t) : a \in \mathbb{R},\, t \geq 0\}$ is a stochastic process depending on more than one parameter, and to emphasise this fact we use the notion **random field**. $\diamond$

The proof uses the following estimate for the sum of independent geometric random variables with mean two, which we prove as Exercise 6.5.

Lemma 6.20. *Let $X_1, X_2, \dots$ are independent geometrically distributed random variables with mean 2. Then, for sufficiently small $\varepsilon > 0$, for all nonnegative integers $k \leq m$,*

$$\mathbb{P}\Big\{\Big|\sum_{j=1}^{k}(X_j - 2)\Big| \geq \varepsilon m\Big\} \leq 4\, \exp\big\{-\tfrac{1}{5}\,\varepsilon^2\, m\big\}\,.$$

The following lemma is the heart of the proof of Theorem 6.18.

Lemma 6.21. *Suppose that $a < b$ and let $\{B(t)\colon 0 \leq t \leq T\}$ be a linear Brownian motion stopped at the time T when it first hits a given level above b. Let*

- *D be the number of downcrossings of the interval $[a,b]$,*
- *D_{l} be the number of downcrossings of the interval $\big[a, \frac{a+b}{2}\big]$,*
- *D_{u} be the number of downcrossings of the interval $\big[\frac{a+b}{2}, b\big]$.*

Then, for sufficiently small $\varepsilon > 0$, for all nonnegative integers $k \leq m$,

$$\mathbb{P}\big\{\big|D - \tfrac{1}{2}\, D_{\mathrm{l}}\big| > \varepsilon m,\ \big|D - \tfrac{1}{2}\, D_{\mathrm{u}}\big| > \varepsilon m \,\big|\, D = k\big\} \leq 12\, \exp\big\{-\tfrac{1}{5}\,\varepsilon^2\, m\big\}\,.$$

Proof. By Lemma 6.3 we have that, given $\{D = k\}$, there exist independent random variables $X_0, X_1, X_2 \dots$, such that

$$D_{\mathrm{l}} = X_0 + \sum_{j=1}^{k} X_j\,.$$

and $X_1, X_2, \dots$ are geometrically distributed with mean 2. An inspection of the proof of Theorem 6.3 reveals that X_0 is either zero or also geometrically distributed with mean 2, depending on the starting point of the Brownian motion.

Using Lemma 6.20 and Chebyshev's inequality, we get, if $\varepsilon > 0$ is small enough,

$$\mathbb{P}\big\{\big|\tfrac{1}{2}\,D_{\rm l} - D\big| > \varepsilon m \,\big|\, D = k\big\} \le \mathbb{P}\Big\{\big|\sum_{j=1}^{k}(X_j - 2)\big| > \varepsilon m \,\Big|\, D = k\Big\} + \mathbb{P}\big\{X_0 > \varepsilon m\big\}$$
$$\le 4\exp\big\{-\tfrac{\varepsilon^2}{5}\,m\big\} + 2\,\exp\{-\varepsilon m \log 2\} \le 6\exp\big\{-\tfrac{\varepsilon^2}{5}\,m\big\}\,.$$

The argument is analogous for $D_{\rm u}$, and this completes the proof. ∎

We now fix $\gamma < \frac{1}{2}$ and a large integer N. We stop the Brownian motion at time T_N when it first hits level N, and abbreviate $D^{(n)}(a) := D^{(n)}(a, T_N)$. We denote the nth dyadic grid by

$$\mathcal{D}_n := \mathcal{D}_n(N) := \big\{k2^{-n} \colon k \in \{-N2^n, -N2^n + 1, \dots, N2^n - 1\}\big\}.$$

Lemma 6.22. *Denote by $\Omega(m)$ the event that, for all $n \ge m$,*

(a) $\big|D^{(n)}(a) - \frac{1}{2}\,D^{(n+1)}(a)\big| \le 2^{n(1-\gamma)}$ *for all $a \in [-N, N)$,*

(b) $\big|D^{(n)}(a) - D^{(n)}(b)\big| \le 2\,2^{n(1-\gamma)}$ *for all $a, b \in [-N, N)$ with $|a - b| \le 2^{-n}$.*

Then

$$\lim_{m\uparrow\infty} \mathbb{P}\big(\Omega(m)\big) = 1\,.$$

Proof. Item (a) follows by combining the following three items,

(i) $\big|D^{(n)}(a) - \frac{1}{2}\,D^{(n+1)}(a)\big| \le \frac{1}{n^2}\,2^{-n\gamma}\,D^{(n)}(a)$ for all $a \in [-N, N)$ with $D^{(n)}(a) \ge 2^n$,

(ii) $\big|D^{(n)}(a) - \frac{1}{2}\,D^{(n+1)}(a)\big| \le 2^{n(1-\gamma)}$ for all $a \in [-N, N)$ with $D^{(n)}(a) < 2^n$,

(iii) $D^{(n)}(a) \le n^2 2^n$ for all $a \in [-N, N)$.

We observe that it is equivalent to show (i),(ii) for all $a \in \mathcal{D}_{n+1}$ and (iii) for all $a \in \mathcal{D}_n$.

To estimate the probability of the first item we use Lemma 6.21 with $\varepsilon = \frac{1}{n^2}\,2^{-n\gamma}$ and $m = k$. We get that

$$\sum_{n=m}^{\infty}\sum_{a\in\mathcal{D}_{n+1}} \mathbb{P}\big\{\big|D^{(n)}(a) - \tfrac{1}{2}\,D^{(n+1)}(a)\big| > \tfrac{1}{n^2}\,2^{-n\gamma}\,D^{(n)}(a) \text{ and } D^{(n)}(a) \ge 2^n\big\}$$
$$\le \sum_{n=m}^{\infty}\sum_{a\in\mathcal{D}_{n+1}} 12\,\exp\big\{-\tfrac{1}{5n^4}\,2^{n(1-2\gamma)}\big\}$$
$$\le (48N)\sum_{n=m}^{\infty} 2^n\,\exp\big\{-\tfrac{1}{5n^4}\,2^{n(1-2\gamma)}\big\} \stackrel{m\to\infty}{\longrightarrow} 0\,.$$

For the second item we get from Lemma 6.21 with $\varepsilon = 2^{-\gamma n}$ and $m = 2^n > k$. This gives that

$$\sum_{n=m}^{\infty} \sum_{a\in\mathcal{D}_{n+1}} \mathbb{P}\big\{\big|D^{(n)}(a) - \tfrac{1}{2}\,D^{(n+1)}(a)\big| > 2^{n(1-\gamma)} \text{ and } D^{(n)}(a) < 2^n\big\}$$

$$\leq \sum_{n=m}^{\infty} \sum_{a\in\mathcal{D}_{n+1}} 12\,\exp\big\{-\tfrac{1}{5}\,2^{n(1-2\gamma)}\big\} \leq (48N) \sum_{n=m}^{\infty} 2^n\,\exp\big\{-\tfrac{1}{5}\,2^{n(1-2\gamma)}\big\} \stackrel{m\to\infty}{\longrightarrow} 0\,.$$

For the third item we use that the random variable $D^{(n)}(a)$ is geometrically distributed with parameter $\frac{2^{-n}}{N-a} \geq \frac{2^{-n}}{2N}$. We therefore obtain, for some sequence $\delta_n \to 0$,

$$\mathbb{P}\big\{D^{(n)}(a) > n^2 2^n\big\} \leq \big(1 - \tfrac{2^{-n}}{2N}\big)^{n^2 2^n - 1} \leq \exp\{-n^2 \tfrac{1+\delta_n}{2N}\},$$

hence, for sufficiently large m,

$$\sum_{n=m}^{\infty} \sum_{a\in\mathcal{D}_n} \mathbb{P}\big\{D^{(n)}(a) > n^2 2^n\big\} \leq \sum_{n=m}^{\infty} (2N) 2^n \exp\big\{-n^2 \tfrac{1+\delta_n}{2N}\big\} \stackrel{m\to\infty}{\longrightarrow} 0.$$

This completes the estimates needed for item (a).

Item (b) need only be checked for all $a, b \in \mathcal{D}_n$ with $|a-b| = 2^{-n}$. Note that $D^{(n)}(a)$, resp. $D^{(n)}(b)$, are the number of downcrossings of the lower, resp. upper, half of an interval of length 2^{-n+1}, which may or may not be dyadic. Denote by $\tilde{D}^{(n-1)}(a) = \tilde{D}^{(n-1)}(b)$ the number of downcrossings of this interval. Then

$$\mathbb{P}\{\big|D^{(n)}(a) - D^{(n)}(b)\big| > 2\,2^{n(1-\gamma)}\}$$

$$\leq \mathbb{P}\{\big|D^{(n)}(a) - \tfrac{1}{2}\,\tilde{D}^{(n-1)}(a)\big| > 2^{n(1-\gamma)}\} + \mathbb{P}\{\big|D^{(n)}(b) - \tfrac{1}{2}\,\tilde{D}^{(n-1)}(b)\big| > 2^{n(1-\gamma)}\},$$

and summability of these probabilities over all $a, b \in \mathcal{D}_n$ with $|a-b| = 2^{-n}$ and $n \geq m$ has been established in the proof of item (a). This completes the proof. ∎

Lemma 6.23. *On the set $\Omega(m)$ we have that*

$$L^a(T_N) := \lim_{n\to\infty} 2^{-n-1}\,D^{(n)}(a)$$

exists for every $a \in [-N, N)$.

Proof. We show that the sequence defined by $2^{-n-1}\,D^{(n)}(a)$, for $n \in \mathbb{N}$, is a Cauchy sequence. Indeed, by item (a) in the definition of the set $\Omega(m)$ we get that, for any $a \in [-N, N]$ and $n \geq m$,

$$\big|2^{-n-1} D^{(n)}(a) - 2^{-n-2}\,D^{(n+1)}(a)\big| \leq 2^{-n\gamma}\,.$$

Thus, for any $n \geq m$,

$$\sup_{k\geq n} \big|2^{-n-1} D^{(n)}(a) - 2^{-k-1}\,D^{(k)}(a)\big|$$

$$\leq \sum_{k=n}^{\infty} \big|2^{-k-1} D^{(k)}(a) - 2^{-k-2}\,D^{(k+1)}(a)\big| \leq \sum_{k=n}^{\infty} 2^{-k\gamma} \stackrel{n\to\infty}{\longrightarrow} 0,$$

and thus the sequence is a Cauchy sequence and therefore convergent. ∎

LEMMA 6.24. *On the set $\Omega(m)$ the process $\{L^a(T_N)\colon a \in [-N,N)\}$ is γ-Hölder continuous.*

Proof. Fix $a, b \in [-N,N)$ with $2^{-n-1} \le a-b \le 2^{-n}$ for some $n \ge m$. Then, using item (a) and item (b) in the definition of $\Omega(m)$, for all $k \ge n$,

$$\begin{aligned}
\big|2^{-k-1}D^{(k)}(a) - 2^{-k-1}D^{(k)}(b)\big| &\le \big|2^{-n-1}D^{(n)}(a) - 2^{-n-1}D^{(n)}(b)\big| \\
&\quad + \sum_{j=n}^{k-1}\big|2^{-j-2}D^{(j+1)}(a) - 2^{-j-1}D^{(j)}(a)\big| + \sum_{j=n}^{k-1}\big|2^{-j-2}D^{(j+1)}(b) - 2^{-j-1}D^{(j)}(b)\big| \\
&\le 2\,2^{-n\gamma} + 2\sum_{j=n}^{\infty} 2^{-j\gamma},
\end{aligned}$$

Letting $k \uparrow \infty$, we get

$$|L^a(T_N) - L^b(T_N)| \le \big(2 + \tfrac{2}{1-2^{-\gamma}}\big)\,2^{-n\gamma} \le \big(2^{1+\gamma} + \tfrac{2^{1+\gamma}}{1-2^{-\gamma}}\big)\,|a-b|^\gamma,$$

which completes the proof. ∎

LEMMA 6.25. *For any fixed time $t > 0$, almost surely, the limit*

$$L^a(t) := \lim_{n\to\infty} 2^{-n-1}\,D^{(n)}(a) \qquad \textit{exists for all } a \in \mathbb{R}$$

and moreover $\{L^a(t)\colon a \in \mathbb{R}\}$ is γ-Hölder continuous.

Proof. Given $t > 0$ define the auxiliary Brownian motion $\{B_t(s)\colon s \ge 0\}$ by $B_t(s) = B(t+s)$ and denote by $D_t^{(n)}(a)$ the number of downcrossings associated to the auxiliary Brownian motion. Then, almost surely, $L_t^a(T_N) := \lim_{n\uparrow\infty} 2^{-n-1}\,D_t^{(n)}(a)$ exists for all $a \in \mathbb{R}$ and integers N. On this event we pick N so large that $T_N > t$. Define $L^a(t) := L^a(T_N) - L_t^a(T_N)$, and observe that $\{L^a(t)\colon a \in \mathbb{R}\}$ defined like this is γ-Hölder continuous by Lemma 6.24. It remains to show that this definition agrees with the one stated in the lemma. To this end, observe that

$$D^{(n)}(a,T_N) - D_t^{(n)}(a,T_N) - 1 \le D^{(n)}(a,t) \le D^{(n)}(a,T_N) - D_t^{(n)}(a,T_N).$$

Multiplying by 2^{-n-1} and taking a limit proves the claimed convergence. ∎

LEMMA 6.26. *Almost surely,*

$$L^a(t) := \lim_{n\to\infty} 2^{-n-1}\,D^{(n)}(a,t)$$

exists for every $t \ge 0$ and $a \in \mathbb{R}$ and $\{L^a(t)\colon a \in \mathbb{R}, t \ge 0\}$ is γ-Hölder continuous.

Proof. It suffices to look at $t \in [0,N)$ and $a \in [-N,N)$. Recall the definition of the dyadic points $\mathcal{D}_n$ in $[-N,N)$ and additionally define dyadic points in $[0,N)$ by

$$\mathcal{H}_m = \big\{k2^{-m}\colon k \in \{0,\dots,N2^m - 1\}\big\}, \quad \mathcal{H} = \bigcup_{m=1}^{\infty} \mathcal{H}_m\,.$$

We show that the claimed statements hold on the set

$$\bigcap_{t\in\mathcal{H}} \big\{ L^a(t) \text{ exists for all } a \in [-N,N) \text{ and } a \mapsto L^a(t) \text{ is } \gamma\text{-Hölder continuous } \big\}$$
$$\cap \bigcap_{m>M} \bigcap_{t\in\mathcal{H}_m} \bigcap_{a\in\mathcal{D}_m} \big\{ L^a(t+2^{-m}) - L^a(t) \le 2^{-m\gamma} \big\},$$

which, by choosing M suitably, has probability arbitrarily close to one by Lemma 6.25 and Lemma 6.7.

Given any $t \in [0,N)$ and $a \in [-N,N]$, for any large m, we find $t_1, t_2 \in \mathcal{H}_m$ with $t_2 - t_1 = 2^{-m}$ and $t \in [t_1, t_2]$. We have

$$2^{-n-1}\, D^{(n)}(a,t_1) \le 2^{-n-1}\, D(a,t) \le 2^{-n-1}\, D(a,t_2).$$

Both bounds converge on our set, and the difference of the limits is $L^a(t_2) - L^a(t_1)$. We can then find $b \in \mathcal{H}_k$ for $k \ge M$ with $|L^a(t_1) - L^b(t_1)| < 2^{-m\gamma}$ and $|L^a(t_2) - L^b(t_2)| < 2^{-m\gamma}$ and get

$$0 \le L^a(t_2) - L^a(t_1) \le |L^a(t_2) - L^b(t_2)| + |L^b(t_2) - L^b(t_1)| + |L^a(t_1) - L^b(t_1)| \le 3 2^{-m\gamma},$$

which can be made arbitrarily small by choice of m, proving simultaneous convergence.

For the proof of continuity, suppose $a, b \in [-N,N)$ and $s, t \in [0,N)$ with $2^{-m} \le |a-b| \le 2^{-m}$ and $2^{-m} \le t - s \le 2^{-m}$ for some $m \ge M$. We pick $s_1, s_2 \in \mathcal{H}_m$ and $t_1, t_2 \in \mathcal{H}_m$ such that

$$s - 2^{-m} < s_1 \le s \le s_2 < s + 2^{-m} \qquad \text{and} \qquad t - 2^{-m} < t_1 \le t \le t_2 < t + 2^{-m},$$

and $a_1, b_1 \in \mathcal{D}_m$ with $|a - a_1| \le 2^{-m}$ and $|b - b_1| \le 2^{-m}$. Then

$$\begin{aligned} L^a(t) - L^b(s) &\le L^a(t_2) - L^b(s_1) \\ &\le |L^a(t_2) - L^{a_1}(t_2)| + |L^{a_1}(t_2) - L^{a_1}(s_1)| + |L^{a_1}(s_1) - L^b(s_1)|, \\ L^a(s) - L^b(t) &\le L^a(s_2) - L^b(t_1) \\ &\le |L^a(s_2) - L^{a_1}(s_2)| + |L^{a_1}(s_2) - L^{a_1}(t_1)| + |L^{a_1}(t_1) - L^b(t_1)|, \end{aligned}$$

and all contributions on the right are bounded by constant multiples of $2^{-m\gamma}$, by the construction of our set. This completes the proof of γ-Hölder continuity. ■

This completes the proof of Trotter's theorem, Theorem 6.18.

3. The Ray-Knight theorem

We now have a closer look at the distributions of local times $L^x(T)$ as a function of the level x in the case that Brownian motion is started at an arbitrary point and stopped at the time T when it first hits level zero. The following remarkable distributional identity goes back to the work of Ray and Knight.

Theorem 6.27 (Ray-Knight theorem). *Suppose $a > 0$ and $\{B(t) : 0 \le t \le T\}$ is a linear Brownian motion started at a and stopped at time $T = \inf\{t \ge 0 : B(t) = 0\}$, when it reaches level zero for the first time. Then*

$$\{L^x(T) : 0 \le x \le a\} \overset{\mathrm{d}}{=} \{|W(x)|^2 : 0 \le x \le a\},$$

where $\{W(x) \colon x \ge 0\}$ is a standard planar Brownian motion.

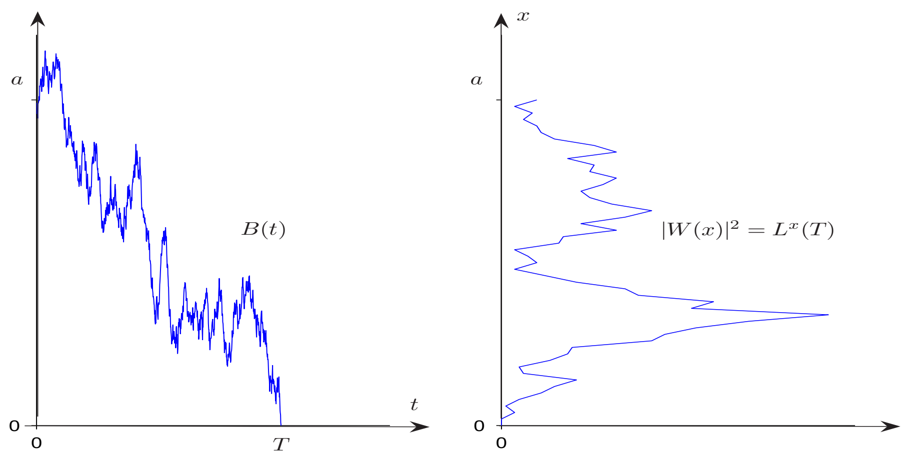

Figure 5. The Brownian path on the left, and its local time as a function of the level, on the right.

Remark 6.28. The process $\{|W(x)|^2 : x \ge 0\}$ of squared norms of a planar Brownian motion is called the squared *two-dimensional Bessel process*. For any fixed x, the random variable $|W(x)|^2$ is exponentially distributed with mean $2x$, see Lemma II.3.8. $\diamond$

We carry out the proof of the Ray-Knight theorem in three steps. As a warm-up, we look at one point $0 < x \le a$. Recall from the downcrossing representation, Theorem 6.1, that

$$\lim_{n\to\infty} \tfrac{2}{n} D_n(x) = L^x(T) \qquad \text{almost surely,}$$

where $D_n(x)$ denotes the number of downcrossings of the interval $[x, x - 1/n]$ before time T.

Lemma 6.29. *For any $0 < x \le a$, we have $\frac{2}{n} D_n(x) \Longrightarrow |W(x)|^2$ as $n \uparrow \infty$.*

Proof. By the strong Markov property and the exit probabilities from an interval described in Theorem 2.45, it is clear that, provided $n > 1/x$, the random variable $D_n(x)$ is geometrically distributed with (success) parameter $1/(nx)$, i.e. $\mathbb{P}\{D_n(x) = k\} = \frac{1}{nx}(1 - \frac{1}{nx})^{k-1}$ for all $k \in \{1, 2, \ldots\}$. Hence, as $n \to \infty$, we obtain that

$$\mathbb{P}\{D_n(x) > ny/2\} = \big(1 - \tfrac{1}{nx}\big)^{\lfloor ny/2 \rfloor} \longrightarrow e^{-y/(2x)},$$

and the result follows, as $|W(x)|^2$ is exponentially distributed with mean $2x$. $\blacksquare$

Lemma 6.29 is the 'one-point version' of Theorem 6.27. The essence of the Ray-Knight theorem is captured in the 'two-point version', which we prove next. We fix two points x and $x+h$ with $0 < x < x+h < a$. The next three lemmas are the crucial ingredients for the proof of Theorem 6.27.

LEMMA 6.30. *Let $0 < x < x+h < a$. Then, for all $n > h$, we have*

$$D_n(x+h) = D + \sum_{j=1}^{D_n(x)} I_j\, N_j\,,$$

where

- *$D = D^{(n)}$ is the number of downcrossings of the interval $[x+h-\frac{1}{n}, x+h]$ before the Brownian motion hits level x,*
- *for any $j \in \mathbb{N}$ the random variable $I_j = I_j^{(n)}$ is Bernoulli distributed with mean $\frac{1}{nh+1}$,*
- *for any $j \in \mathbb{N}$ the random variable $N_j = N_j^{(n)}$ is geometrically distributed with mean $nh+1$,*

and all these random variables are independent of each other and of $D_n(x)$.

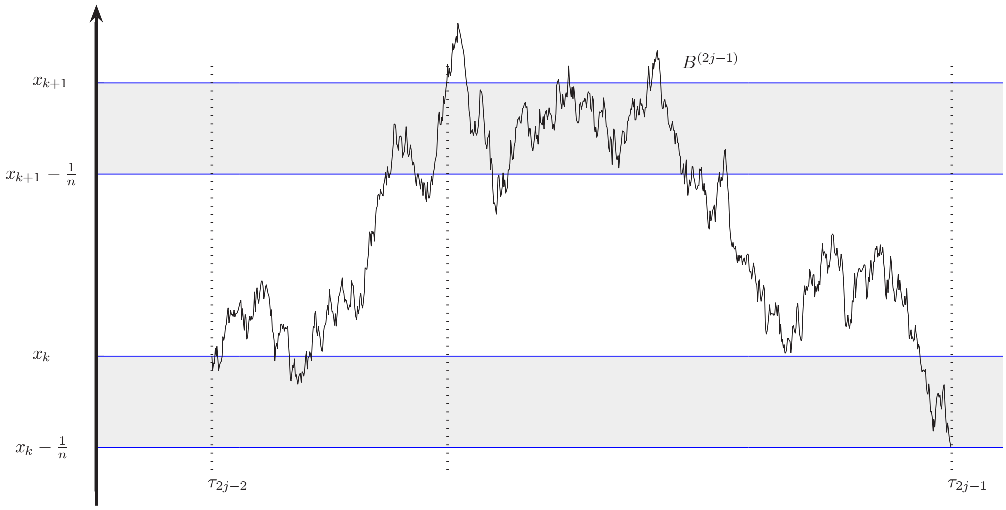

FIGURE 6. The random variables I_j and N_j depend only on the pieces $B^{(2j-1)}$ for $j \geq 1$. For this sample $I_j = 1$ as the path hits $x+h$ before $x - \frac{1}{n}$ and $N_j = 2$, because the path downcrosses $[x+h, x+h-\frac{1}{n}]$ twice before hitting $x - \frac{1}{n}$.

Proof. The decomposition of $D_n(x+h)$ is based on counting the number of downcrossings of the interval $[x+h-1/n, x+h]$ that have taken place between the stopping times in the sequence

$$\tau_0 = \inf\big\{t > 0\colon B(t) = x\big\}, \qquad \tau_1 = \inf\big\{t > \tau_0\colon B(t) = x - \tfrac{1}{n}\big\},$$
$$\tau_{2j} = \inf\big\{t > \tau_{2j-1}\colon B(t) = x\big\}, \qquad \tau_{2j+1} = \inf\big\{t > \tau_{2j}\colon B(t) = x - \tfrac{1}{n}\big\},$$

for $j \geq 1$. By the strong Markov property the pieces

$$B^{(0)} \colon [0, \tau_0] \to \mathbb{R}, \qquad B^{(0)}(s) = B(s)$$
$$B^{(j)} \colon [0, \tau_j - \tau_{j-1}] \to \mathbb{R}, \qquad B^{(j)}(s) = B(\tau_{j-1} + s),\ j \geq 1,$$

are all independent. The crucial observation of the proof is that the vector $D_n(x)$ is a function of the pieces $B^{(2j)}$ for $j \geq 1$, whereas we shall define the random variables D, $I_1, I_2, \ldots$ and $N_1, N_2 \ldots$ depending only on the other pieces $B^{(0)}$ and $B^{(2j-1)}$ for $j \geq 1$.

First, let D be the number of downcrossings of $[x+h, x+h-1/n]$ during the time interval $[0, \tau_0]$. Then fix $j \geq 1$ and hence a piece $B^{(2j-1)}$. Define I_j to be the indicator of the event that $B^{(2j-1)}$ reaches level $x+h$ during its lifetime. By Theorem 2.45 this event has probability $1/(nh+1)$. Observe that the number of downcrossings by $B^{(2j-1)}$ is zero if the event fails. If the event holds, we define N_j as the number of downcrossings of $[x+h, x+h-1/n]$ by $B^{(2j-1)}$, which is a geometric random variable with mean $nh+1$ by the strong Markov property and Theorem 2.45.

The claimed decomposition follows now from the fact that the pieces $B^{(2j)}$ for $j \geq 1$ do not upcross the interval $[x+h, x+h-1/n]$ by definition and that $B^{(2j-1)}$ for $j = 1, \ldots, D_n(x)$ are exactly the pieces that take place before the Brownian motion reaches level zero. ■

Lemma 6.31. *Suppose nu_n are nonnegative, even integers and $u_n \to u$. Then*

$$\frac{2}{n}\, D^{(n)} + \frac{2}{n} \sum_{j=1}^{\frac{nu_n}{2}} I_j^{(n)}\, N_j^{(n)} \Longrightarrow \tilde{X}^2 + \tilde{Y}^2 + 2 \sum_{j=1}^{M} \tilde{Z}_j \qquad \text{as } n \uparrow \infty,$$

where $\tilde{X}$, $\tilde{Y}$ are normally distributed with mean zero and variance h, the random variable M is Poisson distributed with parameter $u/(2h)$ and $\tilde{Z}_1, \tilde{Z}_2, \ldots$ are exponentially distributed with mean h, and all these random variables are independent.

Proof. By Lemma 6.29, we have, for $\tilde{X}$, $\tilde{Y}$ as defined in the lemma,

$$\frac{2}{n}\, D^{(n)} \Longrightarrow |W(h)|^2 \stackrel{d}{=} \tilde{X}^2 + \tilde{Y}^2 \qquad \text{as } n \uparrow \infty.$$

Moreover, we observe that

$$\frac{2}{n} \sum_{j=1}^{\frac{nu_n}{2}} I_j^{(n)}\, N_j^{(n)} \stackrel{d}{=} \frac{2}{n} \sum_{j=1}^{B_n} N_j^{(n)},$$

where B_n is binomial with parameters $nu_n/2 \in \{0, 1, \ldots\}$ and $1/(nh+1) \in (0,1)$ and independent of $N_1^{(n)}, N_2^{(n)}, \ldots$. We now show that, when $n \uparrow \infty$, the random variables B_n converge in distribution to M and the random variables $\frac{1}{n}\, N_j^{(n)}$ converge to $\tilde{Z}_j$, as defined in the lemma. For this purpose it suffices to show convergence of the Laplace transforms, see Proposition II.1.8.

First note that, for $\lambda, \theta > 0$, we have

$$\mathbb{E} \exp\big\{ -\lambda \tilde{Z}_j \big\} = \tfrac{1}{\lambda h + 1}, \quad \mathbb{E}\big[\theta^M\big] = \exp\big\{ -\tfrac{u(1-\theta)}{2h} \big\},$$

and hence

$$\mathbb{E}\exp\Big\{-\lambda\sum_{j=1}^{M}\tilde{Z}_j\Big\} = \mathbb{E}\big(\tfrac{1}{\lambda h+1}\big)^{M} = \exp\big\{-\tfrac{u}{2h}\,\tfrac{\lambda h}{\lambda h+1}\big\} = \exp\big\{-\tfrac{u\lambda}{2\lambda h+2}\big\}\,.$$

Convergence of $\frac{1}{n}\,N_j^{(n)}$ is best seen using tail probabilities

$$\mathbb{P}\big\{\tfrac{1}{n}\,N_j^{(n)} > a\big\} = \big(1-\tfrac{1}{nh+1}\big)^{\lfloor na\rfloor} \longrightarrow \exp\big\{-\tfrac{a}{h}\big\} = \mathbb{P}\{\tilde{Z}_j > a\}\,.$$

Hence, for a suitable sequence $\delta_n \to 0$,

$$\mathbb{E}\exp\big\{-\lambda\tfrac{1}{n}\,N_j^{(n)}\big\} = \frac{1+\delta_n}{\lambda h+1}\,.$$

For the binomial distributions we have

$$\mathbb{E}\big[\theta^{B_n}\big] = \Big(\tfrac{\theta}{nh+1} + \big(1-\tfrac{1}{nh+1}\big)\Big)^{nu_n/2} \longrightarrow \exp\big\{-\tfrac{u(1-\theta)}{2h}\big\}\,,$$

and thus

$$\begin{aligned}\lim_{n\uparrow\infty}\mathbb{E}\exp\Big\{-\lambda\frac{1}{n}\sum_{j=1}^{B_n}N_j^{(n)}\Big\} &= \lim_{n\uparrow\infty}\mathbb{E}\Big[\big(\tfrac{1+\delta_n}{\lambda h+1}\big)^{B_n}\Big] = \exp\big\{-\tfrac{u}{2h}\,\tfrac{\lambda h+\delta_n}{\lambda h+1}\big\}\\ &= \exp\big\{-\tfrac{u\lambda}{2\lambda h+2}\big\} = \mathbb{E}\exp\Big\{-\lambda\sum_{j=1}^{M}\tilde{Z}_j\Big\}.\end{aligned}$$

■

LEMMA 6.32. *Suppose X is standard normally distributed, $Z_1, Z_2, \ldots$ standard exponentially distributed and N Poisson distributed with parameter $\ell^2/2$ for some $\ell > 0$. If all these random variables are independent, then*

$$(X+\ell)^2 \stackrel{\mathrm{d}}{=} X^2 + 2\sum_{j=1}^{N} Z_j\,.$$

Proof. It suffices to show that the Laplace transforms of the random variables on the two sides of the equation agree. Let $\lambda > 0$. Completing the square, we find

$$\begin{aligned}\mathbb{E}\exp\{-\lambda\,(X+\ell)^2\} &= \frac{1}{\sqrt{2\pi}}\int \exp\{-\lambda\,(x+\ell)^2 - x^2/2\}\,dx\\ &= \frac{1}{\sqrt{2\pi}}\int \exp\Big\{-\tfrac{1}{2}\big(\sqrt{2\lambda+1}\,x + \tfrac{2\lambda\ell}{\sqrt{2\lambda+1}}\big)^2 - \lambda\ell^2 + \tfrac{2\lambda^2\ell^2}{2\lambda+1}\Big\}\,dx\\ &= \frac{1}{\sqrt{2\lambda+1}}\,\exp\Big\{-\tfrac{\lambda\ell^2}{2\lambda+1}\Big\}\,.\end{aligned}$$

From the special case $\ell = 0$ we get $\mathbb{E}\exp\{-\lambda\,X^2\} = \frac{1}{\sqrt{2\lambda+1}}$. For any $\theta > 0$,

$$\mathbb{E}[\theta^N] = \exp\{-\ell^2/2\}\sum_{k=0}^{\infty}\tfrac{(\ell^2\theta/2)^k}{k!} = \exp\{(\theta-1)\ell^2/2\}\,.$$

Using this and that $\mathbb{E}\exp\{-2\lambda Z_j\} = \frac{1}{2\lambda+1}$ we get

$$\mathbb{E}\exp\Big\{-\lambda\big(X^2+2\sum_{j=1}^{N} Z_j\big)\Big\} = \frac{1}{\sqrt{2\lambda+1}}\,\mathbb{E}\Big(\frac{1}{2\lambda+1}\Big)^N = \frac{1}{\sqrt{2\lambda+1}}\exp\Big\{-\tfrac{\lambda\ell^2}{2\lambda+1}\Big\},$$

which completes the proof. ■

Remark 6.33. An alternative proof of Lemma 6.32 will be given in Exercise 6.6. ◇

By combining the previous three lemmas we obtain the following convergence result for the conditional distribution of $D_n(x+h)$ given $D_n(x)$, which is the 'two-point version' of the Ray-Knight theorem.

Lemma 6.34. *Suppose nu_n are nonnegative, even integers and $u_n \to u$. For any $\lambda \geq 0$ we have*

$$\lim_{n\to\infty} \mathbb{E}\big[\exp\big\{-\lambda\tfrac{2}{n}D_n(x+h)\big\}\,\big|\,\tfrac{2}{n}D_n(x)=u_n\big] = \mathbb{E}_{(0,\sqrt{u})}\big[\exp\big\{-\lambda|W(h)|^2\big\}\big],$$

where $\{W(x)\colon x\geq 0\}$ denotes a planar Brownian motion started in $(0,\sqrt{u})\in\mathbb{R}^2$.

Proof. Combining Lemmas 6.30 and 6.31 we get

$$\begin{aligned}\lim_{n\to\infty} \mathbb{E}\big[\exp\big\{-\lambda\tfrac{2}{n}D_n(x+h)\big\}\,\big|\,\tfrac{2}{n}D_n(x)=u_n\big] &= \mathbb{E}\Big[\exp\Big\{-\lambda\Big(\tilde{X}^2+\tilde{Y}^2+2\sum_{j=1}^{M}\tilde{Z}_j\Big)\Big\}\Big]\\ &= \mathbb{E}\Big[\exp\Big\{-\lambda h\Big(X^2+Y^2+2\sum_{j=1}^{M}Z_j\Big)\Big\}\Big],\end{aligned}$$

with X, Y are standard normally distributed, $Z_1, Z_2, \ldots$ are standard exponentially distributed and M is Poisson distributed with parameter $\ell^2/2$, for $\ell=\sqrt{u/h}$. By Lemma 6.32 the right hand side can thus be rewritten as

$$\mathbb{E}\big[\exp\big\{-\lambda h\big((X+\sqrt{u/h})^2+Y^2\big)\big] = \mathbb{E}_{(0,\sqrt{u})}\big[\exp\big\{-\lambda|W(h)|^2\big\}\big],$$

which proves the lemma. ■

Now we complete the proof of Theorem 6.27. Note that, as both $\{L^x(T) : x\geq 0\}$ and $\{|W(x)|^2 : x\geq 0\}$ are continuous processes, it suffices to show that, for any

$$0 < x_1 < \cdots < x_m < a$$

the vectors

$$\big(L^{x_1}(T),\ldots,L^{x_m}(T)\big) \qquad \text{and} \qquad \big(|W(x_1)|^2,\ldots,|W(x_m)|^2\big)$$

have the same distribution. The Markov property of the downcrossing numbers, which approximate the local times, allows us to reduce this problem to the study of the 'two-point version'.

Lemma 6.35. *For all sufficiently large integers n, the process*

$$\{D_n(x_k)\colon k=1,\ldots,m\}$$

is a (possibly inhomogeneous) Markov chain.

Proof. Fix $k\in\{2,\ldots,m\}$. By Lemma 6.30 applied to $x=x_{k-1}$ and $h=x_k-x_{k-1}$ we can write $D_n(x_k)$ as a function of $D_n(x_{k-1})$ and various random variables, which by construction, are independent of $D_n(x_1),\ldots,D_n(x_{k-1})$. This establishes the Markov property. ∎

Note that, by rotational invariance of planar Brownian motion, $\{|W(x_k)|^2\colon k=1,\ldots,m\}$ is a Markov chain with transition probabilities given by

$$\mathbb{E}\big[\exp\{-\lambda\,|W(x_{k+1})|^2\}\,\big|\,|W(x_k)|^2=u\big]=\mathbb{E}_{(0,\sqrt{u})}\big[\exp\{-\lambda\,|W(x_{k+1}-x_k)|^2\}\big],$$

for all $\lambda>0$. The following general fact about the convergence of families of Markov chains ensures that we have done enough to complete the proof of Theorem 6.27.

Lemma 6.36. *Suppose, for $n=1,2,\ldots$, that $\{X_k^{(n)}:k=1,\ldots,m\}$ is a Markov chain with discrete state space $\Omega_n\subset[0,\infty)$ and that $\{X_k:k=1,\ldots,m\}$ is a Markov chain with state space $[0,\infty)$. Suppose further that*

(1) *$(X_1^{(n)},\ldots,X_m^{(n)})$ converges almost surely to some random vector $(Y_1,\ldots,Y_m)$,*

(2) *$X_1^{(n)}\Rightarrow X_1$ as $n\uparrow\infty$,*

(3) *for all $k=1,\ldots,m-1$, $\lambda>0$ and $y_n\in\Omega_n$ with $y_n\to y$, we have*

$$\lim_{n\to\infty}\mathbb{E}\big[\exp\{-\lambda X_{k+1}^{(n)}\}\,\big|\,X_k^{(n)}=y_n\big]=\mathbb{E}\big[\exp\{-\lambda X_{k+1}\}\,\big|\,X_k=y\big]\Big|.$$

Then

$$(X_1^{(n)},\ldots,X_m^{(n)})\Longrightarrow(X_1,\ldots,X_m)$$

and, in particular, the vectors $(X_1,\ldots,X_m)$ and $(Y_1,\ldots,Y_m)$ have the same distribution.

Proof. Recall from Proposition II.1.8 that it suffices to show that the Laplace transforms converge. Let $\lambda_1,\ldots,\lambda_m\geq 0$. By assumption (2) we have $X_1^{(n)}\Rightarrow X_1$ and hence we may assume, by way of induction, that for some fixed $k=1,\ldots,m-1$, we have

$$(X_1^{(n)},\ldots,X_k^{(n)})\Longrightarrow(X_1,\ldots,X_k)\,.$$

This implies, in particular, that $(X_1,\ldots,X_k)$ and $(Y_1,\ldots,Y_k)$ have the same distribution. Define

$$\Phi_n\colon\Omega_n\to[0,1],\qquad \Phi_n(y)=\mathbb{E}\big[\exp\{-\lambda_{k+1}X_{k+1}^{(n)}\}\,\big|\,X_k^{(n)}=y\big]$$

and

$$\Phi\colon[0,\infty)\to[0,1],\qquad \Phi(y)=\mathbb{E}\big[\exp\{-\lambda_{k+1}X_{k+1}\}\,\big|\,X_k=y\big]\,.$$

Then, combining assumption (1) and (3), $\Phi_n(X_k^{(n)}) \to \Phi(Y_k)$ almost surely. Hence, using this and once more assumption (1),

$$\mathbb{E}\Big[\exp\Big\{-\sum_{j=1}^{k+1}\lambda_j X_j^{(n)}\Big\}\Big] = \mathbb{E}\Big[\exp\Big\{-\sum_{j=1}^{k}\lambda_j X_j^{(n)}\Big\}\,\Phi_n(X_k^{(n)})\Big] \to \mathbb{E}\Big[\exp\Big\{-\sum_{j=1}^{k}\lambda_j Y_j\Big\}\,\Phi(Y_k)\Big].$$

As the vectors $(X_1,\dots,X_k)$ and $(Y_1,\dots,Y_k)$ have the same distribution the limit can be rewritten as

$$\mathbb{E}\Big[\exp\Big\{-\sum_{j=1}^{k}\lambda_j X_j\Big\}\,\Phi(X_k)\Big] = \mathbb{E}\Big[\exp\Big\{-\sum_{j=1}^{k+1}\lambda_j X_j\Big\}\Big],$$

and this completes the induction step.

Finally, as $(X_1^{(n)},\dots,X_m^{(n)})$ converges almost surely, and hence also in distribution to $(Y_1,\dots,Y_m)$, this vector must have the same distribution as $(X_1,\dots,X_m)$. This completes the proof. ■

Proof of Theorem 6.27. We use Lemma 6.36 with $X_k^{(n)} = \frac{2}{n}D_n(x_k)$, $X_k = |W(x_k)|^2$ and $Y_k = L^{x_k}(T)$. Then assumption (1) is satisfied by the downcrossing representation, assumption (2) follows from Lemma 6.29 and assumption (3) from Lemma 6.34. Lemma 6.36 thus gives that the random vectors $(L^{x_1}(T),\dots,L^{x_m}(T))$ and $(|W(x_1)|^2,\dots,|W(x_m)|^2)$ have the same distribution, which concludes the proof. ■

As an easy application of the Ray-Knight theorem, we answer the question whether, almost surely, *simultaneously* for all levels $x \in [0,a)$ the local times at level x are positive.

THEOREM 6.37 (Ray's theorem). *Suppose $a > 0$ and $\{B(t) : 0 \le t \le T_a\}$ is a linear Brownian motion started at zero and stopped at time $T_a = \inf\{t \ge 0 : B(t) = a\}$, when it reaches level a for the first time. Then, almost surely, $L^x(T_a) > 0$ for all $0 \le x < a$.*

Proof. The statement can be reworded as saying that the process $\{L^{a-x}(T_a) : 0 < x \le a\}$ almost surely does not hit zero. By the Ray-Knight theorem (applied to the Brownian motion $\{a - B(t) \colon t \ge 0\}$) this process agrees with $\{|W(x)|^2 : 0 < x \le a\}$ for a standard planar Brownian motion $\{W(x) : x \ge 0\}$ which, by Theorem 3.19, never returns to the origin. ■

Ray's theorem can be exploited to give a result on the Hausdorff dimension of the level sets of the Brownian motion, which holds *simultaneously* for all levels $a \in \mathbb{R}$.

THEOREM 6.38. *Almost surely,* $\dim\{t \ge 0 : B(t) = a\} \ge \frac{1}{2}$, *for all $a \in \mathbb{R}$.*

Proof. Obviously, it suffices to show that, for every fixed $a > 0$, almost surely,

$$\dim\big\{0 \le t < T_a : B(t) = x\big\} \ge \tfrac{1}{2} \qquad \text{for all } 0 \le x < a\,.$$

This can be achieved using the mass distribution principle. Considering the increasing function $L^x \colon [0,T^a) \to [0,\infty)$ as distribution function of a measure ℓ^x, we infer from Theorem 6.37 and Theorem 6.18 that, almost surely, for every $x \in [0,a)$, the measure ℓ^x is a mass distribution on the set $\{0 \le t < T_a : B(t) = x\}$. By Theorem 6.18, for any $\gamma < 1/2$, almost surely, there exists a (random) $C > 0$ such that, for all $x \in [0, a)$, $t \in [0, T_a)$ and $\varepsilon \in (0, 1)$,

$$\ell^x(t - \varepsilon, t + \varepsilon) \le |L^x(t + \varepsilon) - L^x(t - \varepsilon)| \le C\,(2\varepsilon)^\gamma\,.$$

The claim therefore follows from the mass distribution principle, Theorem 4.19. ■

Remark 6.39. Equality holds in Theorem 6.38. We will obtain the full result later as an easy corollary of Kaufman's dimension doubling theorem, see Theorem 9.28. ◇

4. Brownian local time as a Hausdorff measure

In this section we show that the local time $L^0(t)$ can be obtained as an intrinsically defined measure of the random set $\mathsf{Zero} \cap [0, t]$. The only family of intrinsically defined measures on metric spaces we have encountered so far in this book is the family of α-dimensional Hausdorff measures. As the α-dimensional Hausdorff measure of the zero set is always either zero (if $\alpha \ge \frac{1}{2}$) or infinity (if $\alpha < \frac{1}{2}$) we need to look out for an alternative construction.

We need not look very far. The definition of Hausdorff dimension still makes sense if we evaluate coverings by applying, instead of a simple power, an arbitrary nondecreasing function to the diameters of the sets in a covering.

Definition 6.40. *A nondecreasing function $\phi\colon [0, \varepsilon) \to [0, \infty)$ with $\phi(0) = 0$ defined on a nonempty interval $[0, \varepsilon)$ is called a* **(Hausdorff) gauge function**.

Let X be a metric space and $E \subset X$. For every gauge function ϕ and $\delta > 0$ define

$$\mathcal{H}^\phi_\delta(E) = \inf\Big\{ \sum_{i=1}^\infty \phi(|E_i|) \,:\, E_1, E_2, E_3, \ldots \text{ cover } E, \text{ and } |E_i| \le \delta \Big\}\,.$$

Then

$$\mathcal{H}^\phi(E) = \sup_{\delta > 0} \mathcal{H}^\phi_\delta(E) = \lim_{\delta \downarrow 0} \mathcal{H}^\phi_\delta(E)$$

is the **generalised ϕ-Hausdorff measure** *of the set E.* ◇

Theorem* 6.41. *There exists a constant $c > 0$ such that, almost surely, for all $t > 0$,*

$$L^0(t) = \mathcal{H}^\varphi\big(\mathsf{Zero} \cap [0, t]\big)\,,$$

for the gauge function $\varphi(r) = c\,\sqrt{r \log\log(1/r)}$.

The remainder of this section is devoted to the proof of this theorem. The material developed here will not be used in the remainder of the book. An important tool in the proof is the following classical theorem of Rogers and Taylor.

PROPOSITION 6.42 (Rogers-Taylor Theorem). *Let μ be a Borel measure on $\mathbb{R}^d$ and let ϕ be a Hausdorff gauge function.*

(i) *If $\Lambda \subset \mathbb{R}^d$ is a Borel set and*
$$\limsup_{r\downarrow 0} \frac{\mu\mathcal{B}(x,r)}{\phi(r)} < \alpha$$
for all $x \in \Lambda$, then $\mathcal{H}^\phi(\Lambda) \geq \alpha^{-1}\,\mu(\Lambda)$.

(ii) *If $\Lambda \subset \mathbb{R}^d$ is a Borel set and*
$$\limsup_{r\downarrow 0} \frac{\mu\mathcal{B}(x,r)}{\phi(r)} > \theta$$
for all $x \in \Lambda$, then $\mathcal{H}^\phi(\Lambda) \leq \kappa_d \theta^{-1}\mu(V)$ for any open set $V \subset \mathbb{R}^d$ that contains Λ, where κ_d depends only on d.

Moreover, in $d = 1$ one can also obtain an analogue of (i) for one-sided intervals.

(iii) *If $\Lambda \subset \mathbb{R}$ is a closed set and*
$$\limsup_{r\downarrow 0} \frac{\mu[x,x+r]}{\phi(r)} < \alpha$$
for all $x \in \Lambda$, then $\mathcal{H}^\phi(\Lambda) \geq \alpha^{-1}\,\mu(\Lambda)$.

REMARK 6.43. If μ is finite on compact sets, then $\mu(\Lambda)$ is the infimum of $\mu(V)$ over all open sets $V \supset \Lambda$, see for example [**Ru86**, 2.18]. Hence $\mu(V)$ can be replaced by $\mu(\Lambda)$ on the right hand side of the inequality in (ii). $\diamond$

Proof. **(i)** We write
$$\Lambda_\varepsilon = \left\{x \in \Lambda\colon \sup_{r\in(0,\varepsilon)} \tfrac{\mu\mathcal{B}(x,r)}{\phi(r)} < \alpha\right\}$$
and note that $\mu(\Lambda_\varepsilon) \to \mu(\Lambda)$ as $\varepsilon \downarrow 0$.

Fix $\varepsilon > 0$ and consider a cover $\{A_j\}$ of Λ_ε. Suppose that A_j intersects Λ_ε and $r_j = |A_j| < \varepsilon$ for all j. Choose $x_j \in A_j \cap \Lambda_\varepsilon$ for each j. Then $\mu\mathcal{B}(x_j, r_j) < \alpha\phi(r_j)$ for every j, whence
$$\sum_{j\geq 1} \phi(r_j) \geq \alpha^{-1} \sum_{j\geq 1} \mu\mathcal{B}(x_j, r_j) \geq \alpha^{-1}\mu(\Lambda_\varepsilon)\,.$$
Thus $\mathcal{H}^\phi_\varepsilon(\Lambda) \geq \mathcal{H}^\phi_\varepsilon(\Lambda_\varepsilon) \geq \alpha^{-1}\mu(\Lambda_\varepsilon)$. Letting $\varepsilon \downarrow 0$ proves (i).

(ii) Let $\varepsilon > 0$. For each $x \in \Lambda$, choose a positive $r_x < \varepsilon$ such that $\mathcal{B}(x, 2r_x) \subset V$ and $\mu\mathcal{B}(x, r_x) > \theta\phi(r_x)$; then among the dyadic cubes of diameter at most r_x that intersect $\mathcal{B}(x, r_x)$, let Q_x be a cube with $\mu(Q_x)$ maximal. (We consider here dyadic cubes of the form $\prod_{i=1}^d [a_i/2^m, (a_i+1)/2^m)$ where a_i are integers). In particular, $Q_x \subset V$ and $|Q_x| > r_x/2$ so the side-length of Q_x is at least $r_x/(2\sqrt{d})$. Let $N_d = 1 + 8\lceil\sqrt{d}\rceil$ and let Q_x^* be the cube with the same center z_x as Q_x, scaled by N_d (i.e., $Q_x^* = z_x + N_d(Q_x - z_x)$). Observe that Q_x^* contains

$\mathcal{B}(x, r_x)$, so $\mathcal{B}(x, r_x)$ is covered by at most N_d^d dyadic cubes that are translates of Q_x. Therefore, for every $x \in \Lambda$, we have

$$\mu(Q_x) \geq N_d^{-d} \mu \mathcal{B}(x, r_x) > N_d^{-d} \theta \phi(r_x)\,.$$

Let $\{Q_{x(j)} : j \geq 1\}$ be any enumeration of the maximal dyadic cubes among $\{Q_x : x \in \Lambda\}$. Then

$$\mu(V) \geq \sum_{j \geq 1} \mu(Q_{x(j)}) \geq N_d^{-d} \theta \sum_{j \geq 1} \phi(r_{x(j)})\,.$$

The collection of cubes $\{Q^*_{x(j)} : j \geq 1\}$ forms a cover of Λ. Since each of these cubes is covered by N_d^d cubes of diameter at most $r_{x(j)}$, we infer that

$$\mathcal{H}^{\phi}_{\varepsilon}(\Lambda) \leq N_d^d \sum_{j \geq 1} \phi(r_{x(j)}) \leq N_d^{2d} \theta^{-1} \mu(V).$$

Letting $\varepsilon \downarrow 0$ proves (ii).

(iii) Without loss of generality we may assume that μ has no atoms. Given $\varepsilon > 0$ we find $\delta > 0$ such that

$$\Lambda_\delta(\alpha) = \Big\{ t \in \Lambda \colon \sup_{h < \delta} \tfrac{\mu[t, t+h]}{\varphi(h)} \leq \alpha - \delta \Big\}$$

satisfies $\mu(\Lambda_\delta(\alpha)) > (1 - \varepsilon)\, \mu(\Lambda)$. Observe that $\Lambda_\delta(\alpha)$ is closed. Given a cover $\{\tilde{I}_j\}$ of Λ with $|\tilde{I}_j| < \delta$ we look at $I_j = [a_j, b_j]$ where a_j is the maximum and b_j the minimum of the closed set $\mathrm{cl}\, \tilde{I}_j \cap \Lambda_\delta(\alpha)$. Then $\{I_j\}$ covers $\Lambda_\delta(\alpha)$ and hence

$$\begin{aligned} \sum_{j \geq 1} \varphi(|\tilde{I}_j|) \geq \sum_{j \geq 1} \varphi(|I_j|) &\geq (\alpha - \delta)^{-1} \sum_{j \geq 1} \mu(I_j) \\ &\geq (\alpha - \delta)^{-1}\, \mu\big(\Lambda_\delta(\alpha)\big) \geq (\alpha - \delta)^{-1}\, (1 - \varepsilon)\, \mu(\Lambda)\,, \end{aligned}$$

and (iii) follows for $\delta \downarrow 0$, as $\varepsilon > 0$ was arbitrary. ■

For the proof of Theorem 6.41 we first note that, by Theorem 6.10, it is equivalent to show that, for the maximum process $\{M(t) : t \geq 0\}$ of a Brownian motion $\{B(t) : t \geq 0\}$, we have, almost surely,

$$M(t) = \mathcal{H}^{\varphi}\big(\mathsf{Rec} \cap [0, t]\big) \qquad \text{for all } t \geq 0\,,$$

where $\mathsf{Rec} = \{s \geq 0 \colon B(s) = M(s)\}$ denotes the set of record points of the Brownian motion. We define the measure μ on Rec as given by the distribution function M, i.e.

$$\mu(a, b] = M(b) - M(a) \qquad \text{for all intervals } (a, b] \subset \mathbb{R}.$$

Then μ is also the image measure of the Lebesgue measure on $[0, \infty)$ under the mapping

$$a \mapsto T_a := \inf\{s \geq 0 \,:\, B(t) = a\}.$$

The main part is to show that, for closed sets $\Lambda \subset [0, \infty)$,

$$c\, \mu(\Lambda) \leq \mathcal{H}^{\phi}\big(\Lambda \cap \mathsf{Rec}\big) \leq C\, \mu(\Lambda), \tag{4.1}$$

where $\phi(r) = \sqrt{r \log\log(1/r)}$ and c, C are positive constants.

The easier direction, the lower bound for the Hausdorff measure, follows from part **(iii)** of the Rogers-Taylor theorem and the upper bound in the law of the iterated logarithm. Indeed, for any level $a > 0$ let $T_a = \inf\{s \geq 0 \,:\, B(t) = a\}$. Then, by Corollary 5.3 applied to the standard Brownian motion $\{B(T_a + t) - B(T_a) \colon t \geq 0\}$, almost surely,

$$\limsup_{r\downarrow 0} \frac{M(T_a + r) - M(T_a)}{\sqrt{2r \log\log(1/r)}} = \limsup_{r\downarrow 0} \frac{B(T_a + r) - B(T_a)}{\sqrt{2r \log\log(1/r)}} = 1\,.$$

Defining the set

$$A = \big\{ s \in \mathsf{Rec} \colon \limsup_{r\downarrow 0} \mu[s, s+r]/\phi(r) \leq \sqrt{2} \big\},$$

this means that, for every $a > 0$, we have $T_a \in A$ almost surely. By Fubini's theorem,

$$\mathbb{E}\mu(A^c) = \mathbb{E} \int_0^\infty \mathbb{1}\{T_a \notin A\}\, da = \int_0^\infty \mathbb{P}\{T_a \notin A\}\, da = 0,$$

and hence, almost surely, $\mu(A^c) = 0$. By part **(iii)** of the Rogers-Taylor theorem, for every closed set $\Lambda \subset [0, \infty)$,

$$\mathcal{H}^\phi(\Lambda \cap \mathsf{Rec}) \geq \mathcal{H}^\phi(\Lambda \cap A) \geq \tfrac{1}{\sqrt{2}}\, \mu(\Lambda \cap A) = \tfrac{1}{\sqrt{2}}\, \mu(\Lambda)\,,$$

showing the left inequality in (4.1).

For the harder direction, the upper bound for the Hausdorff measure, it is important to note that the lower bound in Corollary 5.3 does not suffice. Instead, we need a law of the iterated logarithm which holds simultaneously for $\mathcal{H}^\phi$-almost all record times.

Lemma 6.44. *For every $\vartheta > 0$ small enough, almost surely,*

$$\mathcal{H}^\phi \Big\{ s \in \mathsf{Rec} \colon \limsup_{h\downarrow 0} \frac{M(s+h) - M(s-h)}{\phi(h)} < \vartheta \Big\} = 0\,.$$

Proof. We only need to prove that, for some $\vartheta > 0$, the set

$$\Lambda(\vartheta) = \big\{ s \in \mathsf{Rec} \cap (0,1) \colon \limsup_{h\downarrow 0} \tfrac{M(s+h) - M(s-h)}{\phi(h)} < \vartheta \big\}$$

satisfies $\mathcal{H}^\phi(\Lambda(\vartheta)) = 0$. Moreover, denoting

$$\Lambda_\delta(\vartheta) = \big\{ s \in \mathsf{Rec} \cap \big[\delta, 1-\delta\big] \colon \sup_{h<\delta} \tfrac{M(s+h) - M(s-h)}{\phi(h)} < \vartheta \big\},$$

we have

$$\Lambda(\vartheta) = \bigcup_{\delta > 0} \Lambda_\delta(\vartheta)\,.$$

It thus suffices to show, for fixed $\delta > 0$, that, almost surely,

$$\liminf_{n\uparrow\infty} \mathcal{H}^\phi_{1/n}(\Lambda_\delta(\vartheta)) = 0\,.$$

Fix $\delta > 0$ and a positive integer n such that $1/\sqrt{n} < \delta$. For parameters

$$A > 1, \theta > \vartheta \text{ and } q > 2,$$

which we choose later, we say that an interval of the form $I = [(k-1)/n, k/n]$ with $k \in \{1, \ldots, n\}$ is *good* if

(i) I contains a record point, in other words,

$$\tau := \inf\big\{t \geq \tfrac{k-1}{n} \colon B(t) = M(t)\big\} \leq \tfrac{k}{n}\,,$$

and either of the following two conditions hold,

(ii) there exists $j \geq 0$ with $1 \leq q^{j+1} \leq \sqrt{n}$ such that

$$B\big(\tau + \tfrac{q^j}{n}\big) - B(\tau) < -A\,\phi\big(\tfrac{q^j}{n}\big)\,;$$

(iii) for all $j \geq 0$ with $1 \leq q^{j+1} \leq \sqrt{n}$ we have that

$$B\big(\tau + \tfrac{q^{j+1}}{n}\big) - B\big(\tau + \tfrac{q^j}{n}\big) < \theta\,\phi\big(\tfrac{q^{j+1}-q^j}{n}\big)\,.$$

We now argue pathwise, and show that, given $A > 1$, $\theta > \vartheta$ we can find $q > 2$ such that the good intervals cover the set $\Lambda_\delta(\vartheta)$. Indeed, suppose that I is not good but contains a minimal record point $\tau \in [(k-1)/n, k/n]$. Then there exists $j \geq 0$ with $1 \leq q^{j+1} \leq \sqrt{n}$ such that

$$B\big(\tau + \tfrac{q^j}{n}\big) - B(\tau) \geq -A\,\phi\big(\tfrac{q^j}{n}\big) \quad \text{and} \quad B\big(\tau + \tfrac{q^{j+1}}{n}\big) - B\big(\tau + \tfrac{q^j}{n}\big) \geq \theta\,\phi\big(\tfrac{q^{j+1}-q^j}{n}\big)\,.$$

This implies that, for any $t \in [(k-1)/n, k/n] \cap \mathsf{Rec}$,

$$\begin{aligned} M\big(t + \tfrac{q^{j+1}}{n}\big) - M\big(t - \tfrac{q^{j+1}}{n}\big) &\geq M\big(\tau + \tfrac{q^{j+1}}{n}\big) - M\big(\tau\big) \geq B\big(\tau + \tfrac{q^{j+1}}{n}\big) - B\big(\tau\big) \\ &\geq \theta\,\phi\big(\tfrac{q^{j+1}-q^j}{n}\big) - A\,\phi\big(\tfrac{q^j}{n}\big) \geq \vartheta\,\phi\big(\tfrac{q^{j+1}}{n}\big), \end{aligned}$$

if q is chosen large enough. Hence the interval I does not intersect $\Lambda_\delta(\vartheta)$ and therefore the good intervals cover this set.

Next we show that, for any $A > \sqrt{2} > \theta$ and suitably chosen $C > 0$, for every $I = [(k-1)/n, k/n]$ with $I \cap [\delta, 1-\delta] \neq \emptyset$,

$$\mathbb{P}\big\{\big[\tfrac{k-1}{n}, \tfrac{k}{n}\big] \text{ is good }\big\} \leq C\,\frac{1}{\sqrt{n}}\,\Big(\frac{1}{\log n}\Big)^{\frac{A^2}{2}-1}. \tag{4.2}$$

By Lemma 4.22 in conjunction with Theorem 6.10 we get, for some constant $C_0 > 0$ depending only on $\delta > 0$,

$$\mathbb{P}\big\{\tau < \tfrac{k}{n}\big\} \leq C_0\,\frac{1}{\sqrt{n}}.$$

We further get, for some constant $C_1 > 0$, for all j with $q^{j+1} \leq \sqrt{n}$,

$$\begin{aligned} \mathbb{P}\Big\{B\big(\tau + \tfrac{q^j}{n}\big) - B(\tau) < -A\,\phi\big(\tfrac{q^j}{n}\big)\Big\} &\leq \mathbb{P}\big\{B(1) < -A\,\sqrt{\log\log(n/q^j)}\big\} \\ &\leq \exp\big\{-\tfrac{A^2}{2}\,\log\log(\sqrt{n})\big\} \leq C_1\,\Big(\frac{1}{\log n}\Big)^{\frac{A^2}{2}}. \end{aligned}$$

Using the independence of these events and summing over all $j \geq 0$ with $1 \leq q^{j+1} \leq \sqrt{n}$, of which there are no more than $C_2 \log n$, we get that

$$\mathbb{P}\big\{\big[\tfrac{k-1}{n}, \tfrac{k}{n}\big] \text{ satisfies (i) and (ii) }\big\} \leq C_0 C_1 C_2\,\frac{1}{\sqrt{n}}\,\Big(\frac{1}{\log n}\Big)^{\frac{A^2}{2}-1}. \tag{4.3}$$

To estimate the probability that $[(k-1)/n, k/n]$ satisfies (i) and (iii) we first note that

$$\mathbb{P}\Big\{B\big(\tfrac{q^{j+1}-q^j}{n}\big) < \theta\,\phi\big(\tfrac{q^{j+1}-q^j}{n}\big)\Big\} \le \mathbb{P}\Big\{B(1) < \theta\,\sqrt{\log\log\big(\tfrac{n}{q-1}\big)}\Big\}$$

$$\le 1 - \frac{\exp\big\{-\frac{\theta^2}{2}\log\log\big(\frac{n}{q-1}\big)\big\}}{\theta\,\sqrt{\log\log\big(\frac{n}{q-1}\big)}}\,,$$

using Lemma II.3.1. From this we infer that, for suitable $c_3 > 0$,

$$\mathbb{P}\Big\{B\big(\tfrac{q^{j+1}-q^j}{n}\big) < \theta\,\phi\big(\tfrac{q^{j+1}-q^j}{n}\big) \text{ for all } 1 \le q^{j+1} \le \sqrt{n}\Big\}$$

$$\le \prod_{j \le \frac{\log n}{2\log q}} \Big(1 - \frac{\exp\big\{-\frac{\theta^2}{2}\log\log n\big\}}{\theta\,\sqrt{\log\log n}}\Big) \le \Big(1 - \frac{1}{\theta\,(\log n)^{\frac{\theta^2}{2}}\,(\log\log n)^{\frac{1}{2}}}\Big)^{\frac{\log n}{2\log q}}$$

$$\le \exp\Big\{-c_3\,\frac{(\log n)^{1-\frac{\theta^2}{2}}}{(\log\log n)^{\frac{1}{2}}}\Big\}.$$

Combining this with the estimate for $\tau < k/n$ we get that

$$\text{(4.4)}\qquad \mathbb{P}\big\{\big[\tfrac{k-1}{n},\tfrac{k}{n}\big] \text{ satisfies (i) and (iii) }\big\} \le C_0\,\frac{1}{\sqrt{n}}\,\exp\Big\{-c_3\,\frac{(\log n)^{1-\frac{\theta^2}{2}}}{(\log\log n)^{\frac{1}{2}}}\Big\}.$$

As $\theta < \sqrt{2}$, the right hand side in (4.4) is of smaller order than the right hand side in (4.3) and hence we have shown (4.2).

Finally, we look at the expected ϕ-values of our covering. We obtain that

$$\mathbb{E}\mathcal{H}^{\phi}_{1/n}(\Lambda_\delta(\vartheta)) \le \sum_{k=\lceil \delta n\rceil}^{\lceil n(1-\delta)\rceil} \phi(1/n)\mathbb{P}\big\{\big[\tfrac{k-1}{n},\tfrac{k}{n}\big] \text{ is good }\big\} \le C\,\frac{\sqrt{\log\log(1/n)}}{(\log n)^{A^2/2-1}} \longrightarrow 0,$$

and, by Fatou's lemma we get, almost surely,

$$\liminf_{n\uparrow\infty} \mathcal{H}^{\phi}_{1/n}(\Lambda_\delta(\vartheta)) = 0\,,$$

as required to complete the proof. ■

The right inequality in (4.1) now follows easily from Lemma 6.44 and part **(ii)** of the Rogers-Taylor theorem. We define the set

$$A = \big\{s \in \mathsf{Rec}\colon \limsup_{r\downarrow 0} \mu\mathcal{B}(s,r)/\phi(r) \ge \vartheta\big\},$$

and note that $\mathcal{H}^{\phi}(\mathsf{Rec} \cap A^{\mathrm{c}}) = 0$, for ϑ sufficiently small. By part **(ii)** of the Rogers-Taylor theorem we get, for every Borel set $\Lambda \subset [0,\infty)$,

$$\mathcal{H}^{\phi}(\Lambda \cap \mathsf{Rec}) = \mathcal{H}^{\phi}(\Lambda \cap A) \le \kappa_1\vartheta^{-1}\,\mu(\Lambda \cap A) \le \kappa_1\vartheta^{-1}\,\mu(\Lambda).$$

This implies the right inequality and hence completes the proof of (4.1).

To complete the proof of Theorem 6.41 we look at the process $\{X(a)\colon a \geq 0\}$ defined by

$$X(a) = \mathcal{H}^{\phi}\big(\mathsf{Rec} \cap [0, T_a]\big)\,.$$

The next lemma will help us to show that this process is trivial.

Lemma 6.45. *Suppose $\{Y(t)\colon t \geq 0\}$ is stochastic process starting in zero with the following properties,*

- *the paths are almost surely continuous,*
- *the increments are independent, nonnegative and stationary,*
- *there exists a $C > 0$ such that, almost surely, $Y(t) \leq C\,t$ for all $t > 0$.*

Then there exists $\tilde{c} \geq 0$ such that, almost surely, $Y(t) = \tilde{c}\,t$ for every $t \geq 0$.

Proof. We first look at the function $m\colon [0,\infty) \to [0,\infty)$ defined by $m(t) = \mathbb{E}Y(t)$. This function is continuous, as the paths of $\{Y(t)\colon t \geq 0\}$ are continuous and bounded on compact sets. Further, because the process $\{Y(t)\colon t \geq 0\}$ has independent and stationary increments, the function m is linear and hence there exists $\tilde{c} \geq 0$ with $m(t) = \tilde{c}\,t$.

It thus suffices to show that the variance of $Y(t)$ is zero. Indeed, for every $n > 0$, we have

$$\begin{aligned}\operatorname{Var} Y(t) &= \sum_{j=1}^{n} \operatorname{Var}\Big(Y\big(\tfrac{kt}{n}\big) - Y\big(\tfrac{(k-1)t}{n}\big)\Big) = n \operatorname{Var} Y\big(\tfrac{t}{n}\big) \leq n\,\mathbb{E}\big[Y\big(\tfrac{t}{n}\big)^2\big] \\ &\leq n\,C^2\,\big(\tfrac{t}{n}\big)^2 \stackrel{n\to\infty}{\longrightarrow} 0,\end{aligned}$$

and hence $Y(t) = \mathbb{E}Y(t) = \tilde{c}\,t$ as claimed. ■

Let us check that $\{X(a)\colon a \geq 0\}$ satisfies the conditions of Lemma 6.45. We first note that

$$X(a+h) - X(a) = \mathcal{H}^{\phi}\big(\mathsf{Rec} \cap [0, T_{a+h}]\big) - \mathcal{H}^{\phi}\big(\mathsf{Rec} \cap [0, T_a]\big) = \mathcal{H}^{\phi}\big(\mathsf{Rec} \cap [T_a, T_{a+h}]\big),$$

as can be seen easily from the definition of the Hausdorff measure $\mathcal{H}^{\phi}$.

Using this, continuity of the paths follows from the fact that, by (4.1),

$$\mathcal{H}^{\phi}\big(\mathsf{Rec} \cap [T_a, T_{a+h}]\big) \leq C\,\big(M(T_{a+h}) - M(T_a)\big) = C\,h\,.$$

The strong Markov property implies that the increments are independent and stationary, and they are obviously nonnegative. And finally, by (4.1), almost surely, for any $a \geq 0$,

$$X(a) = \mathcal{H}^{\phi}\big(\mathsf{Rec} \cap [0, T_a]\big) \leq C\,M(T_a) = C\,a\,.$$

Lemma 6.45 thus implies that there exists $\tilde{c} \geq 0$ with

$$\mathcal{H}^{\phi}\big(\mathsf{Rec} \cap [0, T_a]\big) = \tilde{c}\,a = \tilde{c}\,M(T_a)$$

for all $a \geq 0$. The set $\{T_a\colon a \in \mathbb{R}\}$ is dense in Rec. Indeed, the only elements in $\mathsf{Rec}\setminus\{T_a\colon a \in \mathbb{R}\}$ are the countably many times when the Brownian motion revisits a local maximum for the first time. These times are stopping times and can therefore be approximated from the right by elements of $\{T_a\colon a \in \mathbb{R}\}$.

Using continuity, we infer that, almost surely, $\mathcal{H}^\phi\big(\mathsf{Rec} \cap [0,t]\big) = \tilde{c}\, M(t)$ for all $t \in \mathsf{Rec}$. For general $t \geq 0$ we let $\tau = \max(\mathsf{Rec} \cap [0,t])$ and note that

$$\mathcal{H}^\phi\big(\mathsf{Rec} \cap [0,t]\big) = \mathcal{H}^\phi\big(\mathsf{Rec} \cap [0,\tau]\big) = \tilde{c}\, M(\tau) = \tilde{c}\, M(t)\,.$$

By the lower bound in (4.1) we must have $\tilde{c} > 0$ and hence we can put $c = 1/\tilde{c}$ and get

$$M(t) = c\,\mathcal{H}^\phi\big(\mathsf{Rec} \cap [0,t]\big) = \mathcal{H}^{c\phi}\big(\mathsf{Rec} \cap [0,t]\big)\,,$$

as required to complete the proof of Theorem 6.41.

Exercises

Exercise 6.1. Using the downcrossing representation of the local time process $\{L(t)\colon t \geq 0\}$ given in Theorem 6.1, show that, almost surely, $L(s) = L(t)$ for every interval (s,t) not containing a zero of the Brownian motion. In other words, the local time at zero increases only on the zero set of the Brownian motion.

Exercise 6.2. Show, by reviewing the argument in the proof of Theorem 6.10, that for a standard linear Brownian motion the processes $\{(|B(t)|, L(t))\colon t \geq 0\}$ and $\{(M(t) - B(t), M(t))\colon t \geq 0\}$ have the same distribution.

Hint. In Theorem 7.36 we give an alternative proof of this result using stochastic integration.

Exercise 6.3. Show that $\mathbb{P}_0\{L(t) > 0 \text{ for every } t > 0\} = 1$.

Hint. This follows easily from Theorem 6.10.

Exercise 6.4 $(*)$. Let $\{W(s)\colon s \geq 0\}$ be a standard linear Brownian motion and τ_1 its first hitting time of level 1. Show that

$$\mathbb{E}\int_0^{\tau_1} \mathbb{1}\{0 \leq W(s) \leq 1\}\, ds = 1.$$

Hint. Use Exercise 2.15.

Exercise 6.5 $(*)$. Suppose $X_1, X_2, \ldots$ are independent geometrically distributed random variables with mean 2. Then, for sufficiently small $\varepsilon > 0$, for all nonnegative integers $k \leq m$,

$$\mathbb{P}\Big\{\Big|\sum_{j=1}^{k}(X_j - 2)\Big| \geq \varepsilon m\Big\} \leq 4\,\exp\big\{-\tfrac{1}{5}\,\varepsilon^2\, m\big\}\,.$$

Exercise 6.6 $(*)$. Give an alternative proof of Lemma 6.32 by computing the densities of the random variables $(X+\ell)^2$ and $X^2 + 2\sum_{j=1}^{N} Z_j$.

Exercise 6.7. Use the Ray-Knight theorem and Lévy's theorem, Theorem 6.10, to show that, for a suitable constant $c > 0$, the function

$$\varphi(h) = c\sqrt{h\log(1/h)} \qquad \text{for } 0 < h < 1,$$

is a modulus of continuity for the random field $\{L^a(t)\colon a \in \mathbb{R}, t \geq 0\}$.

Notes and Comments

The study of local times is crucial for the Brownian motion in dimension one and good references are [**RY94**] and the survey article [**Bo89**]. Brownian local times were first introduced by Paul Lévy in [**Le48**] and a thorough investigation is initiated in a paper by Trotter [**Tr58**] who showed that there is a version of local time continuous in time and space. An alternative construction of local times can be given in terms of stochastic integrals, using Tanaka's formula as a definition. We shall explore this direction in Section 7.3.

The equality for the upcrossing numbers in Lemma 6.3 agrees with the functional equation for a branching process with immigration. The relationship between local times and branching processes, which is underlying our entire treatment, can be exploited and extended in various ways. One example of this can be found in Neveu and Pitman [**NP89**], for more recent progress in this direction, see Le Gall and Le Jan, [**LL98**]. A good source for further reading is the discussion of Lévy processes and trees by Duquesne and Le Gall in [**DL02**]. For an introdution into branching processes with and without immigration, see [**AN04**].

In a similar spirit, a result which is often called the second Ray-Knight theorem describes the process $\{L_T^a : a > 0\}$ when $T = \inf\{t > 0 : L_t^0 = x\}$, see [**RY94**] or the original papers by Ray and Knight cited above. The resulting process is a Feller diffusion, which is the canonical process describing critical branching with initial mass x. The local times of Brownian motion can therefore be used to encode the branching information for a variety of processes describing the evolution of particles which undergo critical branching and spatial migration. For more information on this powerful link between Brownian motion and the world of spatial branching processes, see for example [**LG99**].

The concept of local times can be extended to a variety of processes like continuous semimartingales, see e.g. [**RY94**], or Markov processes [**BG68**]. The idea of introducing local times as densities of occupation measure has been fruitful in a variety of contexts, in particular in the introduction of local times on the intersection of Brownian paths. Important papers in this direction are [**GH80**] and [**GHR84**].

The Ray-Knight theorem was discovered by D. Ray and F. Knight independently by different methods in 1963. The proof of Knight uses discretisation, see [**Kn63**] for the original paper and [**Kn81**] for more information. Ray's approach to Theorem 6.27 is less intuitive but more versatile, and is based on the Feynman-Kac formula, see [**Ra63b**] for the original paper. Our presentation is simpler than Knight's method. The distributional identity at its core, see Lemma 6.32, is yet to be explained probabilistically. The analytic proof given in Exercise 6.6 is due to H. Robbins and E.J.G. Pitman [**RP49**].

Extensions of the Ray-Knight theorem includes a characterisation of $\{L^x(T) : x \geq 0\}$ for parameters exceeding a. This is best discussed in the framework of Brownian excursion theory, see for example [**RY94**]. The Ray-Knight theorem can be extended into a deep relationship between the local times of symmetric Markov processes and an associated Gaussian process, which is the subject of the famous Dynkin isomorphism theorem. See Eisenbaum [**Ei94**] or the comprehensive monograph by Marcus and Rosen [**MR06**] for more on this subject.

According to Taylor [**Ta86**], Hausdorff measures with arbitrary gauge functions were introduced by A.S. Besicovitch. General theory of outer measures, as presented in [**Ro99**] shows that $\mathcal{H}^\phi$ indeed defines a measures on the Borel sets of a metric space. The fact that the local time at zero agrees with the $\mathcal{H}^\phi$ Hausdorff measure of the zero set is due to Taylor and Wendel [**TW66**]. It can be shown that the local times $L^a(t)$ agree with the Hausdorff measure of the set $\{s \in [0,t] \colon B(s) = a\}$ simultaneously for all levels a and times t. This delicate result is proved in [**Pe81**] using nonstandard analysis.

The Rogers-Taylor theorem is due to C.A. Rogers and S.J. Taylor in [**RT61**]. The original statement is slightly more general as it allows to replace $\mu(V)$ by $\mu(\Lambda)$ on the right hand side without any regularity condition on μ. Most proofs in the literature of the harder half, statement (ii) in our formulation, use the Besicovitch covering theorem. We give a self-contained proof using dyadic cubes instead.

Other natural measures related to Brownian motion can also be shown to agree with Hausdorff measures with suitable gauge functions. The most notable example is the occupation measure, whose gauge function is

$$\varphi(r) = \begin{cases} c_d\, r^2 \, \log\log(1/r) & \text{if } d \geq 3, \\ c_2\, r^2 \, \log(1/r) \, \log\log\log(1/r) & \text{if } d = 2. \end{cases}$$

This result is due to Ciesielsky and Taylor [**CT62**] in the first case, and to Ray [**Ra63a**] and Taylor [**Ta64**] in the second case. A stimulating survey of this subject is [**LG85**].

CHAPTER 7

Stochastic integrals and applications

In this chapter we first construct an integral with respect to Brownian motion. Amongst the applications are the conformal invariance of Brownian motion, a short look at windings of Brownian motion, the Tanaka formula for Brownian local times, and the Feynman-Kac formula.

1. Stochastic integrals with respect to Brownian motion

1.1. Construction of the stochastic integral. We look at a Brownian motion in dimension one $\{B(t) : t \geq 0\}$ considered as a random continuous function. As we have found in Theorem 1.35, this function is almost surely of unbounded variation, which is why we cannot use *Lebesgue-Stieltjes integration* to define integrals of the form $\int_0^t f(s)\, dB(s)$. There is however an escape from this dilemma, if one is willing to take advantage of the fact that Brownian motions are *random* functions and therefore one can make use of weaker forms of limits. This is the idea of *stochastic integration*.

Before explaining the procedure, we have a look at a reasonable class of integrands, as we would like to admit random functions as integrands, too. At a first reading, the reader might prefer to think of deterministic integrands only and skip the next couple of paragraphs until we begin the construction of the integral after the proof of Lemma 7.2.

A suitable class of random integrands is the class of *progressively measurable processes*. We denote by $(\Omega, \mathcal{A}, \mathbb{P})$ the probability space on which our Brownian motion $\{B(t) \colon t \geq 0\}$ is defined and suppose that $(\mathcal{F}(t) : t \geq 0)$ is a filtration to which the Brownian motion is adapted such that the strong Markov property holds.

Because we also want the integral up to time t to be adapted to our filtration, we assume that the filtration $(\mathcal{F}(t) : t \geq 0)$ is **complete**, i.e. contains all sets of probability zero in $\mathcal{A}$. Note that every filtration can be completed simply by adding all these sets and that the completion preserves the strong Markov property.

Definition 7.1. *A process $\{X(t,\omega) : t \geq 0, \omega \in \Omega\}$ is called* **progressively measurable** *if for each $t \geq 0$ the mapping $X \colon [0,t] \times \Omega \to \mathbb{R}$ is measurable with respect to the σ-algebra $\mathfrak{B}([0,t]) \times \mathcal{F}(t)$.* $\diamond$

Lemma 7.2. *Any processes $\{X(t) \colon t \geq 0\}$, which is adapted and either right- or left-continuous is also progressively measurable.*

Proof. Assume that $\{X(t)\colon t \geq 0\}$ is right-continuous. Fix $t > 0$. For a positive integer n and $0 \leq s \leq t$ define $X_n(0,\omega) = X(0,\omega)$ and

$$X_n(s,\omega) = X\big(\tfrac{(k+1)t}{2^n}, \omega\big), \qquad \text{for } kt2^{-n} < s \leq (k+1)t2^{-n}\,.$$

The mapping $(s,\omega) \mapsto X_n(s,\omega)$ is $\mathfrak{B}([0,t]) \otimes \mathcal{F}(t)$ measurable. By right-continuity we have $\lim_{n\uparrow\infty} X_n(s,\omega) = X(s,\omega)$ for all $s \in [0,t]$ and $\omega \in \Omega$, hence the limit map $(s,\omega) \mapsto X(s,\omega)$ is also $\mathfrak{B}([0,t]) \otimes \mathcal{F}(t)$ measurable, proving progressive measurability. The left-continuous case is analogous. ■

The construction of the integrals is quite straightforward. We start by integrating progressively measurable step processes $\{H(t,\omega)\,:\, t \geq 0,\, \omega \in \Omega\}$ of the form

$$H(t,\omega) = \sum_{i=1}^{k} A_i(\omega) \mathbb{1}_{(t_i,t_{i+1}]}(t), \quad \text{ for } 0 \leq t_1 \leq \ldots \leq t_{k+1}, \text{ and } \mathcal{F}(t_i)\text{-measurable } A_i.$$

In complete analogy to the classical case we define the integral as

$$\int_0^\infty H(s)\, dB(s) := \sum_{i=1}^{k} A_i\big(B(t_{i+1}) - B(t_i)\big).$$

Now let H be a progressively measurable process satisfying $\mathbb{E}\int_0^\infty H(s)^2\, ds < \infty$. Suppose H can be approximated by a family of progressively measurable step processes H_n, $n \geq 1$, then we define

$$\int_0^\infty H(s)\, dB(s) := \mathrm{L}^2 - \lim_{n\to\infty} \int_0^\infty H_n(s)\, dB(s). \tag{1.1}$$

At this stage we focus on L^2-convergence, though we shall see later that the stochastic integral can also be constructed as an almost sure limit, see Remark 7.7. For the approximation of H by progressively measurable step processes we look at the norm

$$\|H\|_2^2 := \mathbb{E}\int_0^\infty H(s)^2\, ds.$$

What we have to show now to complete the definition is that,

(1) every progressively measurable process satisfying $\mathbb{E}\int_0^\infty H(s)^2\, ds < \infty$ can be approximated in the $\|\cdot\|_2$ norm by progressively measurable step processes,
(2) for each approximating sequence the limit in (1.1) exists,
(3) and this limit does not depend on the choice of the approximating step processes.

This is what we check now, beginning with item (1).

Lemma 7.3. *For every progressively measurable process $\{H(s,\omega)\,:\, s \geq 0,\, \omega \in \Omega\}$ satisfying $\mathbb{E}\int_0^\infty H(s)^2\, ds < \infty$ there exists a sequence $\{H_n\,:\, n \in \mathbb{N}\}$ of progressively measurable step processes such that* $\lim_{n\to\infty} \|H_n - H\|_2 = 0$.

Proof. The strategy is to approximate the progressively measurable process successively by

- a bounded progressively measurable process,
- a bounded, almost surely continuous progressively measurable process,
- and finally, by a progressively measurable step process.

Let $\{H(s,\omega) : s \geq 0, \omega \in \Omega\}$ be a progressively measurable process with $\|H\|_2 < \infty$. We *first* define the cut-off at a fixed time $n > 0$ by letting $H_n(s,\omega) = H(s,\omega)$ for $s \leq n$ and $H_n(s,\omega) = 0$ otherwise. Clearly $\lim_{n\uparrow\infty} \|H_n - H\|_2 = 0$.

Second, we approximate any progressively measurable H on a finite interval by truncating its values, i.e. for large n we define H_n by letting $H_n(s,\omega) = H(s,\omega) \wedge n$. Clearly H_n is progressively measurable and $\lim_{n\uparrow\infty} \|H_n - H\|_2 = 0$.

Third, we approximate any uniformly bounded progressively measurable H by a bounded, almost-surely continuous, progressively measurable process. Let $h = 1/n$ and, using the convention $H(s,\omega) = H(0,\omega)$ for $s < 0$ we define

$$H_n(s,\omega) = \frac{1}{h} \int_{s-h}^{s} H(s,\omega)\, ds.$$

Because we only take an average over the past, H_n is again progressively measurable. It is almost surely continuous and it is a well-known fact that, for every $\omega \in \Omega$ and almost every $s \in [0,t]$,

$$\lim_{h\downarrow 0} \frac{1}{h} \int_{s-h}^{s} H(t,\omega)\, dt = H(s,\omega)\,.$$

Since H is uniformly bounded (and using progressive measurability) we can take expectations and an average over time, and obtain from the bounded convergence theorem that

$$\lim_{n\uparrow\infty} \|H_n - H\|_2 = 0\,.$$

Finally, a bounded, almost-surely continuous, progressively measurable process can be approximated by a step process H_n by taking $H_n(s,\omega) = H(j/n,\omega)$ for $j/n \leq s < (j+1)/n$. These functions are again progressively measurable and one easily sees $\lim_{n\uparrow\infty} \|H_n - H\|_2 = 0$. This completes the approximation. ■

The following lemma describes the crucial property of the integral of step processes.

LEMMA 7.4. *Let H be a progressively measurable step process and $\mathbb{E}\int_0^\infty H(s)^2\, ds < \infty$, then*

$$\mathbb{E}\Big[\Big(\int_0^\infty H(s)\, dB(s)\Big)^2\Big] = \mathbb{E}\int_0^\infty H(s)^2\, ds.$$

Proof. We use the Markov property to see that, for every progressively measurable step process $H = \sum_{i=1}^k A_i \mathbb{1}_{(a_i, a_{i+1}]}$,

$$\mathbb{E}\Big[\Big(\int_0^\infty H(s)\,dB(s)\Big)^2\Big] = \mathbb{E}\Big[\sum_{i,j=1}^k A_i A_j \big(B(a_{i+1}) - B(a_i)\big)\big(B(a_{j+1}) - B(a_j)\big)\Big]$$

$$= 2\sum_{i=1}^k \sum_{j=i+1}^k \mathbb{E}\big[A_i A_j \big(B(a_{i+1}) - B(a_i)\big)\mathbb{E}\big[B(a_{j+1}) - B(a_j)\,\big|\,\mathcal{F}(a_j)\big]\big]$$

$$+ \sum_{i=1}^k \mathbb{E}\big[A_i^2 \big(B(a_{i+1}) - B(a_i)\big)^2\big]$$

$$= \sum_{i=1}^k \mathbb{E}\big[A_i^2\big]\,\big(a_{i+1} - a_i\big) = \mathbb{E}\int_0^\infty H(s)^2\,ds.$$ ■

COROLLARY 7.5. *Suppose* $\{H_n : n \in \mathbb{N}\}$ *is a sequence of progressively measurable step processes such that*

$$\mathbb{E}\int_0^\infty \big(H_n(s) - H_m(s)\big)^2\,ds \longrightarrow 0, \ \text{as } n, m \to \infty,$$

then

$$\mathbb{E}\Big[\Big(\int_0^\infty H_n(s) - H_m(s)\,dB(s)\Big)^2\Big] \longrightarrow 0, \ \text{as } n, m \to \infty.$$

Proof. Because the difference of two step processes is again a step process, Lemma 7.4 can be applied to $H_n - H_m$ and this gives the statement. ■

The following theorem addresses issues (2) and (3), thus completing our construction of the stochastic integral.

THEOREM 7.6. *Suppose* $\{H_n : n \in \mathbb{N}\}$ *is a sequence of progressively measurable step processes and* H *a progressively measurable process such that*

$$\lim_{n\to\infty} \mathbb{E}\int_0^\infty \big(H_n(s) - H(s)\big)^2\,ds = 0,$$

then

$$\mathrm{L}^2 - \lim_{n\to\infty} \int_0^\infty H_n(s)\,dB(s) =: \int_0^\infty H(s)\,dB(s)$$

exists and is independent of the choice of $\{H_n : n \in \mathbb{N}\}$. *Moreover, we have*

$$\mathbb{E}\Big[\Big(\int_0^\infty H(s)\,dB(s)\Big)^2\Big] = \mathbb{E}\int_0^\infty H(s)^2\,ds. \tag{1.2}$$

Remark 7.7. If the sequence of step processes is chosen such that

$$\sum_{n=1}^{\infty} \mathbb{E} \int_0^{\infty} \big(H_n(s) - H(s)\big)^2 \, ds < \infty,$$

then, by (1.2), we get $\sum_{n=1}^{\infty} \mathbb{E}[(\int_0^{\infty} H_n(s) - H(s)\, dB(s))^2] < \infty$, and therefore, almost surely,

$$\sum_{n=1}^{\infty} \Big[\int_0^{\infty} H_n(s)\, dB(s) - \int_0^{\infty} H(s)\, dB(s) \Big]^2 < \infty.$$

This implies that, almost surely,

$$\lim_{n\to\infty} \int_0^{\infty} H_n(s)\, dB(s) = \int_0^{\infty} H(s)\, dB(s)\,. \qquad \diamond$$

Proof of Theorem 7.6. By the triangle inequality $\{H_n : n \in \mathbb{N}\}$ satisfies the assumption of Corollary 7.5, and hence $\{\int_0^{\infty} H_n(s)\, dB(s) : n \in \mathbb{N}\}$ is a Cauchy sequence in $\mathrm{L}^2(\mathbb{P})$. By completeness of this space, the limit exists, and Corollary 7.5 also shows that the limit is independent of the choice of the approximating sequence. The last statement follows from Lemma 7.4, applied to H_n, by taking the limit $n \to \infty$. ∎

Finally, we describe the stochastic integral as a stochastic process in time. The crucial property of this process are continuity and the *martingale property* stated in the next theorem.

Definition 7.8. *Suppose $\{H(s,\omega)\colon s \geq 0\,, \omega \in \Omega\}$ is progressively measurable with $\mathbb{E}\int_0^t H(s,\omega)^2\, ds < \infty$. Define the progressively measurable process $\{H^t(s,\omega)\colon s \geq 0\,, \omega \in \Omega\}$ by*

$$H^t(s,\omega) = H(s,\omega)\, \mathbb{1}\{s \leq t\}\,.$$

Then the **stochastic integral up to** t *is defined as,*

$$\int_0^t H(s)\, dB(s) := \int_0^{\infty} H^t(s)\, dB(s)\,. \qquad \diamond$$

Definition 7.9. *We say that a stochastic process $\{X(t)\colon t \geq 0\}$ is a* **modification** *of a process $\{Y(t)\colon t \geq 0\}$ if, for every $t \geq 0$, we have $\mathbb{P}\{X(t) = Y(t)\} = 1$.* $\diamond$

Theorem 7.10. *Suppose the process $\{H(s,\omega)\colon s \geq 0\,, \omega \in \Omega\}$ is progressively measurable with*

$$\mathbb{E}\int_0^t H(s,\omega)^2\, ds < \infty \quad \text{for any } t \geq 0.$$

Then there exists an almost surely continuous modification of $\{\int_0^t H(s)\, dB(s)\colon t \geq 0\}$. Moreover, this process is a martingale and hence

$$\mathbb{E}\int_0^t H(s)\, dB(s) = 0 \quad \text{for every } t \geq 0.$$

Proof. Fix a large integer t_0 and let H_n be a sequence of step processes such that $\|H_n - H^{t_0}\|_2 \to 0$, and therefore

$$\mathbb{E}\Big[\Big(\int_0^\infty \big(H_n(s) - H^{t_0}(s)\big)\, dB(s)\Big)^2\Big] \to 0\,.$$

Obviously, for any $s \le t$ the random variable $\int_0^s H_n(u)\, dB(u)$ is $\mathcal{F}(s)$-measurable and $\int_s^t H_n(u)\, dB(u)$ is independent of $\mathcal{F}(s)$, meaning that the process

$$\Big\{\int_0^t H_n(u)\, dB(u) \colon 0 \le t \le t_0\Big\}$$

is a martingale, for every n. For any $0 \le t \le t_0$ define

$$X(t) = \mathbb{E}\Big[\int_0^{t_0} H(s)\, dB(s)\,\Big|\,\mathcal{F}(t)\Big]\,,$$

so that $\{X(t) : 0 \le t \le t_0\}$ is also a martingale and

$$X(t_0) = \int_0^{t_0} H(s)\, dB(s).$$

By Doob's maximal inequality, Proposition 2.39, for $p = 2$,

$$\mathbb{E}\Big[\sup_{0\le t\le t_0}\Big(\int_0^t H_n(s)\, dB(s) - X(t)\Big)^2\Big] \le 4\,\mathbb{E}\Big[\Big(\int_0^{t_0} \big(H_n(s) - H(s)\big)\, dB(s)\Big)^2\Big]\,,$$

which converges to zero, as $n \to \infty$. This implies, in particular, that almost surely, the process $\{X(t) : 0 \le t \le t_0\}$ is a uniform limit of continuous processes, and hence continuous. For fixed $0 \le t \le t_0$, by taking L^2-limits from the step process approximation, the random variable $\int_0^t H(s)\, dB(s)$ is $\mathcal{F}(t)$-measurable and $\int_t^{t_0} H(s)\, dB(s)$ is independent of $\mathcal{F}(t)$ with zero expectation. Therefore $\int_0^t H(s)\, dB(s)$ is a conditional expectation of $X(t_0)$ given $\mathcal{F}(t)$, hence coinciding with $X(t)$ almost surely. ■

We now have a basic stochastic integral at our disposal. Obviously, a lot of bells and whistles can be added to this construction, but we refrain from doing so and keep focused on the essential properties and eventually on the applications to Brownian motion.

1.2. Itô's formula. For stochastic integration Itô's formula plays the same role as the fundamental theorem of calculus for classical integration. Let f be continuously differentiable and $x\colon [0,\infty) \to \mathbb{R}$, then the fundamental theorem can be written as

$$f(x(t)) - f(x(0)) = \int_0^t f'(x(s))\, dx(s)\,,$$

and this formula holds when x is continuous and of bounded variation. Itô's formula offers an analogue of this for the case that x is a Brownian motion. The crucial difference is that a third term enters, which makes the existence of a second derivative of f necessary. The next result, a key step in the derivation of this formula, is an extension of Exercise 1.14.

THEOREM 7.11. *Suppose $f\colon \mathbb{R} \to \mathbb{R}$ is continuous, $t > 0$, and $0 = t_1^{(n)} < \ldots < t_n^{(n)} = t$ are partitions of the interval $[0,t]$, such that the mesh converges to zero. Then, in probability,*

$$\sum_{j=1}^{n-1} f\big(B(t_j^{(n)})\big)\,\big(B(t_{j+1}^{(n)}) - B(t_j^{(n)})\big)^2 \longrightarrow \int_0^t f(B(s))\,ds\,.$$

Proof. Let T be the first exit time from a compact interval. It suffices to prove the statement for Brownian motion stopped at T, as the interval may be chosen to make $\mathbb{P}\{T < t\}$ arbitrarily small. By continuity of f and the definition of the Riemann integral, almost surely,

$$\lim_{n\to\infty} \sum_{j=1}^{n-1} f\big(B(t_j^{(n)} \wedge T)\big)\,\big(t_{j+1}^{(n)} \wedge T - t_j^{(n)} \wedge T\big) = \int_0^{t\wedge T} f(B(s))\,ds\,.$$

It thus suffices to show that

$$\lim_{n\to\infty} \mathbb{E}\Big[\Big(\sum_{j=1}^{n-1} f\big(B(t_j^{(n)} \wedge T)\big)\,\Big(\big(B(t_{j+1}^{(n)} \wedge T) - B(t_j^{(n)} \wedge T)\big)^2 - \big(t_{j+1}^{(n)} \wedge T - t_j^{(n)} \wedge T\big)\Big)\Big)^2\Big] = 0.$$

Recall that $\{B(t)^2 - t \,:\, t \geq 0\}$ is a martingale, by Lemma 2.43, and hence, for all $r \leq s$,

$$\mathbb{E}\big[\big(B(s) - B(r)\big)^2 - (s-r)\,\big|\,\mathcal{F}(r)\big] = 0\,.$$

This allows us to simplify the previous expression as follows,

$$\begin{aligned}
&\mathbb{E}\Big[\Big(\sum_{j=1}^{n-1} f\big(B(t_j^{(n)} \wedge T)\big)\,\Big(\big(B(t_{j+1}^{(n)} \wedge T) - B(t_j^{(n)} \wedge T)\big)^2 - \big(t_{j+1}^{(n)} \wedge T - t_j^{(n)} \wedge T\big)\Big)\Big)^2\Big] \\
&\qquad = \sum_{j=1}^{n-1} \mathbb{E}\Big[f\big(B(t_j^{(n)} \wedge T)\big)^2\Big(\big(B\big(t_{j+1}^{(n)} \wedge T\big) - B\big(t_j^{(n)} \wedge T\big)\big)^2 - \big(t_{j+1}^{(n)} \wedge T - t_j^{(n)} \wedge T\big)\Big)^2\Big].
\end{aligned}$$

We can now bound f by its maximum on the compact interval, and multiplying out the square and dropping a negative cross term we get an upper bound, which is a constant multiple of

$$\sum_{j=1}^{n-1} \mathbb{E}\Big[\big(B(t_{j+1}^{(n)} \wedge T) - B(t_j^{(n)} \wedge T)\big)^4\Big] + \sum_{j=1}^{n-1} \mathbb{E}\Big[\big(t_{j+1}^{(n)} \wedge T - t_j^{(n)} \wedge T\big)^2\Big]. \tag{1.3}$$

Using Brownian scaling on the first term, we see that this expression is bounded by a constant multiple of

$$\sum_{j=1}^{n-1} \big(t_{j+1}^{(n)} - t_j^{(n)}\big)^2 \leq t\,\Delta(n),$$

where $\Delta(n)$ denotes the mesh, which goes to zero. This completes the proof. ∎

We are now able to formulate and prove a first version of Itô's formula.

THEOREM 7.12 (Itô's formula I). *Let $f\colon \mathbb{R} \to \mathbb{R}$ be twice continuously differentiable such that $\mathbb{E}\int_0^t f'\big(B(s)\big)^2\,ds < \infty$ for some $t > 0$. Then, almost surely, for all $0 \leq s \leq t$,*

$$f\big(B(s)\big) - f\big(B(0)\big) = \int_0^s f'\big(B(u)\big)\,dB(u) + \tfrac{1}{2}\int_0^s f''\big(B(u)\big)\,du.$$

Proof. We denote the modulus of continuity of f'' on $[-M, M]$ by

$$\omega(\delta, M) := \sup_{\substack{x,y\in[-M,M]\\|x-y|<\delta}} f''(x) - f''(y)\,.$$

Then, by Taylor's formula, for any $x, y \in [-M, M]$ with $|x - y| < \delta$,

$$\big|f(y) - f(x) - f'(x)(y-x) - \tfrac{1}{2}f''(x)(y-x)^2\big| \le \omega(\delta, M)\,(y-x)^2\,.$$

Now, for any sequence $0 = t_1 < \ldots < t_n = t$ with $\delta_B := \max_{1\le i\le n-1}\big|B(t_{i+1}) - B(t_i)\big|$ and $M_B = \max_{0\le s\le t}|B(s)|$, we get

$$\begin{aligned}\Big|\sum_{i=1}^{n-1}\big(f(B(t_{i+1})) - f(B(t_i))\big) - \sum_{i=1}^{n-1} f'\big(B(t_i)\big)\big(B(t_{i+1}) - B(t_i)\big)\\ - \sum_{i=1}^{n-1}\tfrac{1}{2}f''\big(B(t_i)\big)\big(B(t_{i+1}) - B(t_i)\big)^2\Big| \le \omega(\delta_B, M_B)\sum_{i=1}^{n-1}\big(B(t_{i+1}) - B(t_i)\big)^2\,.\end{aligned}$$

Note that the first sum is simply $f(B(t)) - f(B(0))$. By the definition of the stochastic integral and Theorem 7.11 we can choose a sequence of partitions with mesh going to zero, such that, almost surely, the first subtracted term on the left converges to $\int_0^t f'\big(B(s)\big)\,dB(s)$, the second subtracted term converges to $\frac{1}{2}\int_0^t f''\big(B(s)\big)\,ds$, and the sum on the right hand side converges to t. By continuity of the Brownian path $\omega(\delta_B, M_B)$ converges almost surely to zero. This proves Itô's formula for fixed t, or indeed almost surely for all rational times $0 \le s \le t$. As all the terms in Itô's formula are continuous almost surely, we get the result simultaneously for all $0 \le s \le t$. ■

Next, we provide an enhanced version of Itô's formula, which allows the function f to depend not only on the position of Brownian motion, but also on a second argument, which is assumed to be increasing in time.

THEOREM 7.13 (Itô's formula II). *Suppose $\{\zeta(s) \colon s \ge 0\}$ is an increasing, continuous adapted stochastic process. Let $f \colon \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ be twice continuously differentiable in the x-coordiante, and once continuously differentiable in the y-coordinate. Assume that*

$$\mathbb{E}\int_0^t \Big[\partial_x f(B(s), \zeta(s))\Big]^2\,ds < \infty,$$

for some $t > 0$. Then, almost surely, for all $0 \le s \le t$,

$$\begin{aligned}f\big(B(s), \zeta(s)\big) - f\big(B(0), \zeta(0)\big) = \int_0^s \partial_x f\big(B(u), \zeta(u)\big)\,dB(u) + \int_0^s \partial_y f\big(B(u), \zeta(u)\big)\,d\zeta(u)\\ + \frac{1}{2}\int_0^s \partial_{xx} f\big(B(u), \zeta(u)\big)\,du.\end{aligned}$$

Proof. To begin with, we inspect the proof of Theorem 7.11 and see that it carries over without difficulty to the situation, when f is allowed to depend additionally on an adapted process $\{\zeta(s)\colon s \geq 0\}$, i.e. we have for any partitions $0 = t_1^{(n)} < \ldots < t_n^{(n)} = t$ with mesh going to zero, in probability,

$$\sum_{j=1}^{n-1} f\big(\zeta\big(t_j^{(n)}\big), B\big(t_j^{(n)}\big)\big) \Big(B\big(t_{j+1}^{(n)}\big) - B\big(t_j^{(n)}\big)\Big)^2 \longrightarrow \int_0^t f(\zeta(s), B(s))\, ds\,. \tag{1.4}$$

We denote the modulus of continuity of $\partial_y f$ by

$$\omega_1(\delta, M) = \sup_{\substack{-M \leq x_1, x_2, y_1, y_2 \leq M \\ |x_1 - x_2| \wedge |y_1 - y_2| < \delta}} \partial_y f(x_1, y_1) - \partial_y f(x_2, y_2),$$

and the modulus of continuity of $\partial_{xx} f$ by

$$\omega_2(\delta, M) = \sup_{\substack{-M \leq x_1, x_2, y_1, y_2 \leq M \\ |x_1 - x_2| \wedge |y_1 - y_2| < \delta}} \partial_{xx} f(x_1, y_1) - \partial_{xx} f(x_2, y_2).$$

Now take $x, x_0, y, y_0 \in [-M, M]$ with $|x - x_0| \wedge |y - y_0| < \delta$. By the mean value theorem, there exists a value $\tilde{y} \in [-M, M]$ with the property that$|\tilde{y} - y| \wedge |\tilde{y} - y_0| < \delta$ such that

$$f(x, y) - f(x, y_0) = \partial_y f(x, \tilde{y})\,(y - y_0),$$

and hence

$$\big|f(x, y) - f(x, y_0) - \partial_y f(x_0, y_0)\,(y - y_0)\big| \leq \omega_1(M, \delta)\,(y - y_0).$$

Taylor's formula implies that

$$\big|f(x, y_0) - f(x_0, y_0) - \partial_x f(x_0, y_0)(x - x_0) - \tfrac{1}{2}\,\partial_{xx} f(x_0, y_0)(x - x_0)^2\big| \leq \omega_2(\delta, M)(x - x_0)^2.$$

Combining the last two formulas using the triangle inequality, we get that

$$\begin{aligned} \big|f(x, y) - f(x_0, y_0) &- \partial_y f(x_0, y_0)\,(y - y_0) \\ &- \partial_x f(x_0, y_0)\,(x - x_0) - \tfrac{1}{2}\partial_{xx} f(x_0, y_0)(x - x_0)^2\big| \\ &\leq \omega_1(\delta, M)\,(y - y_0) + \omega_2(\delta, M)(x - x_0)^2. \end{aligned} \tag{1.5}$$

Now, for any sequence $0 = t_1 < \ldots < t_n = t$ define

$$\delta := \max_{1 \leq i \leq n-1} \big|B(t_{i+1}) - B(t_i)\big| \wedge \max_{1 \leq i \leq n-1} \big|\zeta(t_{i+1}) - \zeta(t_i)\big|,$$

and

$$M := \max_{0 \leq s \leq t} |B(s)| \wedge \max_{0 \leq s \leq t} |\zeta(s)|.$$

We get from (1.5),

$$\begin{aligned} \Big|f\big(B(t), \zeta(t)\big) &- f\big(B(0), \zeta(0)\big) - \sum_{i=1}^{n-1} \partial_x f\big(B(t_i), \zeta(t_i)\big)\,\big(B(t_{i+1}) - B(t_i)\big) \\ &- \sum_{i=1}^{n-1} \partial_y f\big(B(t_i), \zeta(t_i)\big)\,\big(\zeta(t_{i+1}) - \zeta(t_i)\big) - \tfrac{1}{2} \sum_{i=1}^{n-1} \partial_{xx} f\big(B(t_i), \zeta(t_i)\big)\,\big(B(t_{i+1}) - B(t_i)\big)^2\Big| \\ &\leq \omega_1(\delta, M)\big(\zeta(t) - \zeta(0)\big) + \omega_2(\delta, M) \sum_{i=1}^{n-1} \big(B(t_{i+1}) - B(t_i)\big)^2\,. \end{aligned}$$

We can choose a sequence of partitions with mesh going to zero, such that, almost surely, the following convergence statements hold,

- the first sum on the left converges to $\int_0^t \partial_x f\big(B(s),\zeta(s)\big)\,dB(s)$ by the definition of the stochastic integral,
- the second sum on the left converges to $\int_0^t \partial_y f\big(B(s),\zeta(s)\big)\,d\zeta(s)$ by definition of the Stieltjes integral,
- the third sum on the left converges to $\frac{1}{2}\int_0^t \partial_{xx} f\big(B(s),\zeta(s)\big)\,ds$ by (1.4),
- the sum on the right hand side converges to t by Theorem 7.11.

By continuity of the Brownian path $\omega_1(\delta,M)$ and $\omega_2(\delta,M)$ converge almost surely to zero. This proves the enhanced Itô's formula for fixed t, and looking at rationals and exploiting continuity as before, we get the result simultaneously for all $0 \le s \le t$. ■

With exactly the same technique, we obtain a version of Itô's formula for higher dimensional Brownian motion. The detailed proof will be an exercise, see Exercise 7.3. To give a pleasant formulation, we introduce some notation for functions $f\colon \mathbb{R}^{d+m} \to \mathbb{R}$, where we interpret the argument as two vectors, $x \in \mathbb{R}^d$ and $y \in \mathbb{R}^m$. We write ∂_j for the partial derivative in direction of the jth coordinate, and

$$\nabla_x f = (\partial_1 f, \ldots, \partial_d f) \quad \text{and} \quad \nabla_y f = (\partial_{d+1} f, \ldots, \partial_{d+m} f)$$

for the vector of derivatives in the directions of x, respectively y. For integrals we use the scalar product notation

$$\int_0^t \nabla_x f\big(B(u),\zeta(u)\big)\cdot dB(u) = \sum_{i=1}^d \int_0^t \partial_i f(B(u),\zeta(u))\,dB_i(u),$$

and

$$\int_0^t \nabla_y f\big(B(u),\zeta(u)\big)\cdot d\zeta(u) = \sum_{i=1}^m \int_0^t \partial_{d+i} f(B(u),\zeta(u))\,d\zeta_i(u).$$

Finally, for the Laplacian in the x-variable we write

$$\Delta_x f = \sum_{j=1}^d \partial_{jj} f\,.$$

THEOREM 7.14 (Multidimensional Itô's formula). *Let $\{B(t) : t \ge 0\}$ be a d-dimensional Brownian motion and suppose $\{\zeta(s)\colon s \ge 0\}$ is a continuous, adapted stochastic process with values in $\mathbb{R}^m$ and increasing components. Let $f\colon \mathbb{R}^{d+m} \to \mathbb{R}$ be such that the partial derivatives $\partial_i f$ and $\partial_{jk} f$ exist for all $1 \le j,\, k \le d$, $1 \le i \le d+m$ and are continuous. If, for some $t > 0$,*

$$\mathbb{E}\int_0^t \big|\nabla_x f\big(B(s),\zeta(s)\big)\big|^2\,ds < \infty,$$

then, almost surely, for all $0 \le s \le t$,

$$\begin{aligned} f\big(B(s),\zeta(s)\big) - f\big(B(0),\zeta(0)\big) &= \int_0^s \nabla_x f\big(B(u),\zeta(u)\big)\cdot dB(u) \\ &+ \int_0^s \nabla_y f\big(B(u),\zeta(u)\big)\cdot d\zeta(u) + \tfrac{1}{2}\int_0^s \Delta_x f\big(B(u),\zeta(u)\big)\,du\,. \end{aligned} \tag{1.6}$$

REMARK 7.15. As the Itô formula holds almost surely simultaneously for all times $s \in [0,t]$, it also holds for stopping times bounded by t. Suppose now that $f\colon U \to \mathbb{R}$ satisfies the differentiability conditions on an open set U, and $K \subset U$ is compact. Then there exists $f^*\colon \mathbb{R}^m \to \mathbb{R}$ with $f^* = f$ on K, which satisfies the conditions of Theorem 7.14. Let T be the first exit time from K. Applying Theorem 7.14 to f^* yields (1.6) for f, almost surely, for all times $s \wedge T$, for $s \le t$. ◇

To appreciate the following discussion, we introduce a localisation of the notion of a martingale.

DEFINITION 7.16. *An adapted stochastic process* $\{X(t)\colon 0 \le t \le T\}$ *is called a* **local martingale** *if there exist stopping times* T_n*, which are almost surely increasing to* T*, such that* $\{X(t\wedge T_n)\colon t \ge 0\}$ *is a martingale, for every* n*.* ◇

The following theorem is a substantial extension of Corollary 2.49.

THEOREM 7.17. *Let* $D \subset \mathbb{R}^d$ *be a domain and* $f\colon D \to \mathbb{R}$ *be harmonic on* D*. Suppose that* $\{B(t)\colon 0 \le t \le T\}$ *is a Brownian motion started inside* D *and stopped at the time* T *when it first exits the domain* D*.*

(a) *The process* $\{f(B(t))\colon 0 \le t \le T\}$ *is a local martingale.*

(b) *If we have*

$$\mathbb{E}\int_0^{t\wedge T} \big|\nabla f(B(s))\big|^2\,ds < \infty \qquad \textit{for all } t > 0,$$

then $\{f(B(t\wedge T))\,:\, t \ge 0\}$ *is a martingale.*

Proof. Suppose that K_n, $n \in \mathbb{N}$, is an increasing sequence of compact sets whose union is D, and let T_n be the associated exit times. By Theorem 7.14 in conjunction with Remark 7.15,

$$f\big(B(t\wedge T_n)\big) = \int_0^{t\wedge T_n} \nabla f\big(B(s)\big)\cdot dB(s)\,,$$

whence $\{f\big(B(t\wedge T_n)\big)\,:\, t \ge 0\}$ is a martingale, which proves (a).

Obviously, almost surely,

$$f\big(B(t\wedge T)\big) = \lim_{n\uparrow\infty} f\big(B(t\wedge T_n)\big)\,. \tag{1.7}$$

For any $t \geq 0$, the process $\{f(B(t \wedge T_n)) : n \in \mathbb{N}\}$ is a discrete-time martingale by the optional stopping theorem. By our integrability assumption,

$$\mathbb{E}\Big[\Big(f\big(B(t \wedge T_n)\big)\Big)^2\Big] = \mathbb{E}\int_0^{T_n \wedge t} \big|\nabla f\big(B(s)\big)\big|^2\, ds \leq \mathbb{E}\int_0^{T \wedge t} \big|\nabla f\big(B(s)\big)\big|^2\, ds < \infty,$$

so that the martingale is L^2-bounded and convergence in (1.7) holds in the L^1-sense. Taking limits in the equation

$$\mathbb{E}\Big[f\big(B(t \wedge T_m)\big)\,\Big|\,\mathcal{F}(s \wedge T_n)\Big] = f\big(B(s \wedge T_n)\big), \qquad \text{for } m \geq n \text{ and } t \geq s\,,$$

first for $m \uparrow \infty$, then $n \uparrow \infty$, gives

$$\mathbb{E}\big[f\big(B(t \wedge T)\big)\,\big|\,\mathcal{F}(s \wedge T)\big] = f\big(B(s \wedge T)\big), \qquad \text{for } t \geq s.$$

This shows that $\{f(B(t \wedge T)) : t \geq 0\}$ is a martingale and completes the proof. ■

Example 7.18. The radially symmetric functions (related to the radial potential),

$$f(x) = \begin{cases} \log|x| & \text{if } d = 2, \\ |x|^{2-d} & \text{if } d \geq 3. \end{cases}$$

are harmonic on the domain $\mathbb{R}^d \setminus \{0\}$. For a d-dimensional Brownian motion $\{B(t) \colon t \geq 0\}$ with $B(0) \neq 0$, the process $\{f(B(t)) \colon t \geq 0\}$ is however not a martingale. Indeed, it is a straightforward calculation to verify that

$$\lim_{t \uparrow \infty} \mathbb{E} \log|B(t)| = \infty, \qquad \text{if } d = 2,$$

and

$$\lim_{t \uparrow \infty} \mathbb{E}[|B(t)|^{2-d}] = 0, \qquad \text{if } d \geq 3,$$

contradicting the martingale property. Hence the integrability condition in Theorem 7.17(b) cannot be dropped without replacement, a local martingale is not necessarily a martingale. ◇

2. Conformal invariance and winding numbers

We now focus on planar Brownian motion $\{B(t) : t \geq 0\}$ and formulate an invariance property which is at the heart of the role of Brownian motion in the context of planar random curves. Throughout this section we use the identification of $\mathbb{R}^2$ and $\mathbb{C}$ and use complex notation when it is convenient.

To motivate the main result suppose that $f \colon \mathbb{C} \to \mathbb{C}$ is **analytic**, i.e. everywhere complex differentiable, and write $f = f_1 + \mathbf{i} f_2$ for the decomposition of f into a real and an imaginary part. Then, by the Cauchy-Riemann equations $\partial_1 f_1 = \partial_2 f_2$ and $\partial_2 f_1 = -\partial_1 f_2$, we have $\Delta f_1 = \Delta f_2 = 0$. Then Itô's formula (if applicable) states that almost surely, for every $t \geq 0$,

$$f\big(B(t)\big) = \int_0^t f'\big(B(s)\big)\, dB(s)\,,$$

where $dB(s)$ is short for $dB_1(s) + \mathbf{i}\, dB_2(s)$ with $B(s) = B_1(s) + \mathbf{i}B_2(s)$. The right hand side defines a continuous process with independent increments, and it is at least plausible that they are Gaussian. Moreover, its expectation vanishes and

$$\mathbb{E}\Big[\Big(\int_0^t f'\big(B(s)\big)\, dB(s)\Big)^2\Big] = \int_0^t \big|f'(B(s))\big|^2\, ds\,,$$

suggesting that $\{f(B(t)) : t \geq 0\}$ is a Brownian motion 'travelling' with the modified speed

$$t \mapsto \int_0^t |f'(B(s))|^2\, ds.$$

To turn this heuristic into a powerful theorem we allow the function to be an analytic map $f\colon U \to V$ between domains in the plane. Recall that such a map is called **conformal** if it is a bijection.

THEOREM 7.19. *Let U be a domain in the complex plane, $x \in U$, and let $f\colon U \to V$ be analytic. Let $\{B(t) : t \geq 0\}$ be a planar Brownian motion started in x and*

$$\tau_U = \inf\big\{t \geq 0 \,:\, B(t) \notin U\big\}$$

its first exit time from the domain U. Then the process $\{f(B(t)) : 0 \leq t \leq \tau_U\}$ is a time-changed Brownian motion, i.e. there exists a planar Brownian motion $\{\widetilde{B}(t) : t \geq 0\}$ such that, for any $t \in [0, \tau_U)$,

$$f(B(t)) = \widetilde{B}(\zeta(t)), \qquad \textit{where} \qquad \zeta(t) = \int_0^t \big|f'(B(s))\big|^2\, ds\,.$$

If, additionally, f is conformal, then $\zeta(\tau_U)$ is the first exit time from V by $\{\widetilde{B}(t) : t \geq 0\}$.

REMARK 7.20. Note that, as f is complex differentiable, the derivative $Df(x)$ is just multiplication by a complex number $f'(x)$, and f can be approximated locally around x by its tangent $z \mapsto f(x) + f'(x)(z - x)$. The derivative of the time change is

$$\partial_t \zeta(t) = |f'(B(t))|^2 = \big(\partial_1 f_1(B(t))\big)^2 + \big(\partial_2 f_1(B(t))\big)^2\,. \qquad \diamond$$

REMARK 7.21. The famous *Riemann mapping theorem* states that for any pair of simply connected open sets $U, V \subsetneq \mathbb{C}$ there exists a conformal mapping $f\colon U \to V$, see [**Ru86**, 14.8]. This ensures that there are plenty of examples for Theorem 7.19. $\diamond$

Proof. Note first that the derivative of f is nonzero except for an at most countable set of points, which does not have a limit point in U. As this set is not hit by Brownian motion, we may remove it from U and note that the resulting set is still open We may therefore assume that f has nonvanishing derivative everywhere on U.

We may also assume, without loss of generality, that f is a mapping between *bounded* domains. Otherwise choose $U_n \subset K_n \subset U$ such that U_n is open with $\bigcup U_n = U$ and K_n is compact, which implies that $V_n = f(U_n)$ is bounded. Then $\{f(B(t)) : t \leq \tau_{U_n}\}$ is a time-changed Brownian motion for all n, and this extends immediately to $\{f(B(t)) : t \leq \tau_U\}$.

The main argument of the proof is based on stochastic integration. Recall that the Cauchy-Riemann equations imply that the vectors ∇f_1 and ∇f_2 are orthogonal and $|\nabla f_1| = |\nabla f_2| = |f'|$. We start by defining for each $t \geq 0$, a stopping time

$$\sigma(t) = \inf\big\{s \geq 0 \,:\, \zeta(s) \geq t\big\},$$

which represents the inverse of the time change. Let $\{\widetilde{B}(t) \,:\, t \geq 0\}$ be a Brownian motion independent of $\{B(t) \,:\, t \geq 0\}$, and define a process $\{W(t) \,:\, t \geq 0\}$ by

$$W(t) = f\big(B(\sigma(t) \wedge \tau_U)\big) + \widetilde{B}(t) - \widetilde{B}(t \wedge \zeta(\tau_U)), \qquad \text{for } t \geq 0.$$

In rough words, at the random time $\zeta(\tau_U)$ an independent Brownian motion is attached at the endpoint of the process $\{f(B(\sigma(t))) \,:\, 0 \leq t \leq \zeta(\tau_U)\}$. Denote by $\mathcal{G}(t)$ the σ-algebra generated by $\{W(s) \,:\, s \leq t\}$. It suffices to prove that the process $\{W(t) \,:\, t \geq 0\}$ is a Brownian motion.

It is obvious that the process is continuous almost surely and hence it suffices to show that its finite dimensional marginal distributions coincide with those of a Brownian motion. Recalling the Laplace transform of the bivariate normal distribution, this is equivalent to showing that, for any $0 \leq s \leq t$ and $\lambda \in \mathbb{C}$,

$$\mathbb{E}\big[e^{\langle \lambda, W(t)\rangle} \,\big|\, \mathcal{G}(s)\big] = \exp\big(\tfrac{1}{2}|\lambda|^2(t-s) + \langle \lambda, W(s)\rangle\big).$$

where we have used $\langle \cdot\,,\,\cdot \rangle$ to denote the scalar product. This follows directly once we show that, for $x \in U$,

$$\mathbb{E}\big[e^{\langle \lambda, W(t)\rangle} \,\big|\, W(s) = f(x)\big] = \exp\big(\tfrac{1}{2}\,|\lambda|^2\,(t-s) + \langle \lambda, f(x)\rangle\big)\,. \tag{2.1}$$

For simplicity of notation we may assume $s = 0$. For the proof we first evaluate the expectation with respect to the independent Brownian motion $\{\widetilde{B}(t) \,:\, t \geq 0\}$ inside, which gives

$$\mathbb{E}\big[e^{\langle \lambda, W(t)\rangle} \,\big|\, W(0) = f(x)\big] = \mathbb{E}_x \exp\big(\langle \lambda, f(B(\sigma(t) \wedge \tau_U))\rangle + \tfrac{1}{2}|\lambda|^2\,\big(t - \zeta(\sigma(t) \wedge \tau_U)\big)\Big).$$

We use the multidimensional Itô's formula for the bounded mapping

$$F(x,u) = \exp\big(\langle \lambda, f(x)\rangle + \tfrac{1}{2}\,|\lambda|^2(t-u)\big)\,,$$

which is defined on $U \times \mathbb{R}$, see Remark 7.15. To prepare this, note that $\partial_{ii} e^g = [\partial_{ii} g + (\partial_i g)^2 e^g]$ and hence

$$\Delta e^g = [\Delta g + |\nabla g|^2]\, e^g\,. \tag{2.2}$$

For $g = \langle \lambda, f\rangle$ we have $\nabla g = \sum_{i=1}^2 \lambda_i \nabla f_i$, which implies $|\nabla g|^2 = |\lambda|^2\,|f'|^2$ as the vectors ∇f_i are orthogonal with norm $|f'|$. Moreover, $\Delta g = 0$ by the analyticity of f. Applying (2.2) gives

$$\Delta\, e^{\langle \lambda, f(x)\rangle} = |\lambda|^2\,|f'(x)|^2\, e^{\langle \lambda, f(x)\rangle}\,.$$

Moreover, we have

$$\partial_u \exp\big(\tfrac{1}{2}\,|\lambda|^2\,(t-u)\big) = -\tfrac{1}{2}\,|\lambda|^2\,\exp\big(\tfrac{1}{2}\,|\lambda|^2\,(t-u)\big)\,.$$

We now let $U_n = \{x \in U \,:\, |x-y| \geq \frac{1}{n} \text{ for all } y \in \partial U\}$. Then $|f'(x)|$ is bounded away from zero on U_n and therefore the stopping time $T = \sigma(t) \wedge \tau_{U_n}$ is bounded. The multidimensional version of Itô's formula gives, almost surely,

$$F\big(B(T),\zeta(T)\big) = F\big(B(0),\zeta(0)\big) + \int_0^T \nabla_x F\big(B(s),\zeta(s)\big) \cdot dB(s)$$
$$+ \int_0^T \partial_u F\big(B(s),\zeta(s)\big)\, d\zeta(s) + \tfrac{1}{2} \int_0^T \Delta_x F\big(B(s),\zeta(s)\big)\, ds\,.$$

Looking back at the two preparatory displays and recalling that $d\zeta(u) = |f'(B(u))|^2\, du$ we see that the two terms in the second line cancel each other. Making use of bounded convergence and the fact that the stochastic integral has zero expectation, see Exercise 7.1, we obtain that

$$\mathbb{E}\big[e^{\langle\lambda,W(t)\rangle}\,\big|\,W(0)=f(x)\big] = \mathbb{E}_x\big[F\big(B(\sigma(t)\wedge\tau_U),\zeta(\sigma(t)\wedge\tau_U)\big)\big]$$
$$= \lim_{n\to\infty} \mathbb{E}_x\big[F\big(B(T),\zeta(T)\big)\big] = F\big(x,0\big) = \exp\Big(\tfrac{1}{2}|\lambda|^2 t + \langle\lambda, f(x)\rangle\Big)\,.$$

This shows (2.1) and completes the proof. ■

As a first application we look at harmonic measure and exploit its conformal invariance in order to calculate it explicitly in a special case.

THEOREM 7.22. *Suppose $U,V \subset \mathbb{R}^2$ are domains and $f\colon \bar U \to \bar V$ is continuous and maps U conformally into V.*

(a) *If $x \in U$, then $\mu_{\partial U}(x,\,\cdot\,)\circ f^{-1} = \mu_{\partial V}(f(x),\,\cdot\,)$.*

(b) *Suppose additionally that $U = K^{\mathrm{c}}$ and $V = L^{\mathrm{c}}$ are the complement of compact sets and $\lim_{x\to\infty} f(x) = \infty$. Then*
$$\mu_K \circ f^{-1} = \mu_L\,.$$

Proof. (a) follows from Theorem 7.19 together with the continuity of f on $\bar U$, which ensures that the first hitting point of ∂U by a Brownian motion is mapped onto the first hitting point of ∂V by its conformal image. For (b) tahe the limit $x \to \infty$ and recall Theorem 3.45. ■

EXAMPLE 7.23. We find the harmonic measure from infinity on the unit interval

$$[0,1] = \big\{x + \mathbf{i}y\colon y = 0, 0 \le x \le 1\big\}.$$

Starting point is the harmonic measure on the circle $\partial\mathcal{B}(0,1)$, which we know is the uniform distribution ϖ. Let U be the complement of the unit ball $\mathcal{B}(0,1)$ and V the complement of the interval $[-,1,1]$, and take the conformal mapping

$$f\colon U \to V, \quad f(z) = \frac{1}{2}\Big(z + \frac{1}{z}\Big),$$

which satisfies our conditions. Hence $\varpi\circ f^{-1}$ is the harmonic measure on $[-1,1]$. If $z = x + \mathbf{i}y = \cos\theta + \mathbf{i}\sin\theta \in \partial\mathcal{B}(0,1)$, then $|f'(z)|^2 = \sin^2\theta$, and hence $|f'(z)| = |y| = \sqrt{1-x^2}$. Recalling that every $x \in [-1,1]$ has two preimages, we get that the density of $\varpi \circ f^{-1}$ at $x = \cos\theta$ is

$$\frac{2}{2\pi|f'(e^{\mathbf{i}\theta})|} = \frac{1}{\pi}\,\frac{1}{\sqrt{1-x^2}}.$$

Mapping V via $z \mapsto z^2$ onto the complement of $[0,1]$, noting that $|f'(z)| = 2|z|$ and that again we have two preimages, we obtain that the harmonic measure on $[0,1]$ is

$$d\mu_{[0,1]}(x) = \frac{1}{\pi} \frac{1}{\sqrt{x(1-x)}}\, dx,$$

which is the Beta$(\frac{1}{2}, \frac{1}{2})$ distribution. ◇

As a further important application of conformal invariance we calculate the probability that a planar Brownian motion exits a cone before leaving a disc, see Figure 1.

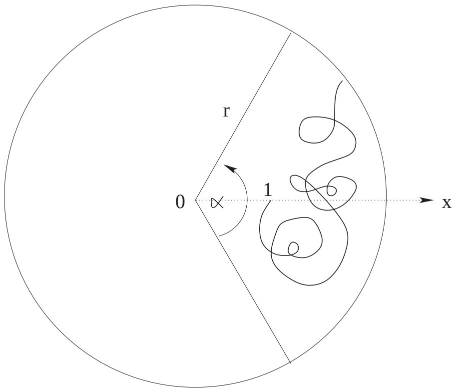

FIGURE 1. The Brownian path does not exit the cone before leaving the disc.

THEOREM 7.24. *Let $\alpha \in (0, 2\pi]$ and denote by $W[\alpha]$ an open cone with vertex in the origin, symmetric about the x-axis, with opening angle α. Let $\{B(t) : t \geq 0\}$ be planar Brownian motion started in $x = (1,0)$, and denote $T(r) = \inf\{t \geq 0 : |B(t)| = r\}$. Then, for $r > 1$,*

$$\mathbb{P}\big\{B[0,T(r)] \subset W[\alpha]\big\} = \frac{2}{\pi} \arctan\Big(\frac{2r^{\frac{\pi}{\alpha}}}{r^{\frac{2\pi}{\alpha}} - 1}\Big).$$

Proof. For ease of notation we identify $\mathbb{R}^2$ with the complex plane. In the first step we use the conformal map $f\colon W[\alpha] \to W[\pi]$ defined by $f(x) = x^{\pi/\alpha}$ to map the cone onto a halfspace. Let $B^* = f \circ B$, which by conformal invariance is a time-changed Brownian motion started in the point $B^*(0) = 1$. We thus have that

$$\big\{B[0,T(r)] \subset W[\alpha]\big\} = \big\{B^*[0, T(r^{\pi/\alpha})] \subset W[\pi]\big\}.$$

It therefore suffices to show the result in the case $\alpha = \pi$. So let $\{B(t) : t \geq 0\}$ be a Brownian motion started in $B(0) = 1$ and look at the stopping time $S = \min\{t \geq 0 : \mathfrak{Re}(B(t)) \leq 0\}$. We use reflection on the imaginary axis, i.e. for $f(x,y) = (-x,y)$ we let

$$\widetilde{B}(t) = \begin{cases} B(t) & \text{if } t \leq S, \\ f(B(t)) & \text{if } t \geq S. \end{cases}$$

Then $\widetilde{B}$ is a Brownian motion started in $\widetilde{B}(0) = 1$ and, for $\widetilde{T}(r) = \inf\{t \geq 0 : |\widetilde{B}(t)| = r\}$,

$$\begin{aligned} \mathbb{P}\{\mathfrak{Re}(B(T(r))) > 0\} &= \mathbb{P}\{\mathfrak{Re}(B(T(r))) > 0, T(r) < S\} + \mathbb{P}\{\mathfrak{Re}(B(T(r))) > 0, T(r) > S\} \\ &= \mathbb{P}\{T(r) < S\} + \mathbb{P}\{\mathfrak{Re}(\widetilde{B}(\widetilde{T}(r))) < 0\}. \end{aligned}$$

As $\{T(r) < S\}$ is the event whose probability we need to bound, it just remains to find

$$\mathbb{P}\{\mathfrak{Re}(B(T(r))) > 0\} - \mathbb{P}\{\mathfrak{Re}(B(T(r))) < 0\}.$$

By Brownian scaling we may assume that the Brownian motion is started at $B(0) = 1/r$ and $T = \min\{t \geq 0 : |B(t)| = 1\}$. We apply the conformal map

$$f \colon \mathcal{B}(0,1) \to \mathcal{B}(0,1), \qquad f(z) = \frac{z - 1/r}{1 - z/r},$$

which is a Möbius transformation mapping the starting point of the Brownian motion to the origin and fixing the point 1. As this maps the segment $\{z \in \partial\mathcal{B}(0,1) : \mathfrak{Re}(z) < 0\}$ onto a segment of length $2\arctan\frac{r^2-1}{2r}$ we obtain the result. ■

The next result represents planar Brownian motion in polar coordinates. Again we identify $\mathbb{R}^2$ with the complex plane.

Theorem 7.25 (Skew-product representation). *Suppose $\{B(t) : t \geq 0\}$ is a planar Brownian motion with $B(0) = 1$. Then there exist two independent linear Brownian motions $\{W_i(t) : t \geq 0\}$, for $i = 1, 2$, such that*

$$B(t) = \exp\big(W_1(H(t)) + \mathbf{i}\, W_2(H(t))\big), \text{ for all } t \geq 0,$$

where

$$H(t) = \int_0^t \frac{ds}{|B(s)|^2} = \inf\Big\{u \geq 0 : \int_0^u \exp(2W_1(s))\, ds > t\Big\}.$$

Remark 7.26. By the result, both the logarithm of the radius, and the continuous determination of the angle of a planar Brownian motion are time-changed Brownian motions. The time-change itself depends only on the radius of the motion and ensures that the angle changes slowly away from the origin, but rapidly near the origin. ◇

Proof. Note first that $H(t)$ itself is well-defined by Corollary 2.23. Moreover, the claimed equality for $H(t)$ follows easily from the fact that both sides have the same value at $t = 0$ and the same derivative. Let $\{W(t) : t \geq 0\}$ be planar Brownian motion and $W(t) = W_1(t) + \mathbf{i}\, W_2(t)$ its decomposition into real and imaginary part. By Theorem 7.19,

$$\exp\big(W(t)\big) = B(\zeta(t)), \tag{2.3}$$

where $\{B(t) : t \geq 0\}$ is a planar Brownian motion and

$$\zeta(t) = \int_0^t \exp(2W_1(s))\, ds.$$

By definition H is the inverse function of ζ. Hence, using (2.3) for $t = H(s)$, we get

$$B(s) = \exp\big(W(H(s))\big) = \exp\big(W_1(H(s)) + \mathbf{i}\, W_2(H(s))\big),$$

which is the desired result. ■

Example 7.27. By the skew-product representation, for a planar Brownian motion $\{B(t) : t \geq 0\}$, we have $\log |B(t)| = W_1(H(t))$ and hence the process $\{\log |B(t)| : t \geq 0\}$ is a time-changed Brownian motion in dimension one. However, recall from Example 7.18 that it is *not* a martingale. $\diamond$

For further applications, we need to study the asymptotics of the random clock $H(t)$ more carefully. To state the next result let $\{W_1(t) : t \geq 0\}$ be a linear Brownian motion as in Theorem 7.25 and, for $a > 0$, let $\{W_1^a(t) : t \geq 0\}$ be the Brownian motion given by $W_1^a(t) = a^{-1} W_1(a^2 t)$. For each such Brownian motion we look at the first hitting time of level b,

$$T_b^a := \inf \big\{ t \geq 0 : W_1^a(t) = b \big\}.$$

Theorem 7.28. *For every $\varepsilon > 0$ we have*

$$\lim_{t \to \infty} \mathbb{P}\Big\{ \Big| \frac{4H(t)}{(\log t)^2} - T_1^{\frac12 \log t} \Big| > \varepsilon \Big\} = 0.$$

The proof uses the following simple fact, sometimes known as *Laplace's method.*

Lemma 7.29. *For any continuous $f\colon [0,t] \to \mathbb{R}$ and $t > 0$,*

$$\lim_{a \uparrow \infty} \frac{1}{a} \log \int_0^t \exp(a f(v))\, dv = \max_{0 \leq s \leq t} f(s).$$

Proof. The upper bound is obvious, by replacing f by its maximum. For the lower bound, let $s \in [0,t]$ be a point where the maximum is taken. We use continuity to find, for any $\varepsilon > 0$, some $0 < \delta < 1$ such that $f(r) \geq f(s) - \varepsilon$ for all $r \in (s-\delta, s+\delta)$. Restricting the limit to this interval gives a lower bound of $\max_{0 \leq s \leq t} f(s) - \varepsilon$, and the result follows as $\varepsilon > 0$ was arbitrary. ∎

Proof of Theorem 7.28. By scaling one may assume that $W_1(0) = 0$. We abbreviate $a = a(t) = \frac12 \log t$. As we have, for any $\delta > 0$,

$$\lim_{\varepsilon \downarrow 0} \mathbb{P}\Big\{ T_{1+\varepsilon}^{\frac12 \log t} - T_{1-\varepsilon}^{\frac12 \log t} > \delta \Big\} = \lim_{\varepsilon \downarrow 0} \mathbb{P}\Big\{ T_{1+\varepsilon}^{1} - T_{1-\varepsilon}^{1} > \delta \Big\} = 0$$

it suffices to show that

$$\lim_{t \uparrow \infty} \mathbb{P}\Big\{ \frac{4H(t)}{(\log t)^2} > T_{1+\varepsilon}^{\frac12 \log t} \Big\} = 0, \quad \text{and} \quad \lim_{t \uparrow \infty} \mathbb{P}\Big\{ \frac{4H(t)}{(\log t)^2} < T_{1-\varepsilon}^{\frac12 \log t} \Big\} = 0.$$

We first show that

$$\lim_{t \uparrow \infty} \mathbb{P}\Big\{ \frac{4H(t)}{(\log t)^2} > T_{1+\varepsilon}^{\frac12 \log t} \Big\} = 0. \tag{2.4}$$

We have

$$\begin{aligned}
\Big\{ \frac{4H(t)}{(\log t)^2} > T_{1+\varepsilon}^{\frac12 \log t} \Big\} &= \Big\{ \int_0^{a^2 T_{1+\varepsilon}^a} \exp(2W_1(u))\, du < t \Big\} \\
&= \Big\{ \frac{1}{2a} \log \int_0^{a^2 T_{1+\varepsilon}^a} \exp(2W_1(u))\, du < 1 \Big\},
\end{aligned}$$

recalling that $a = \frac{1}{2}\log t$. Note now that
$$\frac{1}{2a}\log\int_0^{a^2T^a_{1+\varepsilon}} \exp(2W_1(u))\,du = \frac{\log a}{a} + \frac{1}{2a}\log\int_0^{T^a_{1+\varepsilon}} \exp(2aW^a_1(u))\,du,$$
and the right hand side has the same distribution as
$$\frac{\log a}{a} + \frac{1}{2a}\log\int_0^{T^1_{1+\varepsilon}} \exp(2aW_1(u))\,du.$$
Laplace's method gives that, almost surely,
$$\lim_{a\uparrow\infty}\frac{1}{2a}\log\int_0^{T^1_{1+\varepsilon}} \exp(2aW_1(u))\,du = \sup_{0\le s\le T^1_{1+\varepsilon}} W_1(s) = 1+\varepsilon.$$
Hence,
$$\lim_{a\uparrow\infty}\mathbb{P}\Big\{\Big|\frac{\log a}{a} + \frac{1}{2a}\log\int_0^{T^1_{1+\varepsilon}} \exp(2aW_1(u))\,du - (1+\varepsilon)\Big| > \varepsilon\Big\} = 0.$$
This proves (2.4). In the same way one can show that
$$\lim_{t\uparrow\infty}\mathbb{P}\Big\{\frac{4H(t)}{(\log t)^2} > T^{\frac{1}{2}\log t}_{1-\varepsilon}\Big\} = 0,$$
and this completes the proof. ■

Remark 7.30. As $\{W^a_1(t) : t\ge 0\}$ is a Brownian motion for every $a>0$, the law of T^a_1 does not depend on $a>0$. Therefore, Theorem 7.28 implies that
$$\frac{4H(t)}{(\log t)^2} \Longrightarrow T_1,$$
where $T_1 = \inf\{s\ge 0 : W(s) = 1\}$. The distribution of T_1 is, by Theorem 2.32 given by the density $(2\pi s^3)^{-1/2}\exp(-1/(2s))$. ◇

We now determine the asymptotic law of the winding numbers $\theta(t) = W_2(H(t))$, as $t\to\infty$.

Theorem 7.31 (Spitzer's law). *For any $x\in\mathbb{R}$,*
$$\lim_{t\to\infty}\mathbb{P}\Big\{\frac{2}{\log t}\,\theta(t)\le x\Big\} = \int_{-\infty}^x \frac{dy}{\pi(1+y^2)}.$$
In other words, the law of $2\theta(t)/\log t$ converges to a standard symmetric Cauchy distribution.

Proof. We define $\{W^a_2(t) : t\ge 0\}$ by $W^a_2(t) = (1/a)W_2(a^2t)$. Then,
$$a^{-1}\theta(t) = a^{-1}W_2(H(t)) = W^a_2(a^{-2}H(t)).$$
Hence, by Theorem 7.28, for $a = a(t) = \frac{1}{2}\log t$,
$$\lim_{t\to\infty}\mathbb{P}\Big\{\Big|\frac{2\theta(t)}{\log t} - W^a_2\big(T^a_1\big)\Big| > \varepsilon\Big\} = 0.$$

The law of the random variable $W_2^a(T_1^a)$ does not depend on the choice of a. By Theorem 2.33, see also Exercise 7.4, it is Cauchy distributed. ■

3. Tanaka's formula and Brownian local time

In this section we establish a deep connection between Itô's formula and Brownian local times for linear Brownian motion $\{B(t)\colon t \geq 0\}$. The basic idea is to give an analogue of Itô's formula for the function $f\colon \mathbb{R} \to \mathbb{R}$, $f(t) = |\, t - a|$. Note that this function is not twice continuously differentiable, so Itô's formula cannot be applied directly.

To see what we are aiming at, let's apply Itô's formula informally. We have in the distributional sense that $f'(x) = \mathrm{sign}(x - a)$ and $f''(x) = 2\delta_a$. Hence Itô's formula would give

$$|B(t) - a| - |B(0) - a| = \int_0^t \mathrm{sign}(B(s) - a)\, dB(s) + \int_0^t \delta_a(B(s))\, ds,$$

The last integral can be interpreted as the time spent by Brownian motion at level a and hence it is natural to expect that it is the local time $L^a(t)$. Tanaka's formula confirms this intuition.

Theorem 7.32 (Tanaka's formula). *Let $\{B(t) : t \geq 0\}$ be linear Brownian motion. Then, for every $a \in \mathbb{R}$, almost surely, for all $t > 0$,*

$$|B(t) - a| - |B(0) - a| = \int_0^t \mathrm{sign}(B(s) - a)\, dB(s) + L^a(t),$$

where $\mathrm{sign}\, x = \mathbb{1}_{\{x>0\}} - \mathbb{1}_{\{x\leq 0\}}$.

Remark 7.33. To explore the relation of Tanaka's and Itô's formula further, suppose that $f\colon \mathbb{R} \to \mathbb{R}$ is twice differentiable such that f' has compact support, but do not assume that f'' is continuous. Then, for a suitable constant c,

$$f'(x) = \tfrac{1}{2} \int \mathrm{sign}(x - a)\, f''(a)\, da + c\,,$$

and, for a suitable constant b,

$$f(x) = \tfrac{1}{2} \int |x - a|\, f''(a)\, da + cx + b\,.$$

Integrating Tanaka's formula with respect to $\frac{1}{2} f''(a)\, da$ and exchanging this integral with the stochastic integral, which is justified by Exercise 7.7 below, gives

$$f\big(B(t)\big) - f\big(B(0)\big) = \int_0^t f'\big(B(s)\big)\, dB(s) + \tfrac{1}{2} \int L^a(t)\, f''(a)\, da\,.$$

By Theorem 6.17 the last term equals $\frac{1}{2}\int_0^t f''\big(B(s)\big)\, ds$. Hence, we learn that Itô's formula does not require the continuity requirement for the second derivative. ◇

For the proof of the Tanaka formula we define

$$\tilde{L}^a(t) := |B(t) - a| - |B(0) - a| - \int_0^t \operatorname{sign}(B(s) - a)\, dB(s)$$

and show that this defines a density of the occupation measure.

LEMMA 7.34. *For every $t \geq 0$ and $a \in \mathbb{R}$, almost surely,*

$$\tilde{L}^a(t) = \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \int_0^t \mathbb{1}_{(a,a+\varepsilon)}(B(s))\, ds, \qquad \textit{in probability.}$$

Proof. Using the strong Markov property the statement can be reduced to the case $a = 0$. The main idea of the proof is now to use convolution to make $|\,x|$ smooth, and then use Itô's formula for the smooth function.

For this purpose, recall that, for any $\delta > 0$ we can find smooth functions $g, h\colon \mathbb{R} \to [0,1]$ with compact support such that $g \leq \mathbb{1}_{(0,1)} \leq h$ and $\int g = 1 - \delta$, $\int h = 1 + \delta$. This reduces the problem to showing that

$$\tilde{L}^0(t) = \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \int_0^t g\big(\varepsilon^{-1} B(s)\big)\, ds, \qquad \text{in probability,}$$

for $g\colon \mathbb{R} \to [0,1]$ smooth, with compact support in $[0, \infty)$ and $\int g = 1$. Let

$$f_\varepsilon(x) = \varepsilon^{-1} \int |\,x - a|\, g(\varepsilon^{-1} a)\, da = \int |\,x - \varepsilon a|\, g(a)\, da.$$

f_ε is smooth and $f_\varepsilon'(x) = \int \operatorname{sign}(x - \varepsilon a) g(a)\, da$, $f_\varepsilon''(x) = 2\varepsilon^{-1} g(\varepsilon^{-1} x)$. Itô's formula gives

$$f_\varepsilon(B(t)) - f_\varepsilon(B(0)) - \int_0^t f_\varepsilon'(B(s))\, dB(s) = \varepsilon^{-1} \int_0^t g\big(\varepsilon^{-1} B(s)\big)\, ds. \tag{3.1}$$

Now we let $\varepsilon \downarrow 0$ for each term. The sequence of probability measures $\varepsilon^{-1} g(\varepsilon^{-1} z)\, dz$ converges weakly to δ_0, which implies that $f_\varepsilon(x) \to |\,x|$ for all x. From the definition of f_ε we infer that all functions f_ε are Lipschitz with Lipschitz constant one. Hence, given $\delta > 0$ and a compact interval $I = [a, b]$, we can find finitely many points $a = x_1 < \ldots < x_n = b$ with $x_{k+1} - x_k < \delta/3$. There exists $\varepsilon_0 > 0$ such that $|f_\varepsilon(x_k) - |\,x_k|| < \delta/3$ for all $0 < \varepsilon < \varepsilon_0$. Then, for all $x \in I$ there is $x_k \leq x \leq x_{k+1}$ and

$$|f_\varepsilon(x) - |\,x|\,| \leq |f_\varepsilon(x) - f_\varepsilon(x_k)| + |f_\varepsilon(x_k) - |\,x_k|\,| + ||\,x_k| - |\,x|\,| \leq \delta.$$

In other words, $f_\varepsilon(x) \to |\,x|$ uniformly on compact intervals. Given $\delta > 0$ one can find M such that $|B(s)| \leq M$ on $[0, t]$ with probability exceeding $1 - \delta$. On this event we have

$$f_\varepsilon(B(s)) \to |B(s)| \text{ uniformly in } s \in [0, t].$$

This ensures convergence in probability of the first two terms on the left hand side of (3.1). To deal with the third term, we differentiate f_ε and get

$$f_\varepsilon'(x) = \int \operatorname{sign}\big(x - \varepsilon a\big)\, g(a)\, da \uparrow \operatorname{sign}(x) \text{ as } \varepsilon \downarrow 0.$$

Now we use the isometry property (1.2) to infer that

$$\mathbb{E}\Big[\Big(\int_0^t \operatorname{sign}(B(s))\, dB(s) - \int_0^t f_\varepsilon'(B(s))\, dB(s)\Big)^2\Big] \;=\; \mathbb{E} \int_0^t (\operatorname{sign}(B(s)) - f_\varepsilon'(B(s)))^2\, ds.$$

The right hand side converges to zero, by the bounded convergence theorem. Hence we have shown that, in probability,

$$\begin{aligned}\lim_{\varepsilon\downarrow 0}\varepsilon^{-1}\int_0^t g\big(\varepsilon^{-1}B(s)\big)\,ds &= \lim_{\varepsilon\downarrow 0} f_\varepsilon(B(t)) - f_\varepsilon(B(0)) - \int_0^t f_\varepsilon'(B(s))\,dB(s)\\ &= |B(t)| - |B(0)| - \int_0^t \operatorname{sign}(B(s))\,dB(s) = \tilde{L}^0(t).\end{aligned}$$

■

Proof of Theorem 7.32. Convergence in probability implies that a subsequence converges almost surely and hence, for every t, we obtain from Lemma 7.34 that, almost surely, the process $\{\tilde{L}^a(t)\colon a \in \mathbb{R}\}$ is a density of the occupation measure. From Theorem 6.17 we therefore get that $\tilde{L}^a(t) = L^a(t)$ for almost every $a \in \mathbb{R}$. Averaging over t gives that, almost surely,

$$\tilde{L}^a(t) = L^a(t) \qquad \text{for almost every } a \in \mathbb{R} \text{ and } t \geq 0.$$

As the random field $\{L^a(t)\colon t \geq 0, a \in \mathbb{R}\}$ is continuous by Theorem 6.18, it therefore is a continuous modification of $\{\tilde{L}^a(t)\colon t \geq 0, a \in \mathbb{R}\}$. In particular, for every a, almost surely, the process $\{L^a(t)\colon t \geq 0\}$ agrees with $\{\tilde{L}^a(t)\colon t \geq 0\}$. ■

Corollary 7.35. *For every $a \in \mathbb{R}$, almost surely, for all $t \geq 0$,*

$$\tfrac{1}{2}L^a(t) = (B(t)-a)^+ - (B(0)-a)^+ - \int_0^t \mathbb{1}_{\{B(s)>a\}}\,dB(s),$$

and

$$\tfrac{1}{2}L^a(t) = (B(t)-a)^- - (B(0)-a)^- + \int_0^t \mathbb{1}_{\{B(s)\leq a\}}\,dB(s).$$

Proof. The right sides in these formulas add up to $L^a(t)$, while their difference is zero. ■

We now use Tanka's formula to prove Lévy's theorem describing the joint law of the modulus and local time of a Brownian motion.

Theorem 7.36 (Lévy). *The processes*

$$\{\,(|B(t)|, L^0(t))\,:\, t \geq 0\} \qquad \textit{and} \qquad \{\,(M(t)-B(t), M(t))\,:\, t \geq 0\}$$

have the same distribution.

Remark 7.37. This result extends both Theorem 2.31 where it was shown that the processes $\{|B(t)|\,:\, t \geq 0\}$ and $\{M(t)-B(t)\,:\, t \geq 0\}$ have the same distribution, and Theorem 6.10 where it was shown that $\{L^0(t)\,:\, t \geq 0\}$ and $\{M(t)\,:\, t \geq 0\}$ have the same distribution. Exercise 6.2 suggests an alternative proof using random walk methods. ◇

As a preparation for the proof we find the law of the process given by integrating the sign of a Brownian motion with respect to that Brownian motion.

Lemma 7.38. *For every $a \in \mathbb{R}$, the process $\{W(t) : t \geq 0\}$ given by*

$$W(t) = \int_0^t \operatorname{sign}(B(s) - a)\, dB(s)$$

is a standard Brownian motion.

Proof. Assume, without loss of generality, that $a < 0$. Suppose that $T = \inf\{t > 0 : B(t) = a\}$. Then $W(t) = B(t)$ for all $t \leq T$ and hence $\{W(t) : 0 \leq t \leq T\}$ is a (stopped) Brownian motion. By the strong Markov property the process $\{\widetilde{B}(t) : t \geq 0\}$ given by $\widetilde{B}(t) = B(t+T) - a$ is a Brownian motion started in the origin, which is independent of $\{W(t) : 0 \leq t \leq T\}$. As

$$W(t+T) = W(T) + \int_T^{t+T} \operatorname{sign}(B(s) - a)\, dB(s) = B(T) + \int_0^t \operatorname{sign}(\widetilde{B}(s))\, d\widetilde{B}(s),$$

it suffices to show that the second term is a Brownian motion to complete the proof. Hence we may henceforth assume that $a = 0$.

Now fix $0 \leq s < t$ and recall that $W(t) - W(s)$ is independent of $\mathcal{F}(s)$. For the proof it hence suffices to show that $W(t) - W(s)$ has a centred normal distribution with variance $t - s$. Choose $s = t_1^{(n)} < \ldots < t_n^{(n)} = t$ with mesh $\Delta(n) \downarrow 0$, and approximate the progressively measurable process $H(u) = \operatorname{sign}(B(u))$ by the step processes

$$H_n(u) = \operatorname{sign}\big(B(t_j^{(n)})\big) \qquad \text{if } t_j^{(n)} < u \leq t_{j+1}^{(n)}.$$

It follows from the fact that the zero set of Brownian motion is a closed set of measure zero, that $\lim \mathbb{E} \int_s^t (H_n(u) - H(u))^2\, du = 0$, and hence

$$\begin{aligned} W(t) - W(s) &= \int_s^t H(u)\, dB(u) = \mathrm{L}^2 - \lim_{n\to\infty} \int_s^t H_n(u)\, dB(u) \\ &= \mathrm{L}^2 - \lim_{n\to\infty} \sum_{j=1}^{n-1} \operatorname{sign}\big(B(t_j^{(n)})\big)\, \big(B(t_{j+1}^{(n)}) - B(t_j^{(n)})\big). \end{aligned}$$

From the independence of the Brownian increments and elementary properties of the normal distribution, one can see that the random variables on the right all have a centred normal distribution with variance $t - s$. Hence this also applies to the limit $W(t) - W(s)$. ■

Proof of Theorem 7.36. By Tanaka's formula we have

$$|B(t)| = \int_0^t \operatorname{sign}(B(s))\, dB(s) + L^0(t) = W(t) + L^0(t)\,.$$

Define a standard Brownian motion $\{\widetilde{W}(t) : t \geq 0\}$ by

$$\widetilde{W}(t) = -W(t) \qquad \text{for all } t \geq 0,$$

and let $\{\widetilde{M}(t) : t \geq 0\}$ be the associated maximum process. We show that

$$\widetilde{M}(t) = L^0(t) \qquad \text{for all } t \geq 0,$$

which implies that $\{(|B(t)|, L^0(t)) : t \geq 0\}$ and $\{(\widetilde{M}(t) - \widetilde{W}(t), \widetilde{M}(t)) : t \geq 0\}$ agree pointwise, and the result follows as the latter process agrees in distribution with

$$\{(M(t) - B(t), M(t)) : t \geq 0\}.$$

To show that $\widetilde{M}(t) = L^0(t)$ we first note that

$$\widetilde{W}(s) = L^0(s) - |B(s)| \leq L^0(s),$$

and hence, taking the maximum over all $s \leq t$, we get $\widetilde{M}(t) \leq L^0(t)$. On the other hand, the process $\{L^0(t) : t \geq 0\}$ increases only on the set $\{t : B(t) = 0\}$ and on this set we have $L^0(t) = \widetilde{W}(t) \leq \widetilde{M}(t)$. Hence the proof is complete, since $\{\widetilde{M}(t) : t \geq 0\}$ is increasing. ■

4. Feynman-Kac formulas and applications

In this section we answer some natural questions about Brownian motion that involve time. For example, we find the probability that linear Brownian motion exits a given interval by a fixed time. Our main tool is the close relationship between the expectation of certain functionals of the Brownian path and the heat equation with dissipation term. This goes under the name of *Feynman-Kac formula*, and the theorems that make up this theory establish a strong link between parabolic partial differential equations and Brownian motion.

DEFINITION 7.39. *Let $U \subset \mathbb{R}^d$ be either open and bounded, or $U = \mathbb{R}^d$. A twice differentiable function $u\colon (0,\infty) \times U \to [0,\infty)$ is said to solve the* **heat equation with heat dissipation rate** *$V\colon U \to \mathbb{R}$ and initial condition $f\colon U \to [0,\infty)$ on U if we have*

- $\lim\limits_{\substack{x \to x_0 \\ t \downarrow 0}} u(t,x) = f(x_0)$, *whenever $x_0 \in U$,*
- $\lim\limits_{\substack{x \to x_0 \\ t \to t_0}} u(t,x) = 0$, *whenever $x_0 \in \partial U$,*
- $\partial_t u(t,x) = \frac{1}{2}\Delta_x u(t,x) + V(x)u(t,x)$ *on $(0,\infty) \times U$,*

where the Laplacian Δ_x acts on the space variables x. ◇

REMARK 7.40. The solution $u(t,x)$ describes the temperature at time t at x for a heat flow with *cooling* with rate $-V(x)$ on the set $\{x \in U\colon V(x) < 0\}$, and *heating* with rate $V(x)$ on the set $\{x \in U\colon V(x) > 0\}$, where the initial temperature distribution is given by $f(x)$ and the boundary of U is kept at zero temperature. ◇

Instead of going for the most general results linking the heat equation to Brownian motion, we give some of the more basic forms of the Feynman-Kac formula together with applications. Our first theorem in this spirit, an existence result for the heat equation in the case $U = \mathbb{R}^d$, will lead to a new, more analytic proof of the second arcsine law, Theorem 5.28.

Theorem 7.41. *Suppose $V\colon \mathbb{R}^d \to \mathbb{R}$ is bounded. Then $u\colon [0,\infty)\times\mathbb{R}^d \to \mathbb{R}$ defined by*

$$u(t,x) = \mathbb{E}_x\Big\{ \exp\Big(\int_0^t V\big(B(r)\big)\, dr\Big)\Big\},$$

solves the heat equation on $\mathbb{R}^d$ with dissipation rate V and initial condition one.

Proof. The easiest proof is by a direct calculation. Expand the exponential in a power series, then the terms in the expansion are $a_0(x) := 1$ and, for $n \geq 1$,

$$\begin{aligned}
a_n(x) &:= \frac{1}{n!}\,\mathbb{E}_x\Big[\int_0^t \cdots \int_0^t V(B(t_1))\cdots V(B(t_n))dt_1\ldots dt_n\Big] \\
&= \mathbb{E}_x\Big[\int_0^t dt_1 \int_{t_1}^t dt_2 \cdots \int_{t_{n-1}}^t dt_n\, V(B(t_1))\cdots V(B(t_n))\Big] \\
&= \int_0^t dt_1 \cdots \int_{t_{n-1}}^t dt_n \int dx_1 \cdots \int dx_{n+1} \prod_{i=1}^n V(x_i) \prod_{i=1}^{n+1} \mathfrak{p}(t_i - t_{i-1}, x_{i-1}, x_i)\,,
\end{aligned}$$

with the conventions $x_0 = x$, $t_0 = 0$ and $t_{n+1} = t$. Differentiating with respect to t and using $\partial_t \mathfrak{p}(t, x_1, x_2) = \frac{1}{2}\Delta_x \mathfrak{p}(t, x_1, x_2)$, we get

$$\begin{aligned}
\partial_t a_1(x) &= \partial_t \int_0^t dt_1 \int dx_1\, V(x_1) \int dx_2\, \mathfrak{p}(t_1, x, x_1)\, \mathfrak{p}(t-t_1, x_1, x_2) \\
&= \int_0^t dt_1 \int dx_1\, V(x_1) \int dx_2\, \mathfrak{p}(t_1, x_1, x)\, \tfrac{1}{2}\Delta_x \mathfrak{p}(t-t_1, x_2, x_1) \\
&\qquad\qquad + \int dx_1\, V(x)\, \mathfrak{p}(t, x, x_1) \\
&= \tfrac{1}{2}\Delta_x\, a_1(x) + V(x) a_0(x)\,.
\end{aligned}$$

Analogously, $\frac{\partial}{\partial t}\, a_n(x) = \frac{1}{2}\Delta_x\, a_n(x) + V(x)a_{n-1}(x)$. Adding up all these terms, and noting that differentiation under the summation sign is allowed, verifies the validity of the differential equation. The requirement on the initial condition follows easily from the boundedness of V. ■

As an application we give a proof of the second arcsine law, Theorem 5.28, which does not rely on the first arcsine law. We use Theorem 7.41 with $V(x) = \lambda\, \mathbb{1}_{[0,\infty)}(x)$. Then

$$u(t,x) := \mathbb{E}_x\Big[\exp\Big(-\lambda \int_0^t \mathbb{1}_{[0,\infty)}(B(s))\, ds\Big)\Big]$$

solves

$$\partial_t u(t,x) = \tfrac{1}{2}\, \partial_{xx} u(t,x) - \lambda\, \mathbb{1}_{[0,\infty)}(x)\, u(t,x)\,, \qquad u(0,x) = 1 \text{ for all } x \in \mathbb{R}.$$

To turn this partial differential equation into an ordinary differential equations, we take the Laplace transform

$$g(x) = \int_0^\infty u(t,x)\, e^{-\rho t}\, dt\,,$$

which satisfies the equation

$$\rho\, g(x) + \lambda V(x)\, g(x) - \tfrac{1}{2}\, g''(x) = 1\,.$$

This can be rewritten as

$$(\rho+\lambda)\, g(x) - \tfrac{1}{2}\, g''(x) = 1 \text{ if } x > 0\,,$$
$$\rho\, g(x) - \tfrac{1}{2}\, g''(x) = 1 \text{ if } x < 0\,.$$

Solving these two linear ordinary differential equations gives

$$g(x) = \tfrac{1}{\lambda+\rho} + A\, e^{\sqrt{2\rho}x} + Be^{-\sqrt{2\rho}x} \text{ if } x > 0\,,$$
$$g(x) = \tfrac{1}{\rho} + C\, e^{\sqrt{2\rho}x} + De^{-\sqrt{2\rho}x} \text{ if } x < 0\,.$$

As g must remain bounded as $\rho \uparrow \infty$, we must have $A = D = 0$. Moreover, g must be continuously differentiable in zero, hence C and B can be calculated from matching conditions. After an elementary calculation we obtain

$$g(0) = \frac{1}{\sqrt{\rho(\rho+\lambda)}}\,.$$

On the other hand, with

$$X(t) = \frac{1}{t}\int_0^t \mathbb{1}_{[0,\infty)}(B(s))\, ds$$

we have, using Brownian scaling in the second step,

$$\begin{aligned} g(0) &= \mathbb{E}_0\Big[\int_0^\infty \exp\big(-\rho t - \lambda t X(t)\big)\, dt\Big] \\ &= \mathbb{E}_0\Big[\int_0^\infty \exp\big(-\rho t - \lambda t X(1)\big)\, dt\Big] = \mathbb{E}_0\Big[\frac{1}{\rho + \lambda X(1)}\Big]\,. \end{aligned}$$

Now we let $\rho = 1$ and from

$$\mathbb{E}_0\Big[\frac{1}{1+\lambda X(1)}\Big] = \frac{1}{\sqrt{1+\lambda}}$$

and the expansions

$$\frac{1}{\sqrt{1+\lambda}} = \sum_{n=0}^\infty (-\lambda)^n\, \frac{\frac{1}{2}\frac{3}{2}\cdots\frac{2n-1}{2}}{n!}\,,$$

and

$$\int_0^1 x^{n-\frac{1}{2}}\,(1-x)^{-\frac{1}{2}}\, dx = \frac{\Gamma(\frac{2n+1}{2})\,\Gamma(\frac{1}{2})}{\Gamma(n+1)} = \pi\,\frac{\frac{1}{2}\frac{3}{2}\cdots\frac{2n-1}{2}}{n!}\,,$$

we get for the moments of $X(1)$, by a comparison of coefficients,

$$\mathbb{E}\big[X(1)^n\big] = \frac{1}{\pi}\int_0^1 x^n \frac{1}{\sqrt{x(1-x)}}\, dx,$$

which implies that $X(1)$ is arcsine distributed.

Our second version of the Feynman-Kac formula is a uniqueness result for the case of zero dissipation rate, which will allow us the express the probability that linear Brownian motion exits an interval before a fixed time t in two different ways.

THEOREM 7.42. *If u is a bounded, twice continuously differentiable solution of the heat equation on the domain U, with zero dissipation rate and continuous initial condition f, then*

$$u(t,x) = \mathbb{E}_x\Big[f\big(B(t)\big)\,\mathbb{1}\{t<\tau\}\Big]\,, \tag{4.1}$$

where τ is the first exit time from the domain U.

Proof. The proof is based on Itô's formula, Theorem 7.14, and Remark 7.15. We let $K \subset U$ be compact and denote by σ the first exit time from K. Fixing $t>0$ and applying Itô's formula with $f(x,y)=u(t-y,x)$ and $\zeta(s)=s$ gives, for all $s<t$,

$$\begin{aligned}
&u(t-s\wedge\sigma, B(s\wedge\sigma)) - u(t,B(0))\\
&\quad= \int_0^{s\wedge\sigma} \nabla_x u(t-v,B(v))\cdot dB(v)\\
&\qquad - \int_0^{s\wedge\sigma} \partial_t u(t-v,B(v))\,dv + \tfrac12\int_0^{s\wedge\sigma} \Delta_x u(t-v,B(v))\,dv.
\end{aligned}$$

As u solves the heat equation, the last two terms on the right cancel. Hence, taking expectations,

$$\mathbb{E}_x\big[u(t-s\wedge\sigma, B(s\wedge\sigma))\big] = \mathbb{E}_x\big[u(t,B(0))\big] = u(t,x),$$

using that the stochastic integral has vanishing expectation. Exhausting U by compact sets, i.e. letting $\sigma\uparrow\tau$, leads to $\mathbb{E}_x[u(t-s,B(s))\,\mathbb{1}\{s<\tau\}] = u(t,x)$. Taking a limit $s\uparrow t$ gives the required result. ■

As an application of Theorem 7.42 we calculate the probability that a linear Brownian motion stays, up to time t, within an interval. As a warm-up we suggest to look at Exercise 7.8 where the easy case of a half-open interval is treated. Here we focus on intervals $[a,b]$, for $a<0<b$, and give two different formulas for the probability of staying in $[a,b]$ up to time t. To motivate the first formula, we start with a heuristic approach, which gives the correct result, and then base the rigorous proof on the Feynman-Kac formula.

For our heuristics we think of the transition (sub–)density, $q_t\colon [0,a]\times[0,a]\to[0,1]$ of a Brownian motion, which is killed upon leaving the interval $[0,a]$. In a first approximation we subtract from the transition density $\mathfrak{p}(t,x,y)$ of an unkilled Brownian motion the transition density for all the paths that reach level 0. By the reflection principle (applied to the first hitting time of level 0) the latter is equal to $\mathfrak{p}(t,x,-y)$.

We then subtract the transition density of all the paths that reach level a, which, again by the reflection principle, equals $\mathfrak{p}(t,x,2a-y)$, then add again the density of all the paths that reach level 0 after hitting a, as these have already been subtracted in the first step. This gives the approximation term $\mathfrak{p}(t,x,y)-\mathfrak{p}(t,x,-y)-\mathfrak{p}(t,x,2a-y)+\mathfrak{p}(t,x,2a+y)$.

Of course the iteration does not stop here (for example we have double-counted paths that reach level 0 after hitting a). Eventually we have to consider an infinite series of alternating reflections at levels 0 and a to obtain the density

$$q_t(x,y) = \sum_{k=-\infty}^{\infty} \big\{\mathfrak{p}(t,x,2ka+y) - \mathfrak{p}(t,x,2ka-y)\big\}\,.$$

Integrating this over $y \in [0,a]$ makes the following theorem plausible.

THEOREM 7.43. *Let $0 < x < a$. Then*

$$\begin{aligned}(4.2)\qquad &\mathbb{P}_x\big\{B(s) \in (0,a) \text{ for all } 0 \le s \le t\big\} \\ &= \sum_{k=-\infty}^{\infty} \Big\{ \Phi\big(\tfrac{2ka+a-x}{\sqrt{t}}\big) - \Phi\big(\tfrac{2ka-x}{\sqrt{t}}\big) - \Phi\big(\tfrac{2ka+a+x}{\sqrt{t}}\big) + \Phi\big(\tfrac{2ka+x}{\sqrt{t}}\big) \Big\},\end{aligned}$$

where $\Phi(x)$ is the distribution function of a standard normal distribution.

Proof. The left hand side in (4.2) agrees with the right hand side in Theorem 7.42 for $U = (0,a)$ and $f = 1$. The series on the right hand side is absolutely convergent, and hence satisfies the boundary conditions at $x = 0$ and $x = a$. It is also not difficult to verify that it is bounded. Elementary calculus gives

$$\partial_t \Phi\big(\tfrac{2ka+a-x}{\sqrt{t}}\big) = -\tfrac{2ka+a-x}{2t^{3/2}}\, \mathfrak{p}(t,x,2ka+a) = \tfrac{1}{2}\, \partial_{xx} \Phi\big(\tfrac{2ka+a-x}{\sqrt{t}}\big),$$

and similarly for the other summands. Hence termwise differentiation shows that the right hand side satisfies the heat equation. To see that the initial condition is fulfilled, note that (as $t \downarrow 0$) the sums over all $k > 0$ converge to zero by cancellation, and the sum over all $k < 0$ obviously converge to zero. Among the four terms belonging to $k = 0$, two terms with positive sign and one term with negative sign converge to one, whereas one term converges to zero. ∎

The solution of the heat equation is not in the form one would get by a naïve separation of variables approach. This approach yields a different, equally valuable, expression for the probability of interest. Indeed, writing $u(t,x) = v(t)w(x)$ one expects w to be an eigenfunction of $\frac{1}{2}\,\partial_{xx}$ on $(0,a)$ with zero boundary conditions. These eigenfunctions are

$$\sin\big(\tfrac{k\pi(2x-a)}{2a}\big), \qquad \text{for } k \text{ even}, \qquad \cos\big(\tfrac{k\pi(2x-a)}{2a}\big), \qquad \text{for } k \text{ odd},$$

with eigenvalues $-k^2\pi^2/(2a^2)$. As we are only interested in solutions symmetric about $a/2$ only the cosine terms will contribute For v we are looking for the eigenfunctions of ∂_t with the same eigenvalues, which are

$$\exp\big(-\tfrac{k^2\pi^2}{2a^2}\, t\big), \qquad \text{for } k \text{ odd},$$

and considering the initial condition (and shifting the cosine by $\pi/2$) leads to the solution

$$(4.3)\qquad u(t,x) = \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{1}{2n+1} \exp\big(-\tfrac{(2n+1)^2\pi^2}{2a^2}\, t\big) \sin\big(\tfrac{(2n+1)\pi x}{a}\big).$$

Therefore (4.3) is an alternative representation of the probability in (4.2). For practical purposes this series is more useful when t is large, as the convergence is faster, whereas the series in the theorem converges fast only for small values of $t > 0$.

We now prove an elliptic, or time-stationary, version of the Feynman-Kac formula. This will enable us to describe the distribution of the total time spent by a transient Brownian motion in unit ball in terms of a Laplace transfrom.

THEOREM 7.44. *Let $d \geq 3$ and $V \colon \mathbb{R}^d \to [0,\infty)$ be bounded. Define*

$$h(x) := \mathbb{E}_x\Big[\exp\Big(-\int_0^\infty V(B(t))\,dt\Big)\Big].$$

Then $h \colon \mathbb{R}^d \to [0,\infty)$ satisfies the equation

$$h(x) = 1 - \int G(x,y)V(y)\,h(y)\,dy \quad \text{for all } x \in \mathbb{R}^d.$$

REMARK 7.45. Informally, the integral equation in Theorem 7.44 implies $\frac{1}{2}\Delta h = Vh$, which is also what one gets from letting $t \uparrow \infty$ in Theorem 7.41. See also Exercise 2.18 for a converse result in a similar spirit. $\diamond$

Proof. Define the 'resolvent operator'

$$R_\lambda^V f(x) := \int_0^\infty e^{-\lambda t}\mathbb{E}_x\big[f(B(t))\,e^{-\int_0^t V(B(s))\,ds}\big]\,dt.$$

Using the fundamental theorem of calculus in the second step we obtain

$$\begin{aligned} R_\lambda^0 f(x) - R_\lambda^V f(x) &= \mathbb{E}_x \int_0^\infty dt\, e^{-\lambda t - \int_0^t V(B(s))\,ds}\, f(B(t))\,\big(e^{\int_0^t V(B(s))\,ds} - 1\big) \\ &= \mathbb{E}_x \int_0^\infty dt\, e^{-\lambda t - \int_0^t V(B(s))\,ds}\, f(B(t)) \int_0^t V(B(s))\, e^{\int_0^s V(B(r))\,dr}\,ds. \end{aligned}$$

Using Fubini's theorem and the Markov property, we may continue with

$$\begin{aligned} &= \mathbb{E}_x \int_0^\infty ds\, e^{-\lambda s}\, V(B(s)) \int_0^\infty dt\, \exp\big(-\lambda t - \int_0^t V(B(s+u))\,du\big)\, f(B(s+t)) \\ &= \mathbb{E}_x \int_0^\infty ds\, e^{-\lambda s} V(B(s))\, R_\lambda^V f(B(s)) = R_\lambda^0\big(V\, R_\lambda^V f\big)(x). \end{aligned}$$

The function h is related to the resolvent operator by the equation

$$h(x) = \lim_{\lambda \downarrow 0} \lambda R_\lambda^V 1(x).$$

Letting $f \equiv 1$ we obtain $1 - \lambda\, R_\lambda^V 1 = \lambda\, R_\lambda^0\big(V R_\lambda^V 1\big)$, and as $\lambda \downarrow 0$ we get

$$1 - h(x) = R_0^0\big(V\, h\big)(x) = \int G(x,y)V(y)\,h(y)\,dy,$$

for the Green's function $G(x,y) = (2\pi)^{-1}\,|x-y|^{-2}$, as claimed. ■

We use Theorem 7.44 to prove the three-dimensional case of the Ciesielski-Taylor identity, one of the most surprising identities about Brownian motion. Key to this is the following proposition.

PROPOSITION 7.46. *For a standard Brownian motion $\{B(t)\colon t \geq 0\}$ in dimension three let $T = \int_0^\infty \mathbb{1}\{B(t) \in \mathcal{B}(0,1)\}$ be the total occupation time of the unit ball. Then*

$$\mathbb{E}\big[e^{-\lambda T}\big] = \operatorname{sech}(\sqrt{2\lambda})\,.$$

Proof. Let $V(x) = \lambda \mathbb{1}_{\mathcal{B}(0,1)}$ and define $h(x) = \mathbb{E}_x[e^{-\lambda T}]$ as in Theorem 7.44. Then

$$h(x) = 1 - \lambda \int_{\mathcal{B}(0,1)} G(x,y)\,h(y)\,dy \text{ for all } x \in \mathbb{R}^d\,.$$

Clearly, we are looking for a rotationally symmetric function h. The integral on the right can therefore be split into two parts: First, the integral over $\mathcal{B}(0,|x|)$, which is the Newtonian potential due to a symmetric mass distribution on $\mathcal{B}(0,|x|)$ and therefore remains unchanged if the same mass is concentrated at the origin. Second, the integral over $\mathcal{B}(0,1) \setminus \mathcal{B}(0,|x|)$, which is harmonic on the open ball $\mathcal{B}(0,|x|)$ with constant value on the boundary, so itself is constant. Hence, for $x \in \mathcal{B}(0,1)$, to

$$1 - h(x) = \frac{\lambda}{2\pi|x|} \int_{\mathcal{B}(0,|x|)} h(y)\,dy + \lambda \int_{\mathcal{B}(0,1)\setminus\mathcal{B}(0,|x|)} \frac{h(y)}{2\pi|y|}\,dy.$$

With $u(r) = rh(x)$ for $|x| = r$ we have, for $0 < r < 1$,

$$r - u(r) = 2\lambda \int_0^r su(s)\,ds + 2\lambda r \int_r^1 u(s)\,ds\,,$$

and by differentiation $u'' = 2\lambda u$ on $(0,1)$. Hence

$$u(r) = Ae^{\sqrt{2\lambda} r} + Be^{-\sqrt{2\lambda} r}\,.$$

The boundary conditions $u(0) = 0$ and $u'(1) = 1$ give $B = -A$ and

$$A = \frac{1}{\sqrt{2\lambda}}\,\frac{1}{e^{\sqrt{2\lambda}} + e^{-\sqrt{2\lambda}}}\,.$$

Then

$$\begin{aligned} h(x) &= \lim_{r\downarrow 0} \frac{u(r)}{r} = 1 - 2\lambda \int_0^1 u(r)\,dr \\ &= 1 - A\sqrt{2\lambda}\,\big(e^{\sqrt{2\lambda}} + e^{-\sqrt{2\lambda}} - 2\big) = \frac{2}{e^{\sqrt{2\lambda}} + e^{-\sqrt{2\lambda}}} = \operatorname{sech}(\sqrt{2\lambda})\,, \end{aligned}$$

as required to complete the proof. ■

THEOREM 7.47 (Ciesielski-Taylor identity). *The first exit time from the unit ball by a standard Brownian motion in dimension one and the total occupation time of the unit ball by a standard Brownianian motion in dimension $d = 3$ have the same distribution.*

Proof. The Laplace transform of the first exit time from the unit interval $(-1,1)$ is given in Exercise 2.16. It coincides with the Laplace transform of T given in Proposition 7.46. Hence the two distributions coincide. ■

Exercises

Exercise 7.1 (∗). Suppose $\{H(s,\omega)\colon s \geq 0\,, \omega \in \Omega\}$ is progressively measurable and $\{B(t)\colon t \geq 0\}$ a linear Brownian motion. Show that for any stopping time T with

$$\mathbb{E}\Big[\int_0^T H(s)^2\, ds\Big] < \infty,$$

we have

(a) $\mathbb{E}\Big[\displaystyle\int_0^T H(s)\, dB(s)\Big] = 0,$

(b) $\mathbb{E}\Big[\Big(\displaystyle\int_0^T H(s)\, dB(s)\Big)^2\Big] = \mathbb{E}\Big[\displaystyle\int_0^T H(s)^2\, ds\Big].$

Exercise 7.2. Suppose that $f\colon [0,1] \to \mathbb{R}$ is in the Cameron-Martin space, i.e. $f(t) = \int_0^t f'(s)\, ds$ for all $t \in [0,1]$ and $f' \in L^2(0,1)$. Then, almost surely,

$$\int_0^1 f'(s)\, dB(s) = \lim_{n\to\infty} n \sum_{j=0}^{n} \Big(f\big(\tfrac{j+1}{n}\big) - f\big(\tfrac{j}{n}\big)\Big)\Big(B\big(\tfrac{j+1}{n}\big) - B\big(\tfrac{j}{n}\big)\Big).$$

Exercise 7.3 (∗). Give the details of the proof of the multidimensional Itô formula, Theorem 7.14.

Exercise 7.4 (∗). Give an alternative proof of Theorem 2.33 based on a conformal mapping of the halfplanes $\{(x,y)\colon x > t\}$ onto the unit disk, which exploits our knowledge of harmonic measure on spheres.

Exercise 7.5 (∗). Let $\{B(t)\colon t \geq 0\}$ be a planar Brownian motion. Show that, if $\theta(t)$ is the continuous determination of the angle of $B(t)$, we have, almost surely,

$$\liminf_{t\uparrow\infty} \theta(t) = -\infty \qquad \text{and} \qquad \limsup_{t\uparrow\infty} \theta(t) = \infty.$$

Exercise 7.6. Formalise and prove the statement that, for every $\varepsilon > 0$, a planar Brownian motion winds around its starting point infinitely often in any time interval $[0,\varepsilon]$.

Exercise 7.7 (*). Show that under suitable conditions, stochastic integrals and ordinary integrals can be interchanged: Suppose $h\colon \mathbb{R} \to [0,\infty)$ is a continuous function with compact support. Then, almost surely,

$$\int_{-\infty}^{\infty} h(a)\Big(\int_0^t \mathrm{sign}(B(s)-a)\, dB(s)\Big)\, da = \int_0^t \Big(\int_{-\infty}^{\infty} h(a)\,\mathrm{sign}(B(s)-a)\, da\Big)\, dB(s).$$

Hint. Write the outer integral on the left hand side as a limit of Riemann sums. For this purpose it is useful to know that, by Tanaka's formula and continuity of the local times, the integrand has a continuous modification.

Exercise 7.8.

(a) Show that the function $u\colon (0,\infty)\times(0,\infty) \to \mathbb{R}$ given by

$$u(t,x) = \sqrt{\frac{2}{\pi t}} \int_0^x e^{-\frac{z^2}{2t}}\, dz$$

solves the heat equation on the domain $(0,\infty)$ with zero dissipation rate and constant initial condition $f = 1$.

(b) Infer from this that, for $x > 0$,

$$\mathbb{P}_x\big\{B(s) > 0 \text{ for all } s \le t\big\} = \sqrt{\frac{2}{\pi t}} \int_0^x e^{-\frac{z^2}{2t}}\, dz.$$

(c) Explain how the result of (b) could have been obtained from the reflection principle.

Exercise 7.9. Prove the Erdős-Kac theorem: Let $X_1, X_2, \ldots$ be a sequence of independent identically distributed random variables with mean zero and variance one. Let $S_n = X_1 + \cdots + X_n$ and $T_n = \max\{|S_1|, \ldots, |S_n|\}$. Then

$$\lim_{n\to\infty} \mathbb{P}\{T_n < x\} = \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \exp\Big(-\frac{(2n+1)^2\pi^2}{8x^2}\Big).$$

Notes and Comments

The first stochastic integral with a random integrand was defined by Itô [**It44**] but stochastic integrals with respect to Brownian motion with deterministic integrands were known to Paley, Wiener and Zygmund already in 1933, see [**PWZ33**]. Our stochastic integral is by far not the most general construction possible, the complete theory of Itô integration is one of the cornerstones of modern probability. Interesting further material can be found, for example, in the books [**CW90**], [**RW00**] or [**Du96**].

Itô's formula, first proved in [**It51**], plays a central role in stochastic analysis, quite like the fundamental theorem of calculus does in real analysis. The version we give is designed to minimise the technical effort to get to the desired applications, but a lot more can be said if the discussion is extended to the concept of semimartingales, the references in the previous paragraph provide the background for this.

Conformal invariance was known to Lévy and a sketch of a proof is given in the book [**Le48**]. It is interesting to note that this fact does not extend to higher dimensions $d \geq 3$. There are not many nontrivial conformally invariant maps anyway, but essentially the only one, inversion on a sphere, fails. This is easy to see, as the image of Brownian motion stopped on the boundary of the punctured domain $\mathcal{B}(0,1) \setminus \{0\}$ has zero probability of not hitting $\mathcal{B}(0,1)$.

There is rich interaction between complex analysis and Brownian motion, which relies on conformal invariance. The conformal invariance of harmonic measure, which we proved in Theorem 7.22, is not easy to obtain by purely analytical means. Another result from complex analysis, which can be proved effectively using Brownian motion is Picard's theorem, see Davis [**Da75**]. The theorem states that a nonconstant entire function has a range which omits at most one point from the complex plane. Only very recently a completely new perspective on conformal invariance has opened up through the theory of conformally invariant random curves developed by Lawler, Schramm, and Werner, see e.g. [**We04**].

The skew-product representation has many nice applications, for more examples see [**LG91**], which also served as the backbone of our exposition. The first result about the windings of Brownian motion is Spitzer's law, first proved by F. Spitzer in [**Sp58**]. There are plenty of extensions including pathwise laws [**Sh98, Mö02**], windings around more than one point, and joint laws of windings and other functionals [**PY86**]. A discussion of some problems related to this can be found in [**Yo92**].

Tanaka's formula offers many fruitful openings, among them the theory of local times for semimartingales, which is presented in [**RY94**]. The formula goes back to the paper by Tanaka [**Ta63**]. Alternative to our approach, Tanaka's formula can be taken as a definition of Brownian local time. Then continuity can be obtained from the Kolmogorov-Čentsov theorem and moment estimates based on the Burkholder-Davis-Gundy inequalities, see for example the book by Karatzas and Shreve [**KS88**].

The Feynman-Kac formula is a classical application in stochastic calculus, which is discussed in more detail in [**KS88**]. It can be exploited to obtain an enormous variety of distributional properties of Brownian motion, see [**BS02**] for (literally!) thousands of examples. The converse, application of Brownian motion to study equations, is of course equally natural. Del Moral [**DM04**] gives an impressive account of the wide applicability of this formula and its variants.

The identity between the two formulas describing the probability that a Brownian motion stays between two barriers serves as a standard example for the Poisson summation formula, see [**Fe66**, X.5 and XIX.5]. According to Feller it was discovered originally in connection with Jacobi's theory of transformations of theta functions, see Landau [**La09**, Satz 277].

The 'iterated reflection' argument, which we have used to determine the transition density of a Brownian motion with absorbing barriers may also be used to determine transition densities for a Brownian motion which is reflected at the barriers, see [**Fe66**, X.5]. In higher dimensions Brownian motion reflected at the boundaries of a domain is an interesting subject, not least because of its connections to partial differential equations with Neumann boundary conditions, see, for example, [**Br76**].

The Erdős-Kac law plays an important rôle for the Kolmogorov-Smirnov test known from non-parametric statistics, see e.g. [**Fe68**]. Plenty of proofs of the arcsine law are known: Besides the two provided in this book, there is also an approach of Kac [**Ka51**] based on the Meyer-Tanaka formula, and Rogers and Williams [**RW00**, III.24] provide a proof based on local time theory.

The Ciesielski-Taylor identity was found by Ciesielski and Taylor in 1962 by an explicit calculation, see [**CT62**]. It extends to arbitrary dimensions d, stating that the law of the exit times from the unit ball by a standard Brownian motion in dimension d equals the law of the total occupation time in the unit ball by the standard Brownian motion in dimension $d + 2$. Our argument is taken from [**Sp64**], see also [**RW00**, III.20]. Many proofs of this fact are known, see for example [**Yo92**], but none provides a geometrically intuitive explanation and it may well be that none exists.

CHAPTER 8

Potential theory of Brownian motion

In this chapter we develop the key facts of the potential theory of Brownian motion. This theory is centred around the notions of a harmonic function, the energy of a measure, and the capacity of a set. The probabilistic problem at the heart of this chapter is to find the probability that Brownian motion visits a given set.

1. The Dirichlet problem revisited

We now take up the study of the Dirichlet problem again and ask for sharp conditions on the domain which ensure the existence of solutions, which allow us to understand the problem for domains with very irregular boundaries, like for example connected components of the complement of a planar Brownian curve.

In this chapter, stochastic integrals and Itô's formula will be a helpful tool. As a warm-up, we suggest to use these tools to give a probabilistic proof of the mean value property of harmonic functions, see Exercise 8.1.

Recall from Example 3.15 that the existence of a solution of the Dirichlet problem may be in doubt by the fact that Brownian motion started at the boundary ∂U may not leave the domain U immediately. Indeed, we show here that this is the only problem that can arise.

DEFINITION 8.1. *A point $x \in A$ is called* **regular** *for the closed set $A \subset \mathbb{R}^d$ if the first hitting time $T_A = \inf\{t > 0 : B(t) \in A\}$ satisfies $\mathbb{P}_x\{T_A = 0\big\} = 1$. A point which is not regular is called* **irregular**. $\diamond$

REMARK 8.2. In the case $d = 1$ we have already seen that for any starting point $x \in \mathbb{R}$, almost surely a Brownian motion started in x returns to x in every interval $[0, \varepsilon)$ with $\varepsilon > 0$. Hence every point is regular for any set containing it. $\diamond$

We already know a condition which implies that a point is regular, namely the Poincaré cone condition introduced in Chapter 3.

THEOREM 8.3. *If the domain $U \subset \mathbb{R}^d$ satisfies the Poincaré cone condition at $x \in \partial U$, then x is regular for the complement of U.*

Proof. Suppose $x \in \partial U$ satisfies the condition, then there is an open cone V with height $h > 0$ and angle $\alpha > 0$ in U^c based at x. Then the first exit time τ_U for the domain satisfies

$$\mathbb{P}_x\{\tau_U \leq t\} \geq \mathbb{P}_x\{B(t) \in V \cap \mathcal{B}(x,h)\} \geq \mathbb{P}_x\{B(t) \in V\} - \mathbb{P}_x\{B(t) \notin \mathcal{B}(x,h)\},$$

By Brownian scaling the last term equals $\mathbb{P}\{B(1) \in V\} - \mathbb{P}_x\{B(1) \notin \mathcal{B}(x, h/\sqrt{t})\}$. For $t \downarrow 0$ the subtracted term goes to zero, and hence $\mathbb{P}_x\{\tau_U = 0\} = \lim_{t\downarrow 0} \mathbb{P}_x\{\tau_U \le t\} = \mathbb{P}\{B(1) \in V\} > 0$. By Blumenthal's zero-one law we have $\mathbb{P}_x\{\tau_U = 0\} = 1$, in other words x is regular for U^{c}. ■

Remark 8.4. At the end of this chapter we will be able to improve this and give a sharp condition for a point to be regular, *Wiener's test* of regularity. ◇

Theorem 8.5 (Dirichlet Problem). *Suppose $U \subset \mathbb{R}^d$ is a bounded domain and φ be a continuous function on ∂U. Define $\tau = \min\{t \ge 0 \colon B(t) \in \partial U\}$, and define $u \colon \overline{U} \to \mathbb{R}$ by*

$$u(x) = \mathbb{E}_x\big[\varphi(B(\tau))\big].$$

(a) *A solution to the Dirichlet problem exists if and only if the function u is a solution to the Dirichlet problem with boundary condition φ.*
(b) *u is a harmonic function on U with $u(x) = \varphi(x)$ for all $x \in \partial U$ and is continuous at every point $x \in \partial U$ that is regular for the complement of U.*
(c) *If every $x \in \partial U$ is regular for the complement of U, then u is the unique continuous function $u \colon \overline{U} \to \mathbb{R}$ which is harmonic on U such that $u(x) = \varphi(x)$ for all $x \in \partial U$.*

Proof. For the proof of (a) let v be any solution of the Dirichlet problem on U with boundary condition φ. Define open sets $U_n \uparrow U$ by $U_n = \big\{x \in U \, : \, |x-y| > \frac{1}{n} \text{ for all } y \in \partial U\big\}$. Let τ_n be the first exit time of U_n and τ the first exit time from U, which are stopping times. As $\Delta v(x) = 0$ for all $x \in U$ we see from the multidimensional version of Itô's formula that

$$v(B(t \wedge \tau_n)) = v(B(0)) + \sum_{i=1}^d \int_0^{t\wedge\tau_n} \frac{\partial v}{\partial x_i}(B(s))\, dB_i(s) + \tfrac{1}{2} \sum_{i=1}^d \int_0^{t\wedge\tau_n} \frac{\partial^2 v}{\partial x_i^2}(B(s))\, ds\,.$$

Note that $\partial v/\partial x_i$ is bounded on the closure of U_n, and thus everything is well-defined. The last term vanishes as $\Delta v(x) = 0$ for all $x \in U$. Taking expectations the second term on the right also vanishes, by Exercise 7.1, and we get that

$$\mathbb{E}_x\big[v(B(t \wedge \tau_n))\big] = \mathbb{E}_x\big[v(B(0))\big] = v(x), \qquad \text{for } x \in U_n\,.$$

Note that v, and hence the integrand on the left hand side, are bounded. Moreover, it is easy to check using boundedness of U and a reduction to the one-dimensional case, that τ is almost surely finite. Hence, as $t \uparrow \infty$ and $n \to \infty$, bounded convergence yields that the left hand side converges to $\mathbb{E}_x[v(B(\tau))] = \mathbb{E}_x[\varphi(B(\tau))]$. The result follows, as the right hand side depends neither on t nor on n.

The harmonicity statement of (b) is included in Theorem 3.8, and $u = \varphi$ on ∂U is obvious from the definition. It remains to show the continuity claim. For a regular $x \in \partial U$ we now show that if Brownian motion is started at a point in $\overline{U}$, which is sufficiently close to x, then with high probability the Brownian motion hits U^{c}, before leaving a given ball $\mathcal{B}(x, \delta)$.

We start by noting that, for every $t > 0$ and $\eta > 0$ the set

$$O(t, \eta) := \big\{z \in U \, : \, \mathbb{P}_z\{\tau \le t\} > \eta\big\}$$

is open. Indeed, if $z \in O(t, \eta)$, then for some small $s > 0$ and $\delta > 0$ and large $M > 0$, we have

$$\mathbb{P}_z\big\{|B(s) - z| \le M,\, B(u) \in U^{\mathrm{c}} \text{ for some } s \le u \le t\big\} > \eta + \delta.$$

By the Markov property the left hand side above can be written as

$$\int_{\mathcal{B}(z,M)} \mathbb{P}_\xi\big\{B(u) \in U^c \text{ for some } 0 \le u \le t-s\big\}\, \mathfrak{p}(s,z,\xi)\, d\xi,$$

Now let $\varepsilon > 0$ so small that $|\mathfrak{p}(s,z,\xi) - \mathfrak{p}(s,y,\xi)| < \delta/\mathcal{L}(\mathcal{B}(0,M))$ for all $|z-y| < \varepsilon$ and $\xi \in \mathbb{R}^d$. Then we have $\mathbb{P}_y\{\tau \le t\} \ge \mathbb{P}_y\big\{B(u) \in U^c \text{ for some } s \le u \le t\big\} > \eta$, hence the ball $\mathcal{B}(z,\varepsilon)$ is in $O(t,\eta)$, which therefore must be open.

Given $\varepsilon > 0$ and $\delta > 0$ we now choose $t > 0$ small enough, such that for $\tau' = \inf\{s > 0 \colon B(s) \notin \mathcal{B}(x,\delta)\}$ we have $\mathbb{P}_z\big\{\tau' < t\big\} < \varepsilon/2$ for all $|\,x - z| < \delta/2$. By regularity we have $x \in O(t, 1-\varepsilon/2)$, and hence we can choose $0 < \theta < \delta/2$ to achieve $\mathcal{B}(x,\theta) \subset O(t, 1-\varepsilon/2)$. We have thus shown that,

$$|\,x - z| < \theta \Rightarrow \mathbb{P}_z\big\{\tau < \tau'\big\} > 1 - \varepsilon. \tag{1.1}$$

To complete the proof, let $\varepsilon > 0$ be arbitrary. Then there is a $\delta > 0$ such that $|\varphi(x) - \varphi(y)| < \varepsilon$ for all $y \in \partial U$ with $|\,x - y| < \delta$. Choose θ as in (1.1). For all $z \in \overline{U}$ with $|z - x| < \delta \wedge \theta$ we get

$$|u(x) - u(z)| = \big|\mathbb{E}_z[\varphi(x) - \varphi(B(\tau))]\big| \le 2\|\varphi\|_\infty\, \mathbb{P}_z\big\{\tau' < \tau\big\} + \varepsilon \le \varepsilon\,(2\|\varphi\|_\infty + 1).$$

As $\varepsilon > 0$ can be arbitrarily small, u is continuous at $x \in \partial U$, and part (c) follows trivially from (b) and the maximum principle. ■

A further classical problem of partial differential equations, the Poisson problem, is related to Brownian motion in a way quite similar to the Dirichlet problem.

Definition 8.6. *Let $U \subset \mathbb{R}^d$ be a bounded domain and $u \colon \overline{U} \to \mathbb{R}$ be a continuous function, which is twice continuously differentiable on U. Let $g \colon U \to \mathbb{R}$ be continuous. Then u is said to be the* **solution of Poisson's problem for** *g if $u(x) = 0$ for all $x \in \partial U$ and*

$$-\tfrac{1}{2}\Delta u(x) = g(x) \qquad \text{for all } x \in U. \qquad \diamond$$

Remark 8.7. The probabilistic approach to the Poisson problem will be discussed in Exercises 8.2 and 8.3. We show that, for g bounded, the solution u of Poisson's problem for g, if it exists, equals

$$u(x) = \mathbb{E}_x\Big[\int_0^T g(B(t))\, dt\Big] \quad \text{for } x \in U\,, \tag{1.2}$$

where $T := \inf\{t > 0 \,:\, B(t) \notin U\}$. Conversely, if g is Hölder continuous and every $x \in \partial U$ is regular for the complement of U, then the function (1.2) solves the Poisson problem for g. ◇

Remark 8.8. If u solves Poisson's problem for $g \equiv 1$ in a domain $U \subset \mathbb{R}^d$, then $u(x) = \mathbb{E}_x[T]$ is the average time it takes a Brownian motion started in x to leave the set U. ◇

2. The equilibrium measure

In Chapter 3 we have studied the distribution of the location of the *first entry* of a Brownian motion into a closed set Λ, the harmonic measure. In the case of a transient (or killed) Brownian motion there is a natural counterpart to this by looking at the distribution of the position of the *last exit* from a closed set. This leads to the notion of the *equilibrium measure*, which we discuss and apply in this section.

To motivate the next steps we first look at a simple random walk $\{X_n : n \in \mathbb{N}\}$ in $d \geq 3$. Let $A \subset \mathbb{Z}^d$ be a bounded set, then by transience the last exit time $\gamma = \max\{n \in \mathbb{N} : X_n \in A\}$ is finite on the event that the random walk ever hits A. Note that γ is *not* a stopping time. Then, for any $x \in \mathbb{Z}^d$ and $y \in A$,

$$\begin{aligned}\mathbb{P}_x\big\{X \text{ hits } A \text{ and } X_\gamma = y\big\} &= \sum_{k=0}^{\infty} \mathbb{P}_x\big\{X_k = y, X_j \notin A \text{ for all } j > k\big\} \\ &= \sum_{k=0}^{\infty} \mathbb{P}_x\big\{X_k = y\big\}\mathbb{P}_y\{\gamma = 0\},\end{aligned}$$

and introducing the Green's function $G(x,y) = \sum_{k=0}^{\infty} \mathbb{P}_x\big\{X_k = y\big\}$ we get, for all $y \in A$,

$$\mathbb{P}_x\big\{X \text{ hits } A \text{ and } X_\gamma = y\big\} = G(x,y)\,\mathbb{P}_y\{\gamma = 0\}.$$

This holds also, trivially, for all $y \in \mathbb{Z}^d \setminus A$. Summing over all $y \in \mathbb{Z}^d$ gives

$$\mathbb{P}_x\big\{X \text{ ever hits } A\big\} = \sum_{y \in \mathbb{Z}^d} G(x,y)\mathbb{P}_y\{\gamma = 0\}.$$

The formula allows us to describe the probability of ever hitting a set as a potential with respect to the measure $y \mapsto \mathbb{P}_y\{\gamma = 0\}$, which is supported on A. Our aim in this section is to extend this to Brownian motion.

Note that the argument above relied heavily on the transience of the random walk. This is no different in the case of Brownian motion. In order to include the two-dimensional case we 'kill' the Brownian motion, either when it exits a large domain or at an independent exponential stopping time. Note that both possibilities preserve the strong Markov property, in the case of exponential killing this is due to the lack-of-memory property of the exponential distribution.

To formally explain our setup we now suppose that $\{B(t) : 0 \leq t \leq T\}$ is a transient Brownian motion in the sense of Chapter 3. Recall that this means that $\{B(t) : 0 \leq t \leq T\}$ is a d-dimensional Brownian motion killed at time T, and one of the following three cases holds:

(1) $d \geq 3$ and $T = \infty$,
(2) $d \geq 2$ and T is an independent exponential time,
(3) $d \geq 2$ and T is the first exit time from a bounded domain D containing 0.

We use the convention that $D = \mathbb{R}^d$ in cases (1),(2). In all cases, transient Brownian motion is a Markov process and, by Proposition 3.29 its transition kernel has a density, which we denote by $\mathfrak{p}^*(t,x,y)$. Note that in case (2,3) the function $\mathfrak{p}^*(t,x,y)$ is only a subprobability density because of the killing, indeed it is strictly smaller than the corresponding density without killing. The associated Green's function

$$G(x,y)=\int_0^\infty \mathfrak{p}^*(t,x,y)\,dt,$$

is always well-defined and finite for all $x\neq y$.

THEOREM 8.9 (Last exit formula). *Suppose $\{B(t)\,:\,0\le t\le T\}$ is a transient Brownian motion and $\Lambda\subset\mathbb{R}^d$ a compact set. Let $\gamma=\sup\big\{t\in(0,T]\,:\,B(t)\in\Lambda\big\}$ be the* last exit time *from Λ, using the convention $\gamma=0$ if the path does not hit Λ. Then there exists a finite measure ν on Λ called the* **equilibrium measure**, *such that, for any Borel set $A\subset\Lambda$ and $x\in D$,*

$$\mathbb{P}_x\big\{B(\gamma)\in A, 0<\gamma\le T\big\}=\int_A G(x,y)\,d\nu(y).$$

REMARK 8.10. Observe that the equilibrium measure is uniquely determined by the last exit formula above. The proof of Theorem 8.9 is similar to the simple calculation in the discrete case, the equilibrium measure is constructed as limit of the measure $\varepsilon^{-1}\mathbb{P}_y\{0<\gamma\le\varepsilon\}\,dy$. ◇

Proof of Theorem 8.9. Let U_ε be a uniform random variable on $[0,\varepsilon]$, independent of the Brownian motion and the killing time. Then, for any bounded and continuous $f\colon D\to\mathbb{R}$,

$$\begin{aligned}
&\mathbb{E}_x\big[f(B(\gamma-U_\varepsilon))\,\mathbb{1}\{U_\varepsilon<\gamma\}\big]\\
&\quad=\varepsilon^{-1}\int_0^\infty \mathbb{E}_x\big[f(B(t))\mathbb{1}\{t<\gamma\le t+\varepsilon\}\big]\,dt\\
&\quad=\varepsilon^{-1}\int_0^\infty \mathbb{E}_x\big[f(B(t))\mathbb{1}\{t\le T\}\,\mathbb{P}_{B(t)}\{0<\gamma\le\varepsilon\}\big]\,dt\,.
\end{aligned}$$

Using the notation $\psi_\varepsilon(x)=\varepsilon^{-1}\mathbb{P}_x\{0<\gamma\le\varepsilon\}$ this equals

$$\begin{aligned}
\int_0^\infty \mathbb{E}_x\big[f\cdot\psi_\varepsilon(B(t))\mathbb{1}\{t\le T\}\big]\,dt&=\int_0^\infty\!\!\int_D \mathfrak{p}^*(t,x,y)\,f\cdot\psi_\varepsilon(y)\,dy\,dt\\
&=\int_D f(y)\,G(x,y)\,\psi_\varepsilon(y)\,dy.
\end{aligned}$$

This means that the subprobability measure η_ε defined by

$$\eta_\varepsilon(A)=\mathbb{P}_x\big\{B(\gamma-U_\varepsilon)\in A,\,U_\varepsilon<\gamma\le T\big\}$$

has the density $G(x,y)\,\psi_\varepsilon(y)$. Therefore also,

$$G(x,y)^{-1}\,d\eta_\varepsilon(y)=\psi_\varepsilon(y)\,dy. \tag{2.1}$$

Observe now that, by continuity of the Brownian path, $\lim_{\varepsilon\downarrow0}\eta_\varepsilon=\eta_0$ in the sense of weak convergence, where the measure η_0 on Λ is defined by

$$\eta_0(A)=\mathbb{P}_x\big\{B(\gamma)\in A,\,0<\gamma\le T\big\},$$

for all Borel sets $A\subset\Lambda$. As, for fixed $x\in D$, the function $y\mapsto G(x,y)^{-1}$ is continuous and bounded on Λ, we infer that, in the sense of weak convergence

$$\lim_{\varepsilon\downarrow0}G(x,y)^{-1}\,d\eta_\varepsilon=G(x,y)^{-1}\,d\eta_0.$$

By (2.1) the measure $\psi_\varepsilon(y)\,dy$ therefore converges weakly to a limit measure ν, which does not depend on x, and satisfies $G(x,y)^{-1}\,d\eta_0(y) = d\nu(y)$ for all $x \in D$. As η_0 has no atom in x we therefore obtain that $d\eta_0(y) = G(x,y)\,d\nu(y)$ for all $x \in D$. Integrating over any Borel set A gives the statement. ■

As a first application we give an estimate for the probability that Brownian motion in $\mathbb{R}^d$, for $d \geq 3$, hits a set contained in an annulus around x.

COROLLARY 8.11. *Suppose $\{B(t) : t \geq 0\}$ is Brownian motion in $\mathbb{R}^d$, with $d \geq 3$, and $\Lambda \subset \mathcal{B}(x,R) \setminus \mathcal{B}(x,r)$ is compact. Then*

$$R^{2-d}\nu(\Lambda) \leq \mathbb{P}_x\big\{\{B(t) : t \geq 0\} \text{ ever hits } \Lambda\big\} \leq r^{2-d}\nu(\Lambda),$$

where ν is the equilibrium measure on Λ.

Proof. By Theorem 8.9 in the case $A = \Lambda$ we have

$$\mathbb{P}_x\big\{\{B(t) : t \geq 0\} \text{ ever hits } \Lambda\big\} = \int_\Lambda G(x,y)\,d\nu(y).$$

Recall that $G(x,y) = |\,x-y|^{2-d}$ and use that $R^{2-d} \leq G(x,y) \leq r^{2-d}$. ■

Theorem 8.5 makes us interested in statements claiming that the set of irregular points of a set A is small. The following fundamental result will play an important role in the next chapter.

THEOREM 8.12. *Suppose $A \subset \mathbb{R}^d$, $d \geq 2$ is a closed set and let A^{r} be the set of regular points for A. Then, for all $x \in \mathbb{R}^d$,*

$$\mathbb{P}_x\big\{B(t) \in A \setminus A^{\mathrm{r}} \text{ for some } t > 0\big\} = 0,$$

in other words, the set of irregular points is polar for Brownian motion.

For the proof of Theorem 8.12 we have to develop a tool of independent interest, the strong maximum principle. A special case of this is the following statement, from which Theorem 8.12 follows without too much effort.

THEOREM 8.13. *Let $\{B(t) : t \geq 0\}$ be a d-dimensional Brownian motion, and T an independent exponential time. Let $\Lambda \subset \mathbb{R}^d$ be a compact set and define $\tau = \inf\{t > 0 : B(t) \in \Lambda\}$. If for some $\vartheta < 1$, we have $\mathbb{P}_x\big\{\tau < T\big\} \leq \vartheta$ for all $x \in \Lambda$, then $\mathbb{P}_x\big\{\tau < T\big\} \leq \vartheta$ for all $x \in \mathbb{R}^d$.*

Proof of Theorem 8.12. We can write the set of irregular points of A as a countable union of compact sets

$$A \setminus A^{\mathrm{r}} = \bigcup_{\ell=1}^{\infty}\bigcup_{m=1}^{\infty}\bigcup_{n=1}^{\infty}\Big\{x \in A \cap \mathcal{B}(0,m) : \mathbb{P}_x\{\tau(A) \leq T(n)\} \leq 1 - \tfrac{1}{\ell}\Big\},$$

where $T(n)$ is an independent exponential time with mean $1/n$ and $\tau(A)$ is the first hitting time of A. It suffices to prove that Brownian motion does not hit any fixed set in the union, so let ℓ, m, n be fixed and take $T = T(n)$, $\vartheta = 1 - 1/\ell$ and a compact set

$$\Lambda = \Big\{x \in A \cap \mathcal{B}(0,m) \,:\, \mathbb{P}_x\{\tau(A) \le T\} \le \vartheta\Big\}.$$

If $x \in \Lambda$, then, writing τ for the first hitting time of $\Lambda \subset A$,

$$\mathbb{P}_x\{\tau \le T\} \le \mathbb{P}_x\{\tau(A) \le T\} \le \vartheta,$$

for all $x \in \Lambda$ and therefore by Theorem 8.13 for all $x \in \mathbb{R}^d$.

Now suppose $x \in \mathbb{R}^d$ is the arbitrary starting point of a Brownian motion $\{B(t) \,:\, t \ge 0\}$ and $\Lambda(\varepsilon) = \{y \in \mathbb{R}^d \,:\, |y - z| \le \varepsilon$ for some $z \in \Lambda\}$. Define τ_ε as the first hitting time of $\Lambda(\varepsilon)$. Clearly, as Λ is closed,

$$\lim_{\varepsilon \downarrow 0} \mathbb{P}_x\big\{\tau_\varepsilon \le T\big\} = \mathbb{P}_x\big\{\tau \le T\big\}.$$

Moreover, by the strong Markov property applied at the stopping time τ_ε and the lack of memory property of exponential random variables,

$$\mathbb{P}_x\big\{\tau \le T\big\} \le \mathbb{P}_x\big\{\tau_\varepsilon \le T\big\} \max_{z \in \Lambda_\varepsilon} \mathbb{P}_z\big\{\tau \le T\big\} \le \mathbb{P}_x\big\{\tau_\varepsilon \le T\big\}(1 - \vartheta),$$

and letting $\varepsilon \downarrow 0$ we obtain

$$\mathbb{P}_x\big\{\tau \le T\big\} \le \mathbb{P}_x\big\{\tau \le T\big\}(1 - \vartheta).$$

As $\vartheta < 1$ this implies $\mathbb{P}_x\big\{\tau \le T\big\} = 0$, and as T is independent of the Brownian motion and can take arbitrarily large values with positive probability, we infer that the Brownian motion started in x never hits Λ. ■

The idea in the proof of Theorem 8.13 is to use the equilibrium measure ν to express $\mathbb{P}_x\big\{\tau < T\big\}$ as a potential, which means that, denoting the parameter of the exponential by $\lambda > 0$,

$$\mathbb{P}_x\big\{\tau < T\big\} = \int G_\lambda(x,y)\, d\nu(y),$$

where G_λ is the Green's function for the Brownian motion stopped at time T, i.e.

$$G_\lambda(x,y) = \int_0^\infty e^{-\lambda t}\, \mathfrak{p}(t,x,y)\, dt.$$

Recall that for any fixed y the function $x \mapsto G_\lambda(x,y)$ is subharmonic on $\mathbb{R}^d \setminus \{y\}$, by Exercise 3.11, and this implies that

$$U_\lambda \nu(x) = \int G_\lambda(x,y) d\nu(y)$$

is subharmonic on Λ^{c}. If $U_\lambda\nu$ was also continuous on the closure of Λ^{c}, then the maximum principle in Theorem 3.5 would tell us that $U_\lambda\nu$ has its maxima on the boundary $\partial\Lambda$ and this would prove Theorem 8.13. However we do not know the continuity of $U_\lambda\nu$ on the closure of Λ^{c}, so we need a strengthening of the maximum principle.

We now let K be a **kernel**, i.e. a measurable function $K\colon \mathbb{R}^d \times \mathbb{R}^d \to [0,\infty]$. Suppose that $x \mapsto K(x,y)$ is subharmonic outside $\{y\}$, and that $K(x,y)$ is a continuous and decreasing function of the distance $|\,x-y|$. For any finite measure μ on the compact set Λ, let

$$U_\mu(x) = \int K(x,y)\,d\mu(y)$$

be the potential of μ at x with respect to the kernel K.

THEOREM 8.14 (Strong maximum principle). *Then, for any $\vartheta < 1$, we have the equivalence*

$$U_\mu(x) \le \vartheta \text{ for all } x \in \Lambda \qquad \Leftrightarrow \qquad U_\mu(x) \le \vartheta \text{ for all } x \in \mathbb{R}^d.$$

REMARK 8.15. Note that this completes the proof of Theorem 8.13 and hence of Theorem 8.12 by applying it to the special case of the kernel $K = G_\lambda$ and the equilibrium measure. $\diamond$

The proof we present relies on a beautiful geometric lemma.

LEMMA 8.16. *There is a number N depending only on the dimension d such that the following holds: For every $x \in \mathbb{R}^d$ and every closed set Λ there are N nonoverlapping closed cones $V_1, \dots, V_N$ with vertex x such that, if ξ_i is a point of $\Lambda \cap V_i$ closest to x, then any point $y \in \Lambda$ with $y \neq x$ is no further to some ξ_i than to x.*

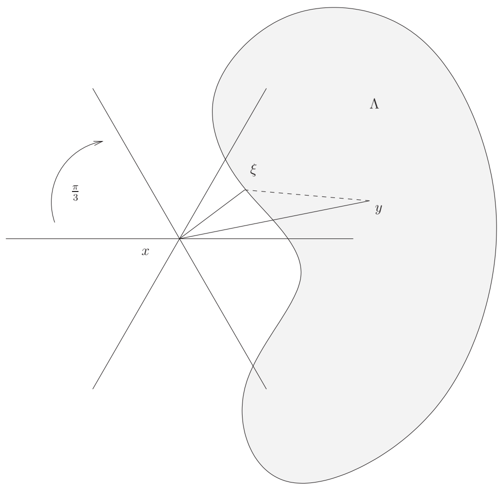

FIGURE 1. The geometric argument in Lemma 8.16.

Proof. The proof is elementary by looking at Figure 1: Let N be the number of closed cones with circular base, vertex in the origin and opening angle $\pi/3$ needed to cover $\mathbb{R}^d$. Let V be a shift of such a cone with vertex in x, ξ be a point in $V \cap \Lambda$ which is closest to x, and $y \in \Lambda$ be arbitrary. The triangle with vertices in x, ξ and y has at most angle $\pi/3$ at the vertex x, and hence by the geometry of triangles, the distance of y and ξ is no larger than the distance of y and x. $\blacksquare$

Proof of Theorem 8.14. Of course, only the implication $\Rightarrow$ needs proof. Take μ satisfying $U_\mu(x) \le \vartheta$ for all $x \in \Lambda$. Note that, by monotone convergence,

$$U_\mu(x) = \lim_{\delta \downarrow 0} \int_{|x-y|>\delta} K(x,y)\, d\mu(y). \tag{2.2}$$

Hence, for a given $\eta > 0$, by Egoroff's theorem, there exists a compact subset $F \subset \Lambda$ such that, $\mu(F) > \mu(\Lambda) - \eta$ and the convergence in (2.2) is uniform on F. If we define μ_1 to be the restriction of μ to F, then we can find, for every $\varepsilon > 0$ some $\delta > 0$ such that

$$\sup_{x \in F} \int_{|x-y| \le \delta} K(x,y)\, d\mu_1(y) < \varepsilon.$$

Now let $\{x_n\} \subset \mathbb{R}^d$ be a sequence converging to $x_0 \in F$. Then, as the kernel K is bounded on sets bounded away from the diagonal,

$$\limsup_{n\to\infty} U_{\mu_1}(x_n) \le \int K(x_0,y)\, d\mu_1(y) + \limsup_{n\to\infty} \int_{|y-x_n|\le\delta} K(x_n,y)\, d\mu_1(y).$$

We now want to compare $K(x_n,y)$ with $K(\xi,y)$ for $\xi \in F$. Here we use Lemma 8.16 for the point $x = x_n$ and obtain $\xi_1, \ldots, \xi_N \in F$ such that

$$K(x_n,y) \le \sum_{i=1}^N K(\xi_i,y),$$

where we have used that K depends only on the distance of the arguments and is decreasing in it. We thus have

$$\int_{|y-x_n|\le\delta} K(x_n,y)\, d\mu_1(y) \le \sum_{i=1}^N \int_{|y-\xi_i|\le\delta} K(\xi_i,y)\, d\mu_1(y) \le N\varepsilon.$$

As $\varepsilon > 0$ was arbitrary we get

$$\limsup_{n\to\infty} U_{\mu_1}(x_n) \le U_{\mu_1}(x_0).$$

As the converse statement

$$\liminf_{n\to\infty} U_{\mu_1}(x_n) \ge U_{\mu_1}(x_0)$$

holds trivially by Fatou's lemma, we obtain the continuity of U_{μ_1} on F. Continuity of U_{μ_1} on F^{c} is obvious from the properties of the kernel, so that we have continuity of U_{μ_1} on all of $\mathbb{R}^d$. By the maximum principle, Theorem 3.5, we infer that $U_{\mu_1}(x) \le \vartheta$.

To complete the proof let $x \notin \Lambda$ be arbitrary, and denote its distance to Λ by ϱ. Then $K(x,y) \le C(\varrho)$ for all $y \in \Lambda$. Therefore

$$U_\mu(z) \le U_{\mu_1}(z) + \eta C(\varrho) \le \vartheta + \eta\, C(\varrho),$$

and the result follows by letting $\eta \downarrow 0$. ∎

3. Polar sets and capacities

One of our ideas to measure the size of sets in Chapter 4 was based on the notion of capacity of the set. While this notion appeared to be useful, but maybe a bit artifical 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 ask, which sets are polar for the range of a Brownian motion $\{B(t) \colon t \geq 0\}$. Recall that a Borel set $A \subset \mathbb{R}^d$ is *polar* for Brownian motion if, for all x,

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

In the case $d = 1$ we already know that only the empty set is polar, whereas by Corollary 2.23 points are polar for Brownian motion in all dimensions $d \geq 2$. The general characterisation of polar sets requires an extension of the notion of capacities to a bigger class of kernels.

Definition 8.17. *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

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

Recall that the α-energy of a measure and the Riesz α-capacity Cap_α of a set defined in Chapter 4 correspond to the kernel $K(x, y) = |\,x - y|^{-\alpha}$. ◇

Remark 8.18. In most of our applications the kernels are of the form $K(x, y) = f(|\,x - y|)$ for some decreasing function $f \colon [0, \infty) \to [0, \infty]$. In this case we simply write I_f instead of I_K and call this the f-energy. We also write Cap_f instead of Cap_K and call this the f-capacity. ◇

Theorem 8.19 (Kakutani's theorem). *A closed set Λ is polar for d-dimensional Brownian motion if and only if it has zero f-capacity for the* **radial potential** *f defined by*

$$f(\varepsilon) := \begin{cases} \big|\log(1/\varepsilon)\big| & \textit{if } d = 2, \\ \varepsilon^{2-d} & \textit{if } d \geq 3. \end{cases}$$

Remark 8.20. We call the kernel $K(x, y) = f(|x - y|)$, where f is the radial potential, the **potential kernel**. Up to constants, it agrees with the Green kernel in $d \geq 3$. ◇

Instead of proving Kakutani's theorem directly, we aim for a stronger result, which gives, for compact sets $\Lambda \subset \mathbb{R}^d$, 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

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

for the radial potential f in Theorem 8.19. 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 modification of the Green kernel G, called the *Martin kernel*, which we now introduce.

We again look at all transient Brownian motions, recall the setup from Section 3.2. All we need is the following property, which is easy to verify directly from the form of the Green's function G in case (1). For the other two cases we give a conceptual proof.

PROPOSITION 8.21. *For every compact set $\Lambda \subset D \subset \mathbb{R}^d$ there exists a constant C depending only on Λ such that, for all $x, y \in \Lambda$ and sufficiently small $\varepsilon > 0$,*

$$\sup_{|x-z|<\varepsilon} \varepsilon^{-d} \int_{\mathcal{B}(y,\varepsilon)} \frac{G(z,\xi)}{G(x,y)}\, d\xi \le C. \tag{3.1}$$

Proof. Fix a compact set $\Lambda \subset D$ and $\varepsilon > 0$ smaller than one tenth of the distance of Λ and D^{c} and let $x, y \in \Lambda$. We abbreviate

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

We first assume that $|x-y| > 4\varepsilon$ and show that in this case

$$\sup_{|x-\widetilde{x}|<\varepsilon} \sup_{|y-\widetilde{y}|<\varepsilon} G(\widetilde{x},\widetilde{y}) \le C G(x,y), \tag{3.2}$$

from which our claim easily follows.

The function $G(\cdot\,,y)$ is harmonic on $D\setminus\{y\}$. Hence, with $\tau = \inf\{0 < t \le T \,:\, B(t) \notin \mathcal{B}(x,2\varepsilon)\}$ we note that, for all $\widetilde{x} \in \mathcal{B}(x,\varepsilon)$,

$$G(\widetilde{x},y) = \mathbb{E}_{\widetilde{x}}\big[G(B(\tau),y),\, \tau \le T\big].$$

This is the average of $G(\cdot\,,y)$ with respect to the harmonic measure $\mu_{\partial\mathcal{B}(x,2\varepsilon)}(\widetilde{x},\,\cdot\,)$. This measure has a density with respect to the uniform measure on the sphere $\partial\mathcal{B}(x,2\varepsilon)$, which is bounded from zero and infinity by absolute constants. In the cases (1),(3) this can be seen directly from Poisson's formula. Therefore $G(\widetilde{x},y) \le CG(x,y)$ and repetition of this argument, introducing now $\widetilde{y} \in \mathcal{B}(y,\varepsilon)$ and fixing $\widetilde{x}$ gives the claim.

Now look at the case $|x-y| \le 4\varepsilon$. We first observe that, for some constant $c > 0$, $G(x,y) \ge c\varepsilon^{2-d}$, which is obvious in all cases. Now let $z \in \mathcal{B}(x,\varepsilon)$. Decomposing the Brownian path on its first exit time τ from $\mathcal{B}(x,8\varepsilon)$ and denoting the uniform distribution on $\partial\mathcal{B}(x,8\varepsilon)$ by ϖ we obtain for constants $C_1, C_2 > 0$,

$$\begin{aligned} h_\varepsilon(z,y) &\le \mathbb{E}[\tau \wedge T] + \mathbb{E}_z\big[h_\varepsilon(B(\tau),y),\, \tau \le T\big] \\ &\le C_1\varepsilon^2 + C_2\varepsilon^d \int G(w,y)\, d\varpi(w), \end{aligned}$$

where we have used (3.2). A similar decomposition gives that $\int G(w,y)\,d\varpi(w) \le C_3 G(x,y)$ and putting all facts together gives $h_\varepsilon(z,y) \le C_4 \varepsilon^d G(x,y)$, as required. ■

Definition 8.22. *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 shows that (in all three cases of transient Brownian motions) Martin capacity is indeed a good estimate of the hitting probability.

Theorem 8.23. *Let $\{B(t) : 0 \le t \le T\}$ be a transient Brownian motion and $A \subset D$ closed. Then*

$$\tfrac{1}{2}\,\mathrm{Cap}_M(A) \le \mathbb{P}_0\{\exists 0 < t \le T \text{ such that } B(t) \in A\} \le \mathrm{Cap}_M(A) \tag{3.3}$$

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

$$\mu(A) = \mathbb{P}_0\{\tau \le T\} = \mathbb{P}_0\{B(t) \in A \text{ for some } 0 < t \le T\}. \tag{3.4}$$

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

$$\mathbb{E}_0 \int_0^T \mathbb{1}\{|B(t) - y| < \varepsilon\}\,dt = \int_{\mathcal{B}(y,\varepsilon)} G(0,z)\,dz. \tag{3.5}$$

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

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

Integrating over t and using Fubini's theorem yields

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

Combining this with (3.5) we infer that

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

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

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

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

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

which by (3.4) 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 $D^c \cup \{0\}$. 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

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

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

$$\begin{aligned}
\mathbb{E}_0 Z_\varepsilon^2 &= 2\mathbb{E}_0 \int_0^T ds \int_s^T dt \iint \frac{\mathbb{1}\{B(s) \in \mathcal{B}(x,\varepsilon),\ B(t) \in \mathcal{B}(y,\varepsilon)\}}{h_\varepsilon(0,x)h_\varepsilon(0,y)}\,d\nu(x)\,d\nu(y) \\
&\le 2\mathbb{E}_0 \iint \int_0^T ds\,\mathbb{1}\{B(s) \in \mathcal{B}(x,\varepsilon)\} \frac{h^*_\varepsilon(x,y)}{h_\varepsilon(0,x)h_\varepsilon(0,y)}\,d\nu(x)\,d\nu(y) \qquad (3.6) \\
&= 2 \iint \frac{h^*_\varepsilon(x,y)}{h_\varepsilon(0,y)}\,d\nu(x)\,d\nu(y).
\end{aligned}$$

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 $\lim_{\varepsilon\downarrow 0} \mathcal{L}(\mathcal{B}(0,\varepsilon))^{-1}\,h_\varepsilon(0,y) = G(0,y)$. Moreover, by (3.1) and the fact that $G(0,y)$ is bounded from zero and infinity for all $y \in A$, for some constant C,

$$\frac{h^*_\varepsilon(x,y)}{h_\varepsilon(0,y)} \le C\,\frac{G(x,y)}{G(0,y)} = C\,M(x,y).$$

Hence, if ν is a measure of finite energy, we can use dominated convergence and obtain,

$$\lim_{\varepsilon\downarrow 0} \mathbb{E} Z_\varepsilon^2 \le 2 \iint \frac{G(x,y)}{G(0,y)}\,d\nu(x)\,d\nu(y) = 2I_M(\nu). \qquad (3.7)$$

Clearly, the hitting probability $\mathbb{P}\{\exists t > 0, y \in A \text{ such that } B(t) \in \mathcal{B}(y,\varepsilon)\}$ is at least

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

where we have used the Paley-Zygmund inequality in the second step. 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 (3.7),

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

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

Remark 8.24. The right-hand inequality in (3.3) can be an equality: look at the case $d=3$, $T=\infty$, our case (1), and take a sphere in $\mathbb{R}^d$ centred at the origin, which has hitting probability and capacity both equal to one. Exercise 8.5 shows that the constant 1/2 on the left cannot be increased. ◇

Proof of Theorem 8.19. It suffices, by taking countable unions, to consider compact sets Λ which have positive distance from the origin.

First consider the case $d=3$. Then $G(0,x)$ is bounded away from zero and infinity. Hence the set Λ is polar if and only if its f-capacity vanishes, where $f(\varepsilon)=\varepsilon^{d-2}$.

In the case $d=2$ we choose a ball $\mathcal{B}(0,R)$ containing Λ. By Exercise 3.13 the Green's function for Brownian motion stopped upon leaving $\mathcal{B}(0,R)$ is given as

$$G^{(R)}(x,y)=\begin{cases} -\frac{1}{\pi}\log|x/R-y/R|+\frac{1}{\pi}\log\left|\frac{x}{|x|}-|x|yR^{-2}\right|, & \text{if } x\neq 0,\ x,y\in\mathcal{B}(0,R),\\ -\frac{1}{\pi}\log|y/R| & \text{if } x=0,\ y\in\mathcal{B}(0,R).\end{cases}$$

Again $G^{(R)}(0,y)$ is bounded over all $y\in\Lambda$, and so is the second summand of $G^{(R)}(x,y)$. Hence only the contribution from $-\log|x-y|$ decides about finiteness of the energy of a probability measure. Therefore, any measure with finite Martin energy has finite f-energy for $f(\varepsilon)=-\log(1/\varepsilon)$, and vice versa. This completes the proof. ■

The estimates in Theorem 8.23 are valid beyond the Brownian motion case. The following proposition, which has a very similar proof to Theorem 8.23, shows that one has an analogous result in a discrete setup. We will see a surprising application of this in Chapter 9.

Proposition 8.25. *Let $\{X_n : n\in\mathbb{N}\}$ be a transient Markov chain on a countable state space S, and set*

$$G(x,y)\;=\;\mathbb{E}_x\left[\sum_{n=0}^{\infty}\mathbb{1}_{\{y\}}(X_n)\right] \quad\textit{and}\quad M(x,y)\;=\;\frac{G(x,y)}{G(\rho,y)}\,.$$

Then for any initial state ρ and any subset Λ of S,

$$\tfrac{1}{2}\,\mathrm{Cap}_M(\Lambda)\leq\mathbb{P}_\rho\big\{\{X_n : n\in\mathbb{N}\}\textit{ hits }\Lambda\big\}\leq\mathrm{Cap}_M(\Lambda)\,.$$

Proof. To prove the right-hand inequality, we may assume that the hitting probability is positive. Let $\tau=\inf\{n : X_n\in\Lambda\}$ and let ν be the measure $\nu(A)=\mathbb{P}_\rho\{\tau<\infty \text{ and } X_\tau\in A\}$.

In general, ν is a sub-probability measure, as τ may be infinite. By the Markov property, for $y \in \Lambda$,

$$\int\limits_{\Lambda} G(x,y)\,d\nu(x) = \sum_{x\in\Lambda} \mathbb{P}_\rho\{X_\tau = x\}\,G(x,y) = G(\rho,y)\,,$$

whence $\int_\Lambda M(x,y)\,d\nu(x) = 1$. Therefore $I_M(\nu) = \nu(\Lambda)$, $I_M\big(\nu/\nu(\Lambda)\big) = [\nu(\Lambda)]^{-1}$; consequently, since $\nu/\nu(\Lambda)$ is a probability measure,

$$\mathrm{Cap}_M(\Lambda) \;\geq\; \nu(\Lambda) \;=\; \mathbb{P}_\rho\big\{\{X_n\} \text{ hits } \Lambda\big\}\,.$$

This yields one inequality. Note that the Markov property was used here.

For the reverse inequality, we use the second moment method. Given a probability measure μ on Λ, set

$$Z = \int_\Lambda \sum_{n=0}^{\infty} \mathbb{1}_{\{y\}}(X_n)\frac{d\mu(y)}{G(\rho,y)}\,.$$

$\mathbb{E}_\rho[Z] = 1$, and the second moment satisfies

$$\begin{aligned}\mathbb{E}_\rho[Z^2] &= \mathbb{E}_\rho \int_\Lambda\int_\Lambda \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \mathbb{1}_{\{x\}}(X_m)\mathbb{1}_{\{y\}}(X_n)\frac{d\mu(x)d\mu(y)}{G(\rho,x)G(\rho,y)}\\ &\leq 2\mathbb{E}_\rho \int_\Lambda\int_\Lambda \sum_{m\leq n} \mathbb{1}_{\{x\}}(X_m)\mathbb{1}_{\{y\}}(X_n)\frac{d\mu(x)d\mu(y)}{G(\rho,x)G(\rho,y)}\,.\end{aligned}$$

Observe that

$$\sum_{m=0}^{\infty}\mathbb{E}_\rho\sum_{n=m}^{\infty} \mathbb{1}_{\{x\}}(X_m)\mathbb{1}_{\{y\}}(X_n) = \sum_{m=0}^{\infty}\mathbb{P}_\rho\{X_m = x\}\,G(x,y) = G(\rho,x)G(x,y)\,.$$

Hence

$$\mathbb{E}_\rho[Z^2] \leq 2\int_\Lambda\int_\Lambda \frac{G(x,y)}{G(\rho,y)}\,d\mu(x)\,d\mu(y) = 2I_M(\mu)\,,$$

and therefore

$$\mathbb{P}_\rho\big\{\{X_n\} \text{ hits } \Lambda\big\} \geq \mathbb{P}_\rho\{Z>0\} \geq \frac{\big(\mathbb{E}_\rho[Z]\big)^2}{\mathbb{E}_\rho[Z^2]} \geq \frac{1}{2I_M(\mu)}\,.$$

We conclude that $\mathbb{P}_\rho\big\{\{X_n\} \text{ hits } \Lambda\big\} \geq \frac{1}{2}\mathrm{Cap}_M(\Lambda)$. ∎

Recall from Corollary 8.11 that we have already seen estimates for the probability that Brownian motion hits a set, which were given in terms of the total mass of the equilibrium measure. The following theorem reveals the relationship between the equilibrium measure and capacities.

Theorem 8.26. *Let $\Lambda \subset \mathbb{R}^d$ be a nonpolar, compact set and $G\colon \mathbb{R}^d \times \mathbb{R}^d \to [0,\infty]$ the Green's function of a transient Brownian motion. Then*

$$\mathrm{Cap}_G(\Lambda) = \Big\{I_G\Big(\frac{\nu}{\nu(\Lambda)}\Big)\Big\}^{-1} = \nu(\Lambda).$$

where ν is the equilibrium measure of Λ.

REMARK 8.27. If Λ is polar, we have $\mathrm{Cap}_G(\Lambda) = 0 = \nu(\Lambda)$. Otherwise, this shows that the probability measure $\nu/\nu(\Lambda)$ minimizes the G-energy over the set of all probability measures on Λ.

For the proof we first note that, for the Green's function G of a transient Brownian motion, the G-energy of a *signed* measure is always nonnegative.

LEMMA 8.28. *Let μ, ν finite measures on $\mathbb{R}^d$ and G the Green's function G of a transient Brownian motion. Then, for $\sigma = \mu - \nu$, we have*

$$\iint G(x,y)\, d\sigma(x)\, d\sigma(y) \geq 0.$$

Proof. From the Chapman-Kolmogorov equation we have

$$\mathfrak{p}^*(t,x,y) = \int \mathfrak{p}^*(t/2,x,z)\, \mathfrak{p}^*(t/2,z,y)\, dz.$$

Integrating with respect to $d\sigma(x)\, d\sigma(y)$ and using the symmetry of $\mathfrak{p}^*(t,\,\cdot\,,\,\cdot\,)$ gives

$$\iint \mathfrak{p}^*(t,x,y)\, d\sigma(x)\, d\sigma(y) = \int \Big(\int \mathfrak{p}^*(t/2,x,z)\, d\sigma(x) \Big)^2 dz \geq 0.$$

Integrating now over time gives the result. ∎

Proof of Theorem 8.26. Let ν be the equilibrium measure and define $\varphi(x) = \int G(x,y)\, d\nu(y)$. By the last exit formula, Theorem 8.9, $\varphi(x)$ is the probability that a Brownian motion started at x hits Λ before time T. Hence $\varphi(x) \leq 1$ for every x and, if x is a regular point for Λ, then $\varphi(x) = 1$. Also by the last exit formula, because irregular points are never hit by a Brownian motion, see Theorem 8.12, we have $\varphi(x) = 1$ for ν-almost every point. This implies that

$$I_G(\nu) = \int_\Lambda \varphi(x)\, d\nu(x) = \nu(\Lambda).$$

Suppose now that μ is an arbitrary measure on Λ with $\mu(\Lambda) = \nu(\Lambda)$ and assume that μ has finite energy. Note that μ does not charge the set of irregular points, as otherwise this set would have positive capacity and would be nonpolar by Theorem 8.23. Then, starting with Lemma 8.28 and using also the symmetry of G,

$$\begin{aligned} 0 &\leq \iint G(x,y)\, d(\nu-\mu)(x)\, d(\nu-\mu)(y) = I_G(\mu) + I_G(\nu) - 2 \iint G(x,y)\, d\nu(x)\, d\mu(y) \\ &= I_G(\mu) + \nu(\Lambda) - 2 \int_\Lambda \varphi(y)\, d\mu(y) \leq I_G(\mu) - \nu(\Lambda), \end{aligned}$$

using in the last step that $\varphi(y) = 1$ on the set of regular points, and thus μ-almost everywhere. This implies that $I_G(\mu) \geq \nu(\Lambda) = I_G(\nu)$, so that $\nu/\nu(\Lambda)$ is a minimiser in the definition of Cap_G. This completes the proof. ∎

4. Wiener's test of regularity

In this section we concentrate on $d \geq 3$ and find a sharp criterion for a point to be regular for a closed set $\Lambda \subset \mathbb{R}^d$. This criterion is given in terms of the capacity of the intersection of Λ with annuli, or shells, concentric about x.

To fix some notation let $k > \ell$ be integers and $x \in \mathbb{R}^d$, and define the annulus

$$A_x(k,\ell) := \big\{ y \in \mathbb{R}^d \,:\, 2^{-k} \leq |y-x| \leq 2^{-\ell} \big\}.$$

Abbreviate $A_x(k) := A_x(k+1,k)$ and let

$$\Lambda_x^k := \Lambda \cap A_x(k)\,.$$

We aim to prove the following result.

THEOREM 8.29 (Wiener's test). *A point $x \in \mathbb{R}^d$ is regular for the closed set $\Lambda \subset \mathbb{R}^d$, $d \geq 3$, if and only if*

$$\sum_{k=1}^{\infty} 2^{k(d-2)} C_{d-2}\big(\Lambda_x^k\big) = \infty\,,$$

where C_{d-2} is the Riesz $(d-2)$-capacity introduced in Definition 4.3.

In the proof, we may assume, without loss of generality, that $x = 0$. We start the proof with an easy observation.

LEMMA 8.30. *There exists a constant $c > 0$, which depends only on the dimension d, such that, for all k, we have*

$$c\, 2^{k(d-2)} C_{d-2}(\Lambda_0^k) \leq \mathrm{Cap}_M(\Lambda_0^k) \leq c\, 2^{(k+1)(d-2)} C_{d-2}(\Lambda_0^k).$$

Proof. Observe that, as $z \in A_0^k$ implies $2^{-k-1} \leq |z| \leq 2^{-k}$, we obtain the statement by estimating the denominator in the Martin kernel M. ■

The crucial step in the proof is a quantitative estimate, from which Wiener's test follows quickly.

LEMMA 8.31. *There exists a constant $c > 0$, depending only on the dimension d, such that*

$$1 - \exp\Big(-c \sum_{j=\ell}^{k-1} \mathrm{Cap}_M(\Lambda_0^j)\Big) \leq \mathbb{P}_0\big\{\{B(t)\,:\, t \geq 0\} \text{ hits } \Lambda \cap A_0(k,\ell)\big\} \leq \sum_{j=\ell}^{k-1} \mathrm{Cap}_M(\Lambda_0^j)\,.$$

Proof. For the *upper* bound we look at the event $D(j)$ that a Brownian motion started in 0 hits Λ_0^j. Then, using Theorem 8.23, we get $\mathbb{P}_0\big(D(j)\big) \leq \mathrm{Cap}_M\big(\Lambda_0^j\big)$. Therefore

$$\mathbb{P}_0\big\{\{B(t)\,:\, t \geq 0\} \text{ hits } \Lambda \cap A_0(k,\ell)\big\} \leq \mathbb{P}_0\Big(\bigcup_{j=\ell}^{k-1} D(j)\Big) \leq \sum_{j=\ell}^{k-1} \mathrm{Cap}_M(\Lambda_0^j),$$

and this completes the proof of the upper bound.

For the *lower* bound we look at the event $E(z,j)$ that a Brownian motion started in some point $z \in \partial\mathcal{B}(0,2^{-j})$ and stopped upon hitting $\partial\mathcal{B}(0,2^{-j+4})$ hits Λ_0^{j-2}. Again we use either

Theorem 8.23, or Corollary 8.11 in conjunction with Theorem 8.26, to get, for constants $c_1, c_2 > 0$ depending only on the dimension d,

$$\mathbb{P}_z\big\{\{B(t) \,:\, t \geq 0\} \text{ ever hits } \Lambda_0^{j-2}\big\} \geq c_1 \left(2^{-(j-1)} - 2^{-j}\right)^{2-d} C_{d-2}\big(\Lambda_0^{j-2}\big),$$

and, for any $y \in \partial\mathcal{B}(0, 2^{-j+4})$,

$$\mathbb{P}_y\big\{\{B(t) \,:\, t \geq 0\} \text{ ever hits } \Lambda_0^{j-2}\big\} \leq c_2 \left(2^{-(j-4)} - 2^{-(j-2)}\right)^{2-d} C_{d-2}\big(\Lambda_0^{j-2}\big).$$

Therefore, for a constant $c > 0$ depending only on the dimension d,

$$\begin{aligned}\mathbb{P}\big(E(z,j)\big) &\geq \mathbb{P}_z\big\{\{B(t)\} \text{ ever hits } \Lambda_0^{j-2}\big\} - \max_{y \in \partial\mathcal{B}(0,2^{-j+4})} \mathbb{P}_y\big\{\{B(t)\} \text{ ever hits } \Lambda_0^{j-2}\big\} \\ &\geq c\, 2^{j(d-2)} C_{d-2}\big(\Lambda_0^{j-2}\big).\end{aligned}$$

Now divide $\{\ell, \ldots, k-1\}$ into four subsets such that each subset I satisfies $|i-j| \geq 4$ for all $i \neq j \in I$. Choose a subset I which satisfies

$$\sum_{j \in I} 2^{j(d-2)} C_{d-2}(\Lambda_0^j) \geq \tfrac{1}{4} \sum_{j=\ell}^{k-1} 2^{j(d-2)}\, C_{d-2}(\Lambda_0^j). \tag{4.1}$$

Now we have with $\tau_j = \inf\{t \geq 0 \,:\, |B(t)| = 2^{-j}\}$,

$$\begin{aligned}\mathbb{P}_0\big\{\ \{B(t) \,: t \geq 0\} \text{ avoids } \Lambda \cap A_0(k,\ell)\big\} &\leq \mathbb{P}_0\Big(\bigcap_{j \in I} E\big(B(\tau_j), j\big)^{\mathrm{c}}\Big) \\ &\leq \prod_{j \in I} \sup_{z \in \partial\mathcal{B}(0,2^{-j})} \mathbb{P}\big(E(z,j)^{\mathrm{c}}\big) \leq \prod_{j \in I} \Big(1 - c\, 2^{j(d-2)} C_{d-2}\big(\Lambda_0^{j-2}\big)\Big) \\ &\leq \exp\Big(-c \sum_{j \in I} 2^{j(d-2)}\, C_{d-2}(\Lambda_0^j)\Big),\end{aligned}$$

using the estimate $\log(1-x) \leq -x$ in the last step. The lower bound, with $c = C/4$, now follows from (4.1) and Lemma 8.30 when we pass to the complement. ■

Proof of Wiener's test. Suppose $\sum_{k=1}^{\infty} 2^{k(d-2)} C_{d-2}\big(\Lambda_0^k\big) = \infty$. Therefore, by Lemma 8.31 and Lemma 8.30, for all $k \in \mathbb{N}$,

$$\mathbb{P}_0\big\{\{B(t) \,: t \geq 0\} \text{ hits } \Lambda \cap \mathcal{B}(0, 2^{-k})\big\} \geq 1 - \exp\Big(-c \sum_{j=k}^{\infty} \mathrm{Cap}_M(\Lambda_0^j)\Big) = 1.$$

Since points are polar, for any $\varepsilon, \delta > 0$ there exists a large k such that

$$\mathbb{P}_0\big\{\{B(t) \,:\, t \geq \varepsilon\} \text{ hits } \mathcal{B}(0, 2^{-k})\big\} < \delta\,.$$

Combining these two facts we get for the first hitting time $\tau = \tau(\Lambda)$ of the set Λ,

$$\begin{aligned}\mathbb{P}_0\big\{\tau < \varepsilon\big\} &\geq \mathbb{P}_0\big\{\{B(t) \,: t \geq 0\} \text{ hits } \Lambda \cap \mathcal{B}(0, 2^{-k})\big\} - \mathbb{P}_0\big\{\{B(t) \,:\, t \geq \varepsilon\} \text{ hits } \mathcal{B}(0, 2^{-k})\big\} \\ &\geq 1 - \delta.\end{aligned}$$

As $\varepsilon, \delta > 0$ were arbitrary, the point 0 must be regular.

Now suppose that $\sum_{k=1}^{\infty} 2^{k(d-2)} C_{d-2}\big(\Lambda_0^k\big) < \infty$. Then

$$\sum_{k=1}^{\infty} \mathbb{P}_0\big\{\{B(t) : t \geq 0\} \text{ hits } \Lambda \cap A_0(k)\big\} \leq \sum_{k=1}^{\infty} \mathrm{Cap}_M(\Lambda_0^k) < \infty.$$

Hence, by the Borel Cantelli lemma, almost surely there exists a ball $\mathcal{B}(0,\varepsilon)$ such that $\{B(t) : t \geq 0\}$ does not hit $\mathcal{B}(0,\varepsilon) \cap \Lambda$. By continuity we therefore must have $\inf\{t > 0\colon B(t) \in \Lambda\} > 0$ almost surely, hence the point 0 is irregular. ■

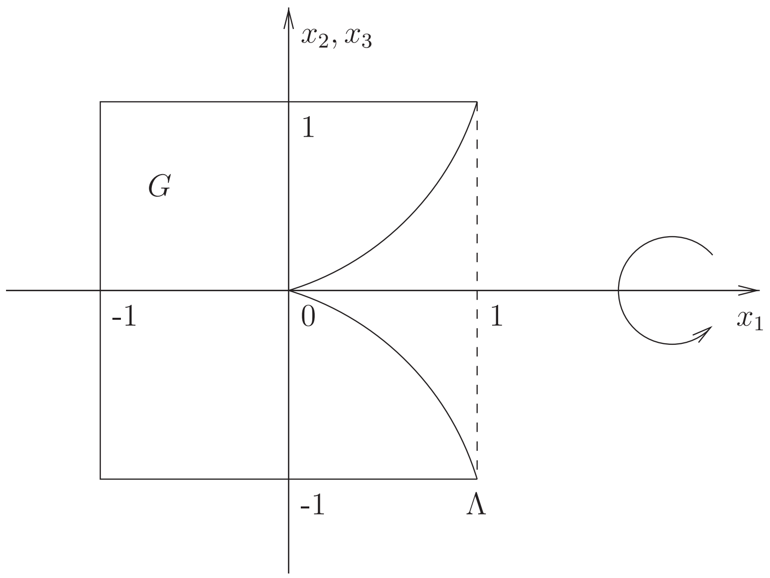

FIGURE 2. Lebesgue's thorn.

EXAMPLE 8.32. The following example is due to Lebesgue [**Le24**], and is usually called **Lebesgue's thorn**. For any $\alpha > 1$ we define an open subset $G \subset (-1,1)^3$ with a *cusp* at zero by

$$G := \Big\{(x_1,x_2,x_3) \in (-1,1)^3 \colon \sqrt{x_2^2 + x_3^2} > f(x_1) \text{ if } x_1 \geq 0\Big\}\,,$$

with $f(x) = x^\alpha$, see Figure 2. Now the origin is an *irregular* point for $\Lambda = G^{\mathrm{c}}$. For the proof it suffices, by Wiener's test, to check that the series $\sum_{k=1}^{\infty} 2^k C_1\big(\Lambda_0^k\big)$ converges. Note that, for any probability measure μ on Λ_0^k, we have $I_1(\mu) \geq 2^{\alpha k}$ and, hence,

$$\sum_{k=1}^{\infty} 2^k C_1\big(\Lambda_0^k\big) \leq \sum_{k=1}^{\infty} 2^{k(1-\alpha)} < \infty\,,$$

verifying Wiener's test of irregularity. ◇

Exercises

Exercise 8.1 (∗). Let $U \subset \mathbb{R}^d$ be a domain and $u\colon U \to \mathbb{R}$ subharmonic. Use Itô's formula to show that, for any ball $\mathcal{B}(x,r) \subset U$,

$$u(x) \le \frac{1}{\mathcal{L}(\mathcal{B}(x,r))} \int_{\mathcal{B}(x,r)} u(y)\, dy.$$

Exercise 8.2 (∗). Suppose g is bounded and u a solution of Poisson's problem for g. Show that this solution has the form

$$u(x) = \mathbb{E}_x\Big[\int_0^T g(B(t))\, dt\Big], \qquad \text{for } x \in U\,,$$

where $T := \inf\{t > 0 \,:\, B(t) \notin U\}$. Observe that his implies that the solution, if it exists, is always uniquely determined.

Exercise 8.3. Let

$$u(x) = \mathbb{E}_x\Big[\int_0^T g(B(t))\, dt\Big], \qquad \text{for } x \in U\,,$$

where $T := \inf\{t > 0 \,:\, B(t) \notin U\}$. Show that,

- if g is Hölder continuous, then the function $u : U \to \mathbb{R}$ solves $-\frac{1}{2}\Delta u = g$.
- if every $x \in \partial U$ is regular for the complement of U, then $u(x) = 0$ for all $x \in \partial U$.

Exercise 8.4. Suppose $\Lambda \subset \mathbb{R}^d$, for $d \ge 3$, is compact and γ the last exit time from Λ defined as in Theorem 8.9. Show that

$$\lim_{x\to\infty} \mathbb{P}_x\big\{B(\gamma) \in A \mid \gamma > 0\big\} = \frac{\nu(A)}{\nu(\Lambda)}.$$

Exercise 8.5. For $d \ge 3$ consider the spherical shell

$$\Lambda_R = \{x \in \mathbb{R}^d \,:\, 1 \le |x| \le R\}.$$

Show that $\lim_{R\to\infty} \mathrm{Cap}_M(\Lambda_R) = 2$.

Exercise 8.6. Let $\{X(a)\colon a \geq 0\}$ be a stable subordinator of index $\frac{1}{2}$ and

$$K(s,t) := \begin{cases} (t-s)^{-1/2} & 0 < s \leq t\,, \\ 0 & s > t > 0\,. \end{cases}$$

Let $M(s,t) = K(s,t)/K(0,t)$, then for any subset Λ of $[0,\infty)$,

$$\tfrac{1}{2}\,\mathrm{Cap}_K(\Lambda) \leq \mathbb{P}_0\big\{\{X(a)\,:\,a \geq 0\} \text{ hits } \Lambda\big\} \leq \mathrm{Cap}_K(\Lambda)\,.$$

Notes and Comments

The proof of the last exit formula is taken from Chung's beautiful paper [**Ch73**], but the existence of an energy minimizing measure is a much older fact. For the case of the Newtonian potential ($d = 3$) it was determined by Gauss as the charge distribution on the surface of a conductor which minimises the electrostatic energy. Classically, the equilibrium measure is defined as the measure ν on Λ that maximizes $\nu(\Lambda)$ among those with potential bounded by one. Then $\nu/\nu(\Lambda)$ is the energy minimizing probability measure, see [**Ca67**]. Rigorous results and extensions to general Riesz-potentials are due to Frostman in his ground-breaking thesis [**Fr35**]. Our discussion of the strong maximum principle follows Carleson [**Ca67**]. Bass [**Ba95**] describes an alternative approach.

Characterising the polar sets for Brownian motion is related to the following question: for which sets $A \subset \mathbb{R}^d$ are there nontrivial bounded harmonic functions on $\mathbb{R}^d \setminus A$? Such sets are called *removable* for bounded harmonic functions. Consider the simplest case first. When A is the empty set, it is obviously polar, and by Liouville's theorem there is no bounded harmonic function on its complement. Nevanlinna (about 1920) proved that for $d \geq 3$ there exist non-constant bounded harmonic functions on $\mathbb{R}^d \setminus A$ if and only if $\mathrm{Cap}_G(A) > 0$, where $G(x, y) = f(|x-y|)$ for the radial potential f as before. Just to make this result more plausible, note that the function $h(x) = \int G(x, y)\mu(dy)$, where μ is a measure on A of finite G-energy, would make a good candidate for such a function, see Theorem 3.34.

Loosely speaking, G-capacity measures whether a set A is big enough to hide a pole of a harmonic function inside. Recall from Theorem 4.36 that $\dim A > d - 2$ implies existence of such functions, and $\dim A < d - 2$ implies nonexistence. Kakutani (1944) showed that there exist bounded harmonic functions on $\mathbb{R}^d \setminus A$ if and only if A is polar for Brownian motion. Kakutani's theorem is proved in [**Ka44a**]. The precise hitting estimates we give are fairly recent, our proof is a variant of the original proof by Benjamini, Pemantle and Peres in [**BPP95**]. Proposition 8.25 goes back to the same paper.

An interesting question is which subsets of a compact sets are charged by the harmonic measure μ_A. Clearly μ_A does not charge polar sets, and in particular, in $d \geq 3$, we have $\mu_A(B) = 0$ for all Borel sets with $\dim(B) < d - 2$. In the plane, by a famous theorem of Makarov, see [**Ma85**], we have that

- any set B of dimension < 1 has $\mu_A(B) = 0$,
- there is a set $S \subset A$ with $\dim S = 1$ such that $\mu_A(S^{\mathrm{c}}) = 0$.

However, the outer boundary, which supports the harmonic measure, may have a dimension much bigger than one. An interesting question arising in the context of self-avoiding curves asks for the dimension of the outer boundary of the image $B[0, 1]$ of a Brownian motion. Based on scaling arguments from polymer physics, Benoit Mandelbrot conjectured in 1982 that this set should have fractal dimension 4/3. Bishop, Jones, Pemantle, and Peres [**BJPP97**] showed that that the outer boundary has dimension > 1. In 2001 Mandelbrot's conjecture was finally proved by Lawler, Schramm and Werner [**LSW01**].

CHAPTER 9

Intersections and self-intersections of Brownian paths

In this chapter we study multiple points of d-dimensional Brownian motion. We shall see, for example, in which dimensions the Brownian path has double points and explore how many double points there are. This chapter also contains some of the highlights of the book: a proof that planar Brownian motion has points of infinite multiplicity, the intersection equivalence of Brownian motion and percolation limit sets, and the surprising dimension-doubling theorem of Kaufman.

1. Intersection of paths: existence and Hausdorff dimension

1.1. Existence of inetrsections. Suppose that $\{B_1(t) : t \geq 0\}$ and $\{B_2(t) : t \geq 0\}$ are two independent d-dimensional Brownian motions started in arbitrary points. The question we ask in this section is, in which dimensions the ranges, or paths, of the two motions have a nontrivial intersection, in other words whether there exist times $t_1, t_2 > 0$ such that $B_1(t_1) = B_2(t_2)$. As this question is trivial if $d = 1$ we assume $d \geq 2$ throughout this section.

We have developed the tools to decide this question in Chapter 4 and Chapter 8. Keeping the path $\{B_1(t) : t \geq 0\}$ fixed, we have to decide whether it is a polar set for the second Brownian motion. By Kakutani's theorem, Theorem 8.19, this question depends on its capacity with respect to the potential kernel. As the capacity is again related to Hausdorff measure and dimension, the results of Chapter 4 are crucial in the proof of the following result.

THEOREM 9.1.

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

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

REMARK 9.2. In the case $d \leq 3$, if the Brownian paths are started in the same point, then almost surely, the paths intersect before any positive time $t > 0$, see Exercise 9.1 (a). $\diamond$

Proof of (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 $d \geq 4$, the capacity with respect to the potential kernel is a multiple of the Riesz $(d-2)$-capacity. By Theorem 4.27 this capacity is zero for sets of finite $(d-2)$-dimensional Hausdorff measure. Now note that if $d \geq 5$ 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 from (1.1) in Chapter 4 that $\mathcal{H}^2(B[0,1]) < \infty$ almost surely, which implies that $\mathrm{Cap}_2(B[0,1]) = 0$ by Theorem 4.27. This implies that an independent Brownian motion almost surely does not hit either of the segments $B[n, n+1]$, and therefore avoids the path entirely. ■

Proof of (b). If $d = 3$, the capacity with respect to the potential kernel is a multiple of the Riesz 1-capacity. As the Hausdorff dimension of a path is two, this capacity is positive by Theorem 4.36. Therefore two Brownian paths in $d = 3$ intersect with positive probability.

Suppose now the two Brownian motions start in different points, we may assume that one is the origin and the other one is denoted x. By rotational invariance, the probability that the paths do not intersect depends only on $|x|$, and by Brownian scaling we see that it is completely independent of the choice of $x \neq 0$. Denote this probability by q and, given any $\varepsilon > 0$, choose a large time t such that

$$\mathbb{P}\big\{B_1(t_1) \neq B_2(t_2) \text{ for all } 0 < t_1, t_2 \leq t\big\} \leq q + \varepsilon.$$

Then, using the Markov property,

$$\begin{aligned} q &\leq \mathbb{P}\big\{B_1(t_1) \neq B_2(t_2) \text{ for all } t_1, t_2 \leq t\big\}\, \mathbb{P}\big\{B_1(t_1) \neq B_2(t_2) \text{ for all } t_1, t_2 > t\big\} \\ &\leq q(q + \varepsilon). \end{aligned}$$

As $\varepsilon > 0$ was arbitrary, we get $q \leq q^2$, and as we know that $q < 1$ we obtain that $q = 0$. This shows that two Brownian paths started in different points intersect almost surely. If they start in the same point, by the Markov property,

$$\mathbb{P}\big\{B_1(t_1) \neq B_2(t_2) \text{ for all } t_1, t_2 > 0\big\} = \lim_{\substack{t \downarrow 0 \\ t > 0}} \mathbb{P}\big\{B_1(t_1) \neq B_2(t_2) \text{ for all } t_1, t_2 > t\big\} = 1,$$

as required to complete the argument in the case $d = 3$. A path in $d \leq 2$ is the projection of a three dimensional path on a lower dimensional subspace, hence if two paths in $d = 3$ intersect almost surely, then so do two paths in $d = 2$. ■

It is equally natural to ask, for integers $p > 2$ and $d \leq 3$, whether a collection of p independent d-dimensional Brownian motions

$$\{B_1(t) : t \geq 0\}, \ldots, \{B_p(t) : t \geq 0\}$$

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

THEOREM 9.3.

(a) *For $d \geq 3$, 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 S 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 9.4. *Suppose S is the intersection of the ranges of two 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*

$$\mathcal{H}^1(S \cap \Lambda) < \infty.$$

Proof. Fix a cube $\mathsf{Cube} \subset \mathbb{R}^3$ of unit sidelength not containing the starting points. It suffices to show that, almost surely, $\mathcal{H}^1(S \cap \mathsf{Cube}) < \infty$. For this purpose let $\mathfrak{C}_n$ be the collection of dyadic subcubes of Cube of sidelength 2^{-n}, and $\mathfrak{I}_n$ be the collection of cubes in $\mathfrak{C}_n$ which are hit by both motions. By our hitting estimates, Theorem 3.17, there exists $C > 0$ such that, for any cube $E \in \mathfrak{C}_n$,

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

Now, for every n, the collection $\mathfrak{I}_n$ is a covering of S, and

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

Therefore, by Fatou's lemma, we obtain

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

Hence, almost surely, for arbitrarily large values of n we have a covering of $S \cap \mathsf{Cube}$ by sets of diameter at most $\sqrt{3}2^{-n}$ with 1-value no more than $C\sqrt{3}$. We infer from this that $\mathcal{H}^1(S \cap \mathsf{Cube}) \leq C\sqrt{3}$ almost surely, and this completes the proof. ∎

Proof of Theorem 9.3 (a). It suffices to show that, for any cube Cube of unit sidelength, which does not contain the origin, we have $\mathrm{Cap}_1(S \cap \mathsf{Cube}) = 0$. This follows directly from Lemma 9.4 and the energy method, Theorem 4.27. ∎

For Theorem 9.3 (b) it would suffice to show that the Hausdorff dimension of the set

$$S = \big\{ x \in \mathbb{R}^d \,:\, \exists t_1, t_2 > 0 \text{ such that } x = B_1(t_1) = B_2(t_2) \big\}.$$

is positive in the case $d = 2$. In fact, it is a natural question to ask for the Hausdorff dimension of the intersection of Brownian paths in any case when the set is nonempty. The problem was raised by Itô and McKean in their influential book [**IM74**], and has since been resolved by Taylor [**Ta66**] and Fristedt [**Fr67**]. The nontrivial 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.

1.2. Stochastic co-dimension and percolation limit sets. 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 positve. In the latter case this indicates that the set is large, and we should therefore get a lower bound on the dimension of A. A natural choice of such a random test set would be the range of Brownian motion. Recall that, for example in the case $d = 3$, if $\mathbb{P}\{\mathsf{Range} \cap A \neq \emptyset\} > 0$ is positive, this implies that $\dim A \geq 1$.

Of course, in order to turn this idea into a systematic technique for finding lower bounds for the Hausdorff dimension, an entire family of test sets is needed to tune the size of the test set in order to give sharp bounds. For this purpose, Taylor [**Ta66**] used stable processes instead of Brownian motion. This is not the easiest way and also limited, because stable processes only exist across a limited range of parameters. The approach we use in this book 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 sidelength 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 9.5 (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,*
(iii) *if* $\dim A > \gamma$, *then*
  (a) *almost surely* $\dim \big(A \cap \Gamma[\gamma]\big) \leq \dim A - \gamma$ *and,*
  (b) *for all* $\varepsilon > 0$, *with positive probability* $\dim \big(A \cap \Gamma[\gamma]\big) \geq \dim A - \gamma - \varepsilon$.

REMARK 9.6. Observe that the first part of the theorem gives a *lower* bound γ for the Hausdorff dimension of a set A, if we can show that $A \cap \Gamma[\gamma] \neq \emptyset$ with positive probability. As with so many ideas in fractal geometry one of the roots of this method lies in the study of trees, more precisely percolation on trees, see [**Ly90**]. ◇

REMARK 9.7.

(a) There is a close kinship of the stochastic co-dimension technique and the energy method: A set A is called *polar for the percolation limit set*, if
$$\mathbb{P}\{A \cap \Gamma[\gamma] \neq \emptyset\} = 0.$$
We shall see in Theorem 9.18 that a set is polar for the percolation limit set if and only if it has γ-capacity zero.

(b) For $d \geq 3$, the criterion for polarity of a percolation limit set with $\gamma = d - 2$ therefore agrees with the criterion for the polarity for Brownian motion. This 'equivalence' between percolation limit sets and Brownian motion has a quantitative strengthening which is discussed in Section 2 of this chapter. ◇

Proof of (i) in Hawkes' theorem. The proof of part (i) 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 $D_1, D_2, \ldots$ 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 $D_1, D_2, \ldots \in \mathfrak{C}$. Suppose that the sidelength of D_i is 2^{-n}, then the probability that $D_i \in \mathfrak{S}_n$ is $2^{-n\gamma}$. By picking from $D_1, D_2, \ldots$ 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
$$\mathbb{P}\{A \cap \Gamma[\gamma] \neq \emptyset\} \leq \mathbb{P}\{N > 0\} \leq \mathbb{E}N = \sum_{i=1}^{\infty} \mathbb{P}\{D_i \in \mathfrak{S}\} = \sum_{i=1}^{\infty} |D_i|^\gamma < \varepsilon.$$
As this holds for all $\varepsilon > 0$ we infer that, almost surely, we have $A \cap \Gamma[\gamma] = \emptyset$. ■

Proof of (ii) in Hawkes' theorem. The proof of part (ii) is based on the *second moment method*, which means that a variance has to be calculated. We also use the nontrivial part of Frostman's lemma in the form of Theorem 4.36, 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

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

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

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

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 sidelength 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

$$|\, x - y| \leq \sqrt{d} 2^{-m}.$$

This gives a handle to estimate the second moment in terms of the energy of μ. We find that

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

Plugging these moment estimates into the easy form of the Paley-Zygmund inequality, Lemma 3.22, gives $\mathbb{P}\{Y_n > 0\} \geq d^{-\gamma/2} I_\gamma(\mu)^{-1}$, as required. ∎

Proof of (iii) in Hawkes' theorem. For part (iii) note that the intersection $\Gamma[\gamma] \cap \Gamma[\delta]$ of two independent percolation limit sets has the same distribution as $\Gamma[\gamma + \delta]$. Suppose first that $\delta > \dim A - \gamma$. Then, by part (i), $A \cap \Gamma[\gamma] \cap \Gamma[\delta] = \emptyset$ almost surely, and hence, by part (ii), $\dim A \cap \Gamma[\gamma] \leq \delta$ almost surely. Letting $\delta \downarrow \dim A - \gamma$ completes the proof of part (a).

Now suppose that $\delta < \dim A - \gamma$. Then, with positive probability, $(A \cap \Gamma[\gamma]) \cap \Gamma[\delta] \neq \emptyset$, by part (ii). And using again part (i) we get that $\dim A \cap \Gamma[\gamma] \geq \delta$ with positive probability, completing the proof of part (b). ∎

1.3. Hausdorff dimension of intersections. We can now use the stochastic codimension approach to find the Hausdorff dimension of the intersection of two Brownian paths, whenever it is nonempty. Note that the following theorem also implies Theorem 9.3 (b).

THEOREM 9.8. *Suppose $d \geq 2$ and $p \geq 2$ are integers sucht that $p(d-1) < d$. Suppose that*

$$\{B_1(t) \,:\, t \geq 0\}, \ldots, \{B_p(t) \,:\, t \geq 0\}$$

are p independent d-dimensional Brownian motions. Let Range_i be the range of $\{B_i(t) \,:\, t \geq 0\}$ for $1 \leq i \leq p$. Then, almost surely,

$$\dim\big(\mathsf{Range}_1 \cap \ldots \cap \mathsf{Range}_p\big) = d - p(d-2).$$

REMARK 9.9. A good way to make this result plausible is by recalling the situation for the intersection of linear subspaces of $\mathbb{R}^d$: If the spaces are in general position, then the co-dimension of the intersection is the sum of the co-dimensions of the subspaces. As the Hausdorff dimension of a Brownian path is two, the plausible codimension of the intersection of p paths is $p(d-2)$, and hence the dimension is $d - p\ (d-2)$. $\diamond$

REMARK 9.10. Under assumption of the theorem, if the Brownian paths are started in the same point, then almost surely, $\dim(B_1[0,t_1] \cap \cdots \cap B_p[0,t_p]) = d - p(d-2)$, for any $t_1, \ldots, t_p > 0$, see Exercise 9.1 (b). $\diamond$

Note that, by Lemma 9.4, we have $\dim(\mathsf{Range}_1 \cap \mathsf{Range}_2) \leq 1$ if $d = 3$, and hence only the lower bounds in Theorem 9.8 remain to be proved. For these we use the stochastic codimension method, but first we provide a useful zero-one law.

LEMMA 9.11. *For any $\gamma > 0$ the probability of the event*

$$\Big\{ \dim\big(\mathsf{Range}_1 \cap \ldots \cap \mathsf{Range}_p\big) \geq \gamma \Big\}$$

is either zero or one, and independent of the starting points of the Brownian motions.

Proof. For $t \in (0, \infty]$ denote by

$$S(t) = \big\{x \in \mathbb{R}^d \colon \text{ there exist } 0 < t_i < t \text{ such that } x = B_1(t_1) = \cdots = B_p(t_p)\big\},$$

and let

$$p(t) = \mathbb{P}\{\dim S(t) \geq \gamma\}.$$

We start by considering the case that all Brownian motions start at the origin. Then, by monotonicity of the events,

$$\mathbb{P}\big\{ \dim S(t) \geq \gamma \text{ for all } t > 0\big\} = \lim_{t \downarrow 0} p(t).$$

The event on the left hand side is in the germ-σ-algebra and hence, by Blumenthal's zero-one law, has probability zero or one. By scaling, however, $p(t)$ does not depend on t at all, so we have either $p(t) = 0$ for all $t > 0$ or $p(t) = 1$ for all $t > 0$.

In the first case we note that, by the Markov property applied at times $t_1, \ldots, t_p$,

$$
\begin{aligned}
0 &= \mathbb{P}\big\{ \dim S(\infty) \geq \gamma \big\} \\
&= \int \mathbb{P}\big\{ \dim S(\infty) \geq \gamma \mid B_1(t) = x_1, \ldots, B_p(t) = x_p \big\}\, d\mu(x_1, \ldots, x_p),
\end{aligned}
$$

where μ is the product of p independent centred, normally distributed random variables with variances t. Therefore $\mathbb{P}\{\dim S(\infty) \geq \gamma\} = 0$ for $\mathcal{L}_{pd}$-almost every vector of starting points. Finally, for an arbitrary configuration of starting points,

$$
\begin{aligned}
&\mathbb{P}\left\{ \dim S(\infty) \geq \gamma \right\} \\
&\quad = \lim_{t \downarrow 0} \mathbb{P}\big\{ \dim\{x \in \mathbb{R}^d \colon \exists t_i \geq t \text{ such that } x = B_1(t_1) = \cdots = B_p(t_p)\} \geq \gamma \big\} = 0.
\end{aligned}
$$

A completely analogous argument can be carried out for the second case. ■

Proof of Theorem 9.8. First we look at $d = 3$ and $p = 2$. Suppose $\gamma < 1$ is arbitrary, and pick $\beta > 1$ such that $\gamma + \beta < 2$. Let $\Gamma[\gamma]$ and $\Gamma[\beta]$ by two independent percolation limit sets, indepedendent of the Brownian motions. Note that then $\Gamma[\gamma] \cap \Gamma[\beta]$ is a percolation limit set with parameter $\gamma + \beta$. Hence, by Theorem 9.5 (ii) and the fact that $\dim(\mathsf{Range}_1) = 2 > \gamma + \beta$, we have

$$
\mathbb{P}\big\{ \mathsf{Range}_1 \cap \Gamma[\gamma] \cap \Gamma[\beta] \neq \emptyset \big\} > 0.
$$

Interpreting $\Gamma[\beta]$ as the test set and using Theorem 9.5 (i) we obtain

$$
\dim \big(\mathsf{Range}_1 \cap \Gamma[\gamma] \big) \geq \beta \qquad \text{with positive probability.}
$$

As $\beta > 1$, given this event, the set $\mathsf{Range}_1 \cap \Gamma[\gamma]$ has positive capacity with respect to the potential kernel in $\mathbb{R}^3$ and is therefore nonpolar with respect to the independent Brownian motion $\{B_2(t) \colon t \geq 0\}$. We therefore have

$$
\mathbb{P}\big\{ \mathsf{Range}_1 \cap \mathsf{Range}_2 \cap \Gamma[\gamma] \neq \emptyset \big\} > 0.
$$

Using Theorem 9.5 (i) we infer that $\dim(\mathsf{Range}_1 \cap \mathsf{Range}_2) \geq \gamma$ with positive probability. Lemma 9.11 shows that this must in fact hold almost surely, and the result follows as $\gamma < 1$ was arbitrary.

Second we look at $d = 3$ and any $p \geq 2$. Suppose $\gamma < 2$ is arbitrary, and pick $\beta_1, \ldots, \beta_p > 0$ such that $\gamma + \beta_1 + \cdots + \beta_p < 2$. Let $\Gamma[\gamma]$ and $\Gamma[\beta_1], \ldots, \Gamma[\beta_p]$ be independent percolation limit sets, indepedendent of the p Brownian motions. Note that then

$$
\Gamma[\gamma] \cap \bigcap_{i=1}^{p} \Gamma[\beta_i]
$$

is a percolation limit set with parameter $\gamma + \beta_1 + \cdots + \beta_p$. Hence, by Theorem 9.5 (ii) and the fact that $\dim(\mathsf{Range}_1) = 2 > \gamma + \beta_1 + \cdots + \beta_p$, we have

$$
\mathbb{P}\Big\{ \mathsf{Range}_1 \cap \Gamma[\gamma] \cap \bigcap_{i=1}^{p} \Gamma[\beta_i] \neq \emptyset \Big\} > 0.
$$

Interpreting $\Gamma[\beta_p]$ as the test set and using Theorem 9.5 (i) we obtain

$$\dim\Big(\mathsf{Range}_1 \cap \Gamma[\gamma] \cap \bigcap_{i=1}^{p-1} \Gamma[\beta_i]\Big) \geq \beta_p \qquad \text{with positive probability.}$$

As $\beta_p > 0$, given this event, the set

$$\mathsf{Range}_1 \cap \Gamma[\gamma] \cap \bigcap_{i=1}^{p-1} \Gamma[\beta_i]$$

has positive capacity with respect to the potential kernel in $\mathbb{R}^2$ and is therefore nonpolar with respect to the independent Brownian motion $\{B_2(t) \colon t \geq 0\}$. We therefore have

$$\mathbb{P}\Big\{\mathsf{Range}_1 \cap \mathsf{Range}_2 \cap \Gamma[\gamma] \cap \bigcap_{i=1}^{p-1} \Gamma[\beta_i] \neq \emptyset\Big\} > 0.$$

Iterating this procedure $p-1$ times we obtain

$$\mathbb{P}\Big\{\bigcap_{i=1}^{p} \mathsf{Range}_i \cap \Gamma[\gamma] \neq \emptyset\Big\} > 0.$$

Using Theorem 9.5 (i) we infer that $\dim(\bigcap_{i=1}^{p} \mathsf{Range}_i) \geq \gamma$ with positive probability. Lemma 9.11 shows that this must in fact hold almost surely, and the result follows as $\gamma < 2$ was arbitrary. ∎

2. Intersection equivalence of Brownian motion and percolation limit sets

The idea of quantitative estimates of hitting probabilities, has a natural extension: two random sets may be called *intersection-equivalent*, if their hitting probabilities for a large class of test sets are comparable. This concept of equivalence allows surprising relationships between random sets which, at first sight, might not have much in common. In this section we establish intersection-equivalence between Brownian motion and suitably defined percolation limit sets, and use this to characterise the polar sets for the intersection of Brownian paths.

We start the discussion by formalising the idea of intersection-equivalence.

Definition 9.12. *Two random closed sets A and B in $\mathbb{R}^d$ are* **intersection-equivalent** *in the compact set U, if there exist two positive constants c, C such that, for any closed set $\Lambda \subset U$,*

$$c\,\mathbb{P}\{A \cap \Lambda \neq \emptyset\} \leq \mathbb{P}\{B \cap \Lambda \neq \emptyset\} \leq C\,\mathbb{P}\{A \cap \Lambda \neq \emptyset\}. \tag{2.1}$$

Using the symbol $a \asymp b$ to indicate that the ratio of a and b is bounded from above and below by positive constants which do not depend on Λ we can write this as

$$\mathbb{P}\{A \cap \Lambda \neq \emptyset\} \asymp \mathbb{P}\{B \cap \Lambda \neq \emptyset\}. \qquad \diamond$$

Remark 9.13. Let $\mathcal{G}$ be the collection of all closed subsets of $\mathbb{R}^d$. Formally, we define a random closed set as a mapping $A \colon \Omega \to \mathcal{G}$ such that, for every compact $\Lambda \subset \mathbb{R}^d$, the set $\{\omega \colon A(\omega) \cap \Lambda = \emptyset\}$ is measurable. ◇

The philosophy of the main result of this section is that we would like to find a class of particularly simple sets, which are intersection-equivalent to the paths of transient Brownian motion. If these sets are easier to study, we can 'translate' easy results about the simple sets into hard ones for Brownian motion.

A good candidate for these simple sets are percolation limit sets: they have excellent features of *self-similarity* and *independence* between the fine structures in different parts. Many of their properties can be obtained from classical facts about Galton-Watson branching processes.

We introduce percolation limit sets with *generation dependent* retention probabilities. Denote by $\mathfrak{C}_n$ the compact dyadic cubes of sidelength 2^{-n}. For any sequence $p_1, p_2, \ldots$ in $(0,1)$ we define families $\mathfrak{S}_n$ of compact dyadic cubes inductively by including any cube in $\mathfrak{C}_n$ which is not contained in a previously rejected cube, independently with probability p_n. Define

$$\Gamma = \bigcap_{n=1}^{\infty} \bigcup_{S \in \mathfrak{S}_n} S,$$

to be the **percolation limit set** for the sequence $p_1, p_2, \ldots$.

To find a suitable sequence of retention probabilities we compare the hitting probabilities of dyadic cubes by a percolation limit set on the one hand and a transient Brownian on the other. (This is obviously necessary to establish intersection-equivalence). We assume that percolation is performed in a cube Cube at positive distance from the origin, at which a transient Brownian motion is started. Supposing for the moment that the retention probabilities are such that the survival probability of any retained cube is bounded from below, for any cube $Q \in \mathfrak{C}_n$, the hitting estimates for the percolation limit set are

$$\mathbb{P}\big\{\Gamma \cap Q \neq \emptyset\big\} \asymp p_1 \cdots p_n \,.$$

By Theorem 8.23, on the other hand,

$$\mathbb{P}\big\{B[0,T] \cap Q \neq \emptyset\big\} \asymp \mathrm{Cap}_M(Q) \asymp 1/f(2^{-n})\,,$$

for the radial potential

$$f(\varepsilon) = \begin{cases} \log_2(1/\varepsilon) & \text{for } d = 2, \\ \varepsilon^{2-d} & \text{for } d \geq 3\,, \end{cases}$$

where we have chosen basis 2 for the logarithm for convenience of this argument. When we choose the sequence $p_1, p_2, \ldots$ of retention probabilities such that $p_1 \cdots p_n = 1/f(2^{-n})$. More explicitly, we choose $p_1 = 2^{2-d}$ and, for $n \geq 2$,

$$p_n = \frac{f(2^{-n+1})}{f(2^{-n})} = \begin{cases} \frac{n-1}{n} & \text{for } d = 2, \\ 2^{2-d} & \text{for } d \geq 3\,. \end{cases} \tag{2.2}$$

The retention probabilities are constant for $d \geq 3$, but generation dependent for $d = 2$.

THEOREM 9.14. *Let $\{B(t)\colon 0 \leq t \leq T\}$ denote transient Brownian motion started in the origin, and $\mathsf{Cube} \subset \mathbb{R}^d$ a compact cube of unit sidelength not containing the origin. Let Γ be a percolation limit set in Cube with retention probabilities chosen as in (2.2). Then the range of the Brownian motion is intersection-equivalent to the percolation limit set Γ in the cube Cube.*

Before discussing the proof, we look at an application of Theorem 9.14 to our understanding of Brownian motion. We first make two easy observations.

LEMMA 9.15. *Suppose that* $A_1, \ldots, A_k, F_1, \ldots, F_k$ *are independent random closed sets, with* A_i *intersection-equivalent to* F_i *for* $1 \leq i \leq k$. *Then* $A_1 \cap A_2 \cap \ldots \cap A_k$ *is intersection-equivalent to* $F_1 \cap F_2 \cap \ldots \cap F_k$.

Proof. By induction, we can reduce this to the case $k = 2$. It then clearly suffices to show that $A_1 \cap A_2$ is intersection-equivalent to $F_1 \cap A_2$. This is done by conditioning on A_2,

$$\begin{aligned} \mathbb{P}\big\{A_1 \cap A_2 \cap \Lambda \neq \emptyset\big\} &= \mathbb{E}\big[\mathbb{P}\{A_1 \cap A_2 \cap \Lambda \neq \emptyset \,|\, A_2\}\big] \\ &\asymp \mathbb{E}\big[\mathbb{P}\{F_1 \cap A_2 \cap \Lambda \neq \emptyset \,|\, A_2\}\big] \\ &= \mathbb{P}\{F_1 \cap A_2 \cap \Lambda \neq \emptyset\}. \end{aligned}$$

■

LEMMA 9.16. *For independent percolation limit sets* Γ_1 *and* Γ_2 *with retention probabilities* $p_1, p_2, \ldots$ *and* $q_1, q_2, \ldots$, *respectively, their intersection* $\Gamma_1 \cap \Gamma_2$ *is a percolation limit set with retention probabilities* $p_1q_1, p_2q_2, \ldots$.

Proof. This is obvious from the definition of percolation limit sets and independence. ■

These results enable us to recover the results about existence of intersection of Brownian paths from the survival criteria of Galton-Watson trees. As an example look at the intersection of two Brownian paths in $\mathbb{R}^d$, $d \geq 3$. By Theorem 9.14 and Lemma 9.15, the intersection of these paths is intersection-equivalent (in any unit cube not containing the starting points) to the intersection of two independent percolation limit sets with constant retention parameters $p = 2^{2-d}$. This intersection, by Lemma 9.16, is another percolation limit set, but now with parameter $p^2 = 2^{4-2d}$. Now observe that this set has a positive probability of being nonempty if and only if a Galton-Watson process with binomial offspring distribution with parameters $n = 2^d$ and $p = 2^{4-2d}$ has a positive survival probability. This is the case if and only if the mean offspring number np strictly exceeds 1, i.e. if $4 - d > 0$. In other words, in $d = 3$ the two paths intersect with positive probability, in all higher dimensions they almost surely do not intersect.

We now give the proof of Theorem 9.14. A key rôle in the proof is played by a fundamental result of R. Lyons concerning survival probabilities of general trees under the percolation process, which has great formal similarity with the quantitative hitting estimates for Brownian paths of Theorem 8.23.

Recall the notation for trees from Page 112. As usual we define, for any kernel $K\colon \partial T \times \partial T \to [0, \infty]$, the K-energy of the measure μ on ∂T as

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

and the K-capacity of the boundary of the tree by

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

Given a sequence $p_1, p_2, \ldots$ of probabilities, *percolation* on T is obtained by removing each edge of T of order n independently with probability $1 - p_n$ and retaining it otherwise, with mutual independence among edges. Say that a ray ξ **survives the percolation** if all the edges on ξ are retained, and say that the tree boundary ∂T survives if some ray of T survives.

THEOREM 9.17 (Lyons). *If percolation with retention probabilities $p_1, p_2, \ldots$ is performed on a rooted tree T, then*

$$\mathrm{Cap}_K(\partial T) \le \mathbb{P}\big\{\, \partial T \textit{ survives the percolation} \,\big\} \le 2\mathrm{Cap}_K(\partial T)\,, \tag{2.3}$$

where the kernel K is defined by $K(x, y) = \prod_{i=1}^{|x\wedge y|} p_i^{-1}$.

Proof. For two vertices v, w we write $v \leftrightarrow w$ if the shortest path between the vertices is retained in the percolation. We also write $v \leftrightarrow \partial T$ if a ray through vertex v survives the percolation and $v \leftrightarrow T_n$ if there is a self-avoiding path of retained edges connecting v to a vertex of generation n. Note that $K(x, y) = \mathbb{P}\{\rho \leftrightarrow x \wedge y\}^{-1}$ by definition of the kernel K. By the finiteness of the degrees,

$$\{\rho \leftrightarrow \partial T\} = \bigcap_n \{\rho \leftrightarrow T_n\}\,.$$

We start with the left inequality in (2.3) and consider the case of a finite tree T first. We extend the definition of the boundary ∂T to finite trees by letting ∂T be the set of leaves, i.e., the vertices with no offspring. Let μ be a probability measure on ∂T and set

$$Y \;=\; \sum_{x\in\partial T} \mu(x) \frac{\mathbb{1}\{\rho \leftrightarrow x\}}{\mathbb{P}\{\rho \leftrightarrow x\}}\,.$$

Then $\mathbb{E}[Y] \;=\; \sum_{x\in\partial T} \mu(x) = 1$, and

$$\begin{aligned}
\mathbb{E}[Y^2] \;&=\; \mathbb{E}\left[\sum_{x\in\partial T}\sum_{y\in\partial T} \mu(x)\mu(y)\; \frac{\mathbb{1}\{\rho \leftrightarrow x,\, \rho \leftrightarrow y\}}{\mathbb{P}\{\rho \leftrightarrow x\}\,\mathbb{P}\{\rho \leftrightarrow y\}}\right] \\
&=\; \sum_{x\in\partial T}\sum_{y\in\partial T} \mu(x)\mu(y) \frac{\mathbb{P}\{\rho \leftrightarrow x \text{ and } \rho \leftrightarrow y\}}{\mathbb{P}\{\rho \leftrightarrow x\}\mathbb{P}\{\rho \leftrightarrow y\}}\,.
\end{aligned}$$

Thus,

$$\mathbb{E}[Y^2] \;=\; \sum_{x,y\in\partial T} \mu(x)\mu(y)\; \frac{1}{\mathbb{P}\{\rho \leftrightarrow x \wedge y\}} \;=\; I_K(\mu).$$

Using the Paley-Zygmund inequality in the second step, we obtain

$$\mathbb{P}\{\rho \leftrightarrow \partial T\} \ge \mathbb{P}\{Y > 0\} \ge \frac{\big(\mathbb{E}[Y]\big)^2}{\mathbb{E}[Y^2]} = \frac{1}{I_K(\mu)}\,.$$

The left-hand side does not depend on μ, so optimising the right-hand side over μ yields

$$\mathbb{P}\{\rho \leftrightarrow \partial T\} \ge \sup_\mu \frac{1}{I_K(\mu)} = \mathrm{Cap}_K(\partial T)\,,$$

which proves the lower bound for finite trees. For T infinite, let μ be any probability measure on ∂T. This induces a probability measure $\widetilde{\mu}$ on the set T_n, consisting of those vertices which become leaves when the tree T is cut off after the n^{th} generation, by letting

$$\widetilde{\mu}(v) = \mu\big\{\xi \in \partial T \,:\, v \in \xi\big\}, \qquad \text{for any vertex } v \in T_n\,.$$

By the finite case considered above,

$$\mathbb{P}\{\rho \leftrightarrow T_n\} \geq \Big(\sum_{x,y \in T_n} K(x,y)\widetilde{\mu}(x)\widetilde{\mu}(y)\Big)^{-1}.$$

Each ray ξ must pass through some vertex $x \in T_n$. This implies that $K(x,y) \leq K(\xi,\eta)$ for $x \in \xi$ and $y \in \eta$. Therefore,

$$\int_{\partial T}\int_{\partial T} K(\xi,\eta)\, d\mu(\xi)\, d\mu(\eta) \geq \sum_{x,y\in T_n} K(x,y)\widetilde{\mu}(x)\widetilde{\mu}(y) \geq \frac{1}{\mathbb{P}\{\rho \leftrightarrow T_n\}}\,.$$

Hence $\mathbb{P}\{\rho \leftrightarrow T_n\} \geq I_K(\mu)^{-1}$ for any probability measure μ on ∂T. Optimising over μ and passing to the limit as $n \to \infty$, we get $\mathbb{P}\{\rho \leftrightarrow \partial T\} \geq \mathrm{Cap}_K(\partial T)$.

It remains to prove the right-hand inequality in (2.3). Assume first that T is finite. There is a Markov chain $\{V_k \,:\, k \in \mathbb{N}\}$ hiding here: Suppose the offspring of each individual is ordered from left to right, and note that this imposes a natural order on all vertices of the tree by saying that x is to the left of y if there are siblings v, w with v to the left of w, such that x is a descendant of v and y is a descendant of w. The random set of leaves that survive the percolation may thus be enumerated from left to right as $V_1, V_2, \ldots, V_r$. The key observation is that the random sequence $\rho, V_1, V_2, \ldots, V_r, \Delta, \Delta, \ldots$ is a Markov chain on the state space $\partial T \cup \{\rho, \Delta\}$, where ρ is the root and Δ is a formal absorbing cemetery.

Indeed, given that $V_k = x$, all the edges on the unique path from ρ to x are retained, so that survival of leaves to the right of x is determined by the edges strictly to the right of the path from ρ to x, and is thus conditionally independent of $V_1, \ldots, V_{k-1}$, see Figure 1.

This verifies the Markov property, so Proposition 8.25 may be applied. The transition probabilities for the Markov chain above are complicated, but it is easy to write down the Green kernel G. For any vertex x let $\mathtt{path}(x)$ be the set of edges on the shortest path from ρ to x. Clearly, $G(\rho, y)$ equals the probability that y survives percolation, so

$$G(\rho,y) = \prod_{n=1}^{|y|} p_n\,.$$

If x is to the left of y, then $G(x,y)$ is equal to the probability that the range of the Markov chain contains y given that it contains x, which is just the probability of y surviving given that x survives. Therefore,

$$G(x,y) = \prod_{n=|x\wedge y|+1}^{|y|} p_n\,,$$

and hence

$$M(x,y) = \frac{G(x,y)}{G(\rho,y)} = \prod_{n=1}^{|x\wedge y|} p_n^{-1}\,.$$

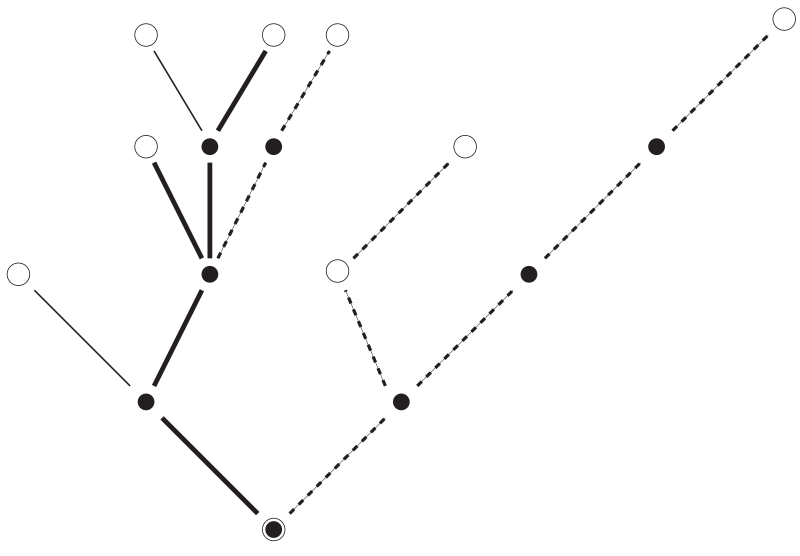

FIGURE 1. The Markov chain embedded in the tree.

Now $G(x,y) = 0$ for x on the right of y; thus (keeping the diagonal in mind)

$$K(x,y) \le M(x,y) + M(y,x)$$

for all $x, y \in \partial T$, and therefore $I_K(\mu) \le 2I_M(\mu)$. Now apply Proposition 8.25 to $\Lambda = \partial T$:

$$\mathrm{Cap}_K(\partial T) \ge \tfrac{1}{2}\mathrm{Cap}_M(\partial T) \ge \tfrac{1}{2}\,\mathbb{P}\big\{\{V_k \,:\, k \in \mathbb{N}\} \text{ hits } \partial T\big\} = \tfrac{1}{2}\,\mathbb{P}\{\rho \leftrightarrow \partial T\}\,.$$

This establishes the upper bound for finite T. The inequality for general T follows from the finite case by taking limits. ■

The main remaining task is to translate Lyons' theorem, Theorem 9.17 into hitting estimates for percolation limit sets using a 'tree representation' as in Figure 2, and relating the capacity of the tree boundary to the capacity of the percolation limit set.

THEOREM 9.18. *Let* Γ *be a percolation limit set in the unit cube* Cube *with retention parameters* $p_1, p_2, \ldots$. *Then, for any closed set* $\Lambda \subset$ Cube *we have*

$$\mathbb{P}\big\{\Gamma \cap \Lambda \ne \emptyset\big\} \asymp \mathrm{Cap}_f(\Lambda)\,,$$

for any decreasing f *satisfying* $f(2^{-k}) = p_1^{-1} \cdots p_k^{-1}$.

REMARK 9.19. This result extends parts (i) and (ii) in Hawkes' theorem, Theorem 9.5, in two ways: It includes generation dependent retention and gives a quantitative estimate. ◇

The key to this lemma is the following representation for the f-energy of a measure.

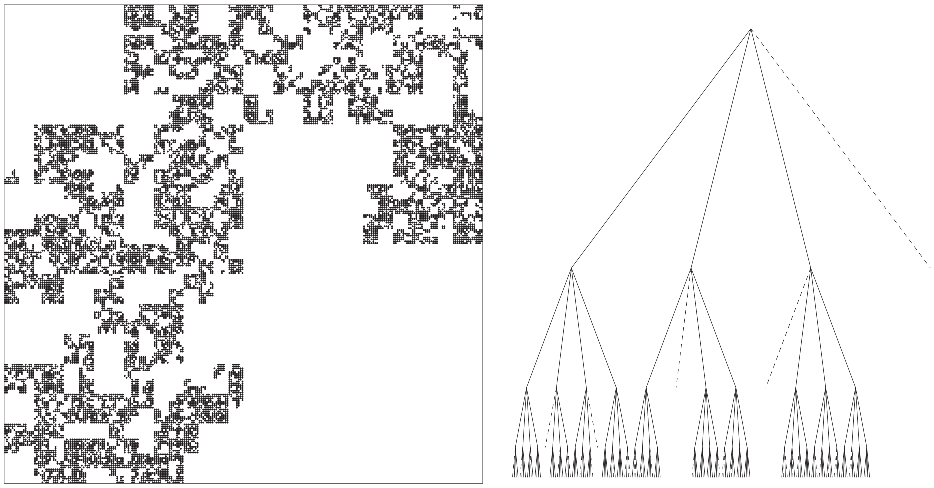

FIGURE 2. Percolation limit set and associated tree

LEMMA 9.20. *Suppose $f\colon (0,\infty) \to (0,\infty)$ is a decreasing function, and denote $h(n) = f(2^{-n}) - f(2^{1-n})$ for $n \geq 1$, and $h(0) = f(1)$. Then, for any measure μ on the unit cube $[0,1)^d$,*

$$I_f(\mu) \asymp \sum_{n=0}^{\infty} h(n) \Big(\sum_{Q \in \mathfrak{D}_n} \mu(Q)^2 \Big),$$

where the implied constants depend only on d.

Proof of the lower bound in Lemma 9.20. Fix an integer ℓ such that $\sqrt{d} \leq 2^\ell$. For any $x, y \in [0,1]^d$ we write $n(x,y) = \max\big\{n : x, y \in D \text{ for some } D \in \mathfrak{D}_n\big\}$. Note that $n(x,y) = n + \ell$ implies $|\,x - y| \leq \sqrt{d}2^{-n-\ell} \leq 2^{-n}$ and hence $f(|\,x-y|) \geq f(2^{-n})$. We thus get

$$\begin{aligned} I_f(\mu) &= \iint f(\,|x-y|)\, d\mu(x)\, d\mu(y) \\ &\geq \sum_{n=0}^{\infty} f(2^{-n})\, \mu \otimes \mu\big\{(x,y) : n(x,y) = n+\ell\big\} = \sum_{n=0}^{\infty} f(2^{-n}) \big[S_{n+\ell}(\mu) - S_{n+\ell+1}(\mu)\big], \end{aligned}$$

where $S_n(\mu) = \sum_{Q \in \mathfrak{D}_n} \mu(Q)^2$. Note that, by the Cauchy-Schwarz inequality,

$$S_n(\mu) = \sum_{Q \in \mathfrak{D}_n} \mu(Q)^2 \leq \sum_{Q \in \mathfrak{D}_n} \Big(\sum_{\substack{V \in \mathfrak{D}_{n+1} \\ V \subset Q}} \mu(V) \Big)^2 \leq 2^d \sum_{V \in \mathfrak{D}_{n+1}} \mu(V)^2 = 2^d\, S_{n+1}(\mu)\,. \tag{2.4}$$

Rearranging the sum and using this ℓ times, we obtain that

$$\sum_{n=0}^{\infty} f(2^{-n}) \big[S_{n+\ell}(\mu) - S_{n+\ell+1}(\mu)\big] = \sum_{n=0}^{\infty} h(n)\, S_{n+\ell}(\mu) \geq \sum_{n=0}^{\infty} h(n)\, 2^{-d\ell}\, S_n(\mu)\,,$$

which is our statement with $c = 2^{-d\ell}$. ∎

Proof of the upper bound in Lemma 9.20. For $2^{1-n} \geq |\,x-y| > 2^{-n}$, we have

$$\sum_{k=0}^{\infty} h(k)\mathbb{1}\{2^{1-k} \geq |\,x-y|\} = f(2^{-n}) \geq f(|\,x-y|)\,,$$

and hence we can decompose the integral as

$$I_f(\mu) = \iint f(|\,x-y|)\,d\mu(x)\,d\mu(y) \leq \iint \sum_{k=0}^{\infty} h(k)\mathbb{1}\{2^{1-k} \geq |\,x-y|\}\,d\mu(x)\,d\mu(y)\,.$$

For cubes $Q_1, Q_2 \in \mathfrak{D}_n$ we write $Q_1 \sim Q_2$ is they are either adjacent or they agree (though note that $\sim$ is not an equivalence relation). Then

$$\begin{aligned}\iint \mathbb{1}\{2^{1-k} \geq |\,x-y|\}\,d\mu(x)\,d\mu(y) &= \mu \otimes \mu\big\{(x,y) \,:\, |\,x-y| \leq 2^{1-k}\big\} \\ &\leq \sum_{\substack{Q_1,Q_2\in\mathfrak{D}_{k-1}\\ Q_1\sim Q_2}} \mu(Q_1)\,\mu(Q_2) \leq \tfrac{1}{2} \sum_{\substack{Q_1,Q_2\in\mathfrak{D}_{k-1}\\ Q_1\sim Q_2}} \Big(\mu(Q_1)^2 + \mu(Q_2)^2\Big)\,,\end{aligned}$$

using the inequality of the geometric and arithmetic mean in the last step. As, for each cube, the number of adjacent or identical dyadic cubes of the same sidelength is 3^d, we obtain that

$$I_f(\mu) \leq \tfrac{3^d+1}{2} \sum_{k=0}^{\infty} h(k) \sum_{Q\in\mathfrak{D}_{k-1}} \mu(Q)^2 \leq 3^{d+1}2^{d-1} \sum_{k=0}^{\infty} h(k) \sum_{Q\in\mathfrak{D}_k} \mu(Q)^2\,,$$

using (2.4) from above. This completes the proof of the upper bound. ∎

Proof. Denote the coordinatewise minimum of Cube by x. We employ the canonical mapping $\mathcal{R}$ from the boundary of a 2^d-ary tree Υ, where every vertex has 2^d children, to the cube Cube. Formally, label the edges from each vertex to its children in a one-to-one manner with the vectors in $\Theta = \{0,1\}^d$. Then the boundary $\partial\Upsilon$ is identified with the sequence space $\Theta^{\mathbb{Z}^+}$ and we define $\mathcal{R}\colon \partial\Upsilon = \Theta^{\mathbb{Z}^+} \to \mathsf{Cube}$ by

$$\mathcal{R}(\omega_1, \omega_2, \ldots) = x + \sum_{n=1}^{\infty} 2^{-n}\omega_n.$$

We now use the representation given in Lemma 9.20 to relate the K-energy of a measure μ on ∂T (with K as in Theorem 9.17) to the f-energy of its image measure $\mu\circ\mathcal{R}^{-1}$ on Cube, showing that

$$I_K(\mu) \asymp I_f(\mu\circ\mathcal{R}^{-1}) \tag{2.5}$$

where the implied constants depend only on the dimension d. Indeed the K-energy of a measure μ on ∂T satisfies, by definition,

$$\begin{aligned} I_K(\mu) &= \iint \prod_{i=1}^{|x\wedge y|} p_i^{-1}\, d\mu(x)\, d\mu(y) = \iint \sum_{v\le x\wedge y} \Big(\prod_{i=1}^{|v|} p_i^{-1} - \prod_{i=1}^{|v|-1} p_i^{-1}\Big)\, d\mu(x)\, d\mu(y) \\ &= \sum_{v\in\Gamma} \Big(\prod_{i=1}^{|v|} p_i^{-1} - \prod_{i=1}^{|v|-1} p_i^{-1}\Big) \iint \mathbb{1}\{x,y\ge v\}\, d\mu(x)\, d\mu(y) \\ &= \sum_{v\in\Gamma} \Big(\prod_{i=1}^{|v|} p_i^{-1} - \prod_{i=1}^{|v|-1} p_i^{-1}\Big)\, \mu\big(\{\xi\in\partial T \,:\, v\in\xi\}\big)^2, \end{aligned}$$

whereas the f-energy of the measure $\mu\circ\mathcal{R}^{-1}$ satisfies, by Lemma 9.20,

$$I_f\big(\mu\circ\mathcal{R}^{-1}\big) \asymp \sum_{k=0}^{\infty} h(k) \sum_{D\in\mathfrak{D}_k} \mu\big(\mathcal{R}^{-1}(D)\big)^2,$$

where

$$h(k) = f(2^{-k}) - f(2^{-k+1}) = p_1^{-1}\cdots p_k^{-1} - p_1^{-1}\cdots p_{k+1}^{-1}$$

by our assumptions on f. Now $\mathcal{R}^{-1}(D)$ is contained in no more than 3^d sets of the form $\big\{\xi\in\partial T \,:\, v\in\xi\big\}$, for $|v|=k$, in such a way that over all cubes $D\in\mathfrak{D}_k$ no such set is used in more than 3^d covers. Conversely each set $\mathcal{R}^{-1}(D)$ contains an individual set of this form entirely, so that we obtain (2.5).

As any measure ν on $\mathcal{R}(\partial T)\subset \mathsf{Cube}$ can be written as $\mu\circ\mathcal{R}^{-1}$ for an appropriate measure μ on ∂T it follows from (2.5) that $\mathrm{Cap}_K(\partial T)\asymp \mathrm{Cap}_f(\mathcal{R}(\partial T))$. Any closed set Λ in the unit cube Cube can be written as the image $\mathcal{R}(\partial T)$ of the boundary of some subtree T of the regular 2^d-ary tree. We perform percolation with retention parameters $p_1,p_2,\ldots$ on this tree. Then, by Theorem 9.17,

$$\begin{aligned} \mathbb{P}\big\{\Gamma\cap\Lambda\neq\emptyset\big\} &= \mathbb{P}\big\{\ \partial T \text{ survives the percolation }\big\} \\ &\asymp \mathrm{Cap}_K(\partial T) \ \asymp\ \mathrm{Cap}_f(\Lambda)\,. \end{aligned}$$

■

Proof of Theorem 9.14. As the cube Cube has positive distance to the starting point of Brownian motion, we can remove the denominator and smaller order terms from the Martin kernel in Theorem 8.23, as in the proof of Theorem 8.19. We thus obtain

$$\mathbb{P}\big\{B[0,T]\cap\Lambda\neq\emptyset\big\} \asymp \mathrm{Cap}_f(\Lambda)\,,$$

where f is the radial potential. For the choice of retention probabilities in (2.2) we can apply Theorem 9.18, which implies

$$\mathrm{Cap}_f(\Lambda) \asymp \mathbb{P}\big\{\Gamma\cap\Lambda\neq\emptyset\big\}\,,$$

and combining the two displays gives the result. ■

The intersection-equivalence approach enables us to characterise the polar sets for the intersection of p independent Brownian motions in $\mathbb{R}^d$ and give a quantitative estimate of the hitting probabilities.

THEOREM 9.21. *Let $B_1, \ldots, B_p$ be independent Brownian motions in $\mathbb{R}^d$ starting in arbitrary fixed points and suppose $p(d-2) < d$. Let*

$$S = \big\{x \in \mathbb{R}^d \colon\ \exists\, t_1, \ldots, t_p > 0 \text{ with } x = B_1(t_1) = \cdots = B_p(t_p)\big\}.$$

Then, for any closed set Λ, we have

$$\mathbb{P}\big\{S \cap \Lambda \neq \emptyset\big\} > 0 \qquad \text{if and only if} \qquad \mathrm{Cap}_{f^p}(\Lambda) > 0\,,$$

where f is the radial potential.

Proof. We may assume that Λ is contained in a unit cube at positive distance from the starting points. Let Γ be a percolation limit set in that cube, with retention probabilities $p_1, p_2, \ldots$ satisfying $p_1 \cdots p_n = 1/f^p(2^{-n})$. By Theorem 9.14 and Lemma 9.15, the random set S is intersection-equivalent to Γ in that cube. Theorem 9.18 characterises the polar sets for Γ, completing the argument. ∎

3. Multiple points of Brownian paths

The results of the previous section also provide the complete answer to the question of the existence of p-fold multiple points of d-dimensional Brownian motion. This is achived by

THEOREM 9.22. *Suppose $d \geq 2$ and $\{B(t) : t \in [0,1]\}$ is a d-dimensional Brownian motion. Then, almost surely,*

- *if $d \geq 4$ 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. To show *nonexistence* of double points in $d \geq 4$ we divide, for every $n \in \mathbb{N}$ the interval $[0,1)$ into 2^n equal subintervals of the form $[k2^{-n}, (k+1)2^{-n})$. Note that, for all $s, t \in [0,1)$ with $s < t$ there exists a unique $n \in \mathbb{N}$ and $k \in \{1, \ldots, 2^n - 1\}$ with $s \in [(k-1)2^{-n}, k2^{-n})$ and $t \in [k2^{-n}, (k+1)2^{-n})$, see Figure 3.

The Brownian motions $\{B_1(t) : 0 \leq t \leq 2^{-n}\}$ and $\{B_2(t) : 0 \leq t \leq 2^{-n}\}$ given by

$$B_1(t) = B(k2^{-n} + t) - B(k2^{-n}) \quad \text{ and } \quad B_2(t) = B(k2^{-n} - t) - B(k2^{-n})$$

are independent and hence, by Theorem 9.1, they do not intersect, almost surely. Hence for each pair $[(k-1)2^{-n}, k2^{-n})$, $[k2^{-n}, (k+1)2^{-n})$ of intervals, almost surely, there is no $s \in [(k-1)2^{-n}, k2^{-n})$ and $t \in [k2^{-n}, (k+1)2^{-n})$ with $B(s) = B(t)$. As this holds almost surely simultaneously for all pairs of intervals and all n, we note that there exists no $s, t \in [0,1)$ with $s \neq t$ but $B(s) = B(t)$.

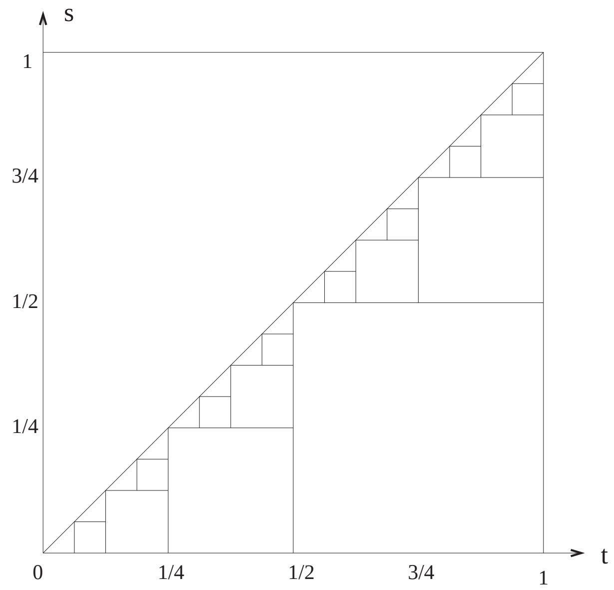

FIGURE 3. Exhausting the triangle $\{(s,t) \in [0,1)^2,\ s < t\}$ by squares of the form $[(k-1)2^{-n}, k2^{-n}) \times [k2^n, (k+1)2^n)$.

To show *existence* of double points in $d \le 3$ we fix an arbitrary pair of adjacent intervals, say

$$[(k-1)2^{-n}, k2^{-n}) \quad \text{and} \quad [k2^{-n}, (k+1)2^{-n}).$$

We apply Theorem 9.1 in conjunction with Remark 9.2, to the independent Brownian motions $\{B_1(t) : 0 \le t \le 2^{-n}\}$ and $\{B_2(t) : 0 \le t \le 2^{-n}\}$ given by

$$B_1(t) = B(k2^{-n} + t) - B(k2^{-n}) \quad \text{and} \quad B_2(t) = B(k2^{-n} - t) - B(k2^{-n}),$$

to see that, almost surely, the two ranges intersect.

To show nonexistence of triple points in $d = 3$ we observe that it suffices to show that for any four rationals $0 < \alpha_1 < \alpha_2 < \alpha_3 < \alpha_4 < 1$, almost surely there is no $0 < s < \alpha_1$, $\alpha_2 < t < \alpha_3$, $\alpha_4 < u < 1$ with $B(s) = B(t) = B(u)$.

Fix four rationals as above and denote by μ the law of the vector

$$\big(B(\alpha_1), B(\alpha_2), B(\alpha_3), B(\alpha_4)\big).$$

Obviously this law has a bounded density with respect to a vector (X_1, X_2, X_3, X_4) of independent standard normal random variables with variances $\alpha_1, \alpha_2, \alpha_3, \alpha_4$. Denoting the upper bound by $C > 0$, we obtain

$$\begin{aligned}
&\mathbb{P}\big\{ \text{ there exist } 0 < s < \alpha_1, \alpha_2 < t < \alpha_3, \alpha_4 < u < 1 \text{ with } B(s) = B(t) = B(u)\big\} \\
&\quad = \int \mathbb{P}\big\{B[0,\alpha_1] \cap B[\alpha_2,\alpha_3] \cap B[\alpha_4, 1] \neq \emptyset \mid B(\alpha_i) = x_i \text{ for } i = 1, \dots, 4\big\}\, d\mu(x_1, \cdots, x_4) \\
&\quad \le C\, \mathbb{E}\Big[\mathbb{P}\big\{B[0,\alpha_1] \cap B[\alpha_2,\alpha_3] \cap B[\alpha_4, 1] \neq \emptyset \mid B(\alpha_i) = X_i \text{ for } i = 1, \dots, 4\big\}\Big],
\end{aligned}$$

where the last expectation is with respect to the vector (X_1, X_2, X_3, X_4). It is easy to see, for example from Lévy's construction of Brownian motion on an interval, that this expectation equals

$$\mathbb{P}\big\{B_1[0,\alpha_1]\cap B_2[\alpha_2,\alpha_3]\cap B_3[\alpha_4,1]\neq\emptyset\big\}$$

where B_1, B_2, B_3 are three independent Brownian motions. Hence Theorem 9.3 shows that the probability is zero, and therefore there are no triple points of Brownian motion in $d = 3$.

To show the existence of p-multiple points in $\mathbb{R}^2$ fix numbers

$$0<\alpha_1<\alpha_2<\cdots<\alpha_{2p}<\varepsilon,$$

and let μ the law of the vector

$$\big(B(\alpha_1),\dots,B(\alpha_{2p})\big).$$

This law has a density with respect to a vector $(X_1,\dots,X_{2p})$ of independent standard normal random variables with variances $\alpha_1,\dots\alpha_{2p}$, which is bounded from below, say by $c>0$. Hence, with $\alpha_0 := 0$,

$$\begin{aligned}
&\mathbb{P}\Big\{\bigcap_{i=0}^{p-1} B[\alpha_{2i},\alpha_{2i+1}]\neq\emptyset\Big\}\\
&\quad=\int\mathbb{P}\Big\{\bigcap_{i=0}^{p-1} B[\alpha_{2i},\alpha_{2i+1}]\neq\emptyset\Big|\,B(\alpha_i)=x_i\text{ for }i=1,\dots,2p\Big\}\,d\mu(x_1,\cdots,x_{2p})\\
&\quad\geq c\,\mathbb{E}\Big[\mathbb{P}\Big\{\bigcap_{i=0}^{p-1} B[\alpha_{2i},\alpha_{2i+1}]=\emptyset\Big|\,B(\alpha_i)=X_i\text{ for }i=1,\dots,2p\Big\}\Big],
\end{aligned}$$

where the last expectation is with respect to the vector $(X_1,\dots,X_{2p})$. The last exepctation equals

$$\mathbb{P}\Big\{\bigcap_{i=0}^{p-1} B_i[\alpha_{2i},\alpha_{2i+1}]\neq\emptyset\Big\}>0.$$

Hence, for every $\varepsilon>0$, p-fold self-intersections of the Brownian path occur before time ε with positive probability. By Brownian scaling this probability does not depend on $\varepsilon>0$, and therefore also the event that p-fold self-intersection occur before any positive time has positive probability. As this event is in the germ-σ-algebra, Blumenthal's zero-one law implies that its probaility must be one, which completes the proof. ■

Theorem 9.23. *Let $\{B(t) : 0\leq t\leq 1\}$ be a planar Brownian motion. Then, almost surely, for every positive integer p, there exists points $x\in\mathbb{R}^2$ which are visited* exactly *p times by the Brownian motion.*

Proof. Note first that it suffices to show this with positive probability. Indeed, by Brownian scaling, the probability that the path $\{B(t) : 0\leq t\leq r\}$ has points of multiplicity exactly p does not depend on r. By Blumenthal's zero-one law it therefore must be zero or one.

The idea of the proof is now to construct a set Λ such that $\mathrm{Cap}_{f^p}(\Lambda) > 0$ but $\mathrm{Cap}_{f^{p+1}}(\Lambda) = 0$ for the radial potential f. By Exercise 9.2 the first condition implies that the probability that Λ contains a p-fold multiple point is positive. The second conditon ensures that it almost surely does not contain a $p+1$-fold multiple point. Hence the p-multiple points found in Λ must be strictly p-multiple.

We construct the set Λ by iteration, starting from a compact unit cube Cube. In the n^{th} construction step we divide each cube retained in the previous step into its four nonoverlapping dyadic subcubes and retain only one of them, say the bottom left cube, except at the steps with number

$$n = \lceil 4^{\frac{k}{p+1}} \rceil, \qquad \text{for } k = p+1, p+2, \ldots,$$

when we retain all four subcubes. The number $k(n)$ of times within the first n steps when we have retained all four cubes satisfies $k(n) \asymp (\log n)\, \frac{p+1}{\log 4}$. Denoting by $\mathfrak{S}_n$ the set of all dyadic cubes retained in the n^{th} step, we define the compact set

$$\Lambda = \bigcap_{n=1}^{\infty} \bigcup_{S \in \mathfrak{S}_n} S\,.$$

The calculation of the capacity of Λ will be based on the formula given in Lemma 9.20. Observe that, if $f^p(\varepsilon) = \log^p(1/\varepsilon)$ is the p^{th} power of the 2-dimensional radial potential, then the associated function is

$$h^{(p)}(n) = f^p(2^{-n}) - f^p(2^{1-n}) \asymp n^p - (n-1)^p \asymp n^{p-1}\,.$$

Note that the number $g(n)$ of cubes kept in the first n steps of the construction satisfies $g(n) \asymp n^{p+1}$. By our construction $\sum_{n=0}^{\infty} n^{p-1}\, g(n)^{-1} < \infty$, but $\sum_{n=0}^{\infty} n^p\, g(n)^{-1} = \infty$. For the measure μ distributing the unit mass equally among the retained cubes of the same sidelength (hence giving mass $g(n)^{-1}$ to each retained cube), we have

$$I_{f^p}(\mu) \asymp \sum_{n=0}^{\infty} h^{(p)}(n) \Big(\sum_{Q \in \mathfrak{D}_n} \mu(Q)^2 \Big) \asymp \sum_{n=0}^{\infty} n^{p-1}\; g(n)^{-1} < \infty\,,$$

and hence $\mathrm{Cap}_{f^p}(\Lambda) > 0$. For the converse statement, note that the equidistributing measure μ minimises the sum $\sum_{Q \in \mathfrak{D}_n} \nu(Q)^2$ over all probability measures ν charging only the retained cubes. Hence, for any probability measure on Λ,

$$I_{f^{p+1}}(\nu) \asymp \sum_{n=0}^{\infty} h^{(p)}(n) \Big(\sum_{Q \in \mathfrak{D}_n} \nu(Q)^2 \Big) \geq \sum_{n=0}^{\infty} h^{(p)}(n)\; g(n)^{-1} \asymp \sum_{n=0}^{\infty} n^p\; g(n)^{-1} = \infty\,,$$

verifying that $\mathrm{Cap}_{f^{p+1}}(\Lambda) = 0$. This completes the proof. ■

Knowing that planar Brownian motion has points of arbitrarily large *finite* multiplicity, it is an interesting question whether there are points of *infinite* multiplicity.

THEOREM* 9.24. *Let $\{B(t) : t \geq 0\}$ be a planar Brownian motion. Then, almost surely, there exists a point $x \in \mathbb{R}^2$ such that the set $\{t \geq 0 : B(t) = x\}$ is uncountable.*

The rest of this section is devoted to the proof of this nontrivial result and will not be used in the remainder of the book. It may be skipped on first reading.

Let us first describe the rough strategy of the proof: We start with a double point, i.e. some $s_1 < s_2$ such that $B(s_1) = B(s_2)$ and suppose s_1 and s_2 are not too close. Forgetting that such times are necessarily random times, we could argue that for a small $\varepsilon_1 > 0$ the four independent Brownian motions

$$\{B(s_1 - t) - B(s_1) \,:\, t \in (0, \varepsilon_1)\}, \{B(s_1 + t) - B(s_1) \,:\, t \in (0, \varepsilon_1)\},$$
$$\{B(s_2 - t) - B(s_1) \,:\, t \in (0, \varepsilon_1)\}, \{B(s_2 + t) - B(s_1) \,:\, t \in (0, \varepsilon_1)\},$$

which all start in the origin, have a point of intersection with probability one. Hence there existed a quadruple point, i.e. times $t_1 < t_2 < t_3 < t_4$ with $B(t_1) = B(t_2) = B(t_3) = B(t_4)$.

Again under the false assumption that these times are fixed, we could iterate this argument with a very small ε_2. Inductively we would construct a sequence x_n of points of multiplicity 2^n converging to some x. Consider the set of times where Brownian motion visits this point. In the closure of each of the 2^n disjoint time intervals of length ε_n, on which the 2^n Brownian motions of the n^{th} stage are defined, there is a time t with $B(t) = x$. Hence there must be at least as many such times as rays in a binary tree, i.e. uncountably many.

While this idea is nice, it cannot be applied directly: we cannot choose the intersection times as stopping times, or even fixed times, for our Brownian motions. In the proof we therefore replace the intersection times by the hitting times of small balls, which are stopping times. However when moving from stage n to stage $n+1$ we only have a 2^{n+1}-fold intersection with *positive* probability, not with probability one. We therefore need to do this simultaneously for many balls and obtain a high probability of intersection by an additional *law of large numbers* effect.

Throughout the proof we use the following notation. For any open or closed sets A_1, A_2, ... and a Brownian motion $B\colon [0,\infty) \to \mathbb{R}^2$ define stopping times

$$\tau(A_1) := \inf\{t \geq 0\colon\; B(t) \in A_1\},$$
$$\tau(A_1, \dots, A_n) := \inf\{t \geq \tau(A_1, \dots, A_{n-1}) \,:\, B(t) \in A_n\}, \qquad \text{for } n \geq 2,$$

where, as usual, the infimum over the empty set is set to infinity. We say the Brownian motion *upcrosses the shell* $\mathcal{B}(x,2r) \setminus \mathcal{B}(x,r)$ *twice* before a stopping time T if,

$$\tau\big(\mathcal{B}(x,r), \mathcal{B}(x,2r)^{\mathrm{c}}, \mathcal{B}(x,r), \mathcal{B}(x,2r)^{\mathrm{c}}\big) < T.$$

We call the paths of Brownian motion between times $\tau(\mathcal{B}(x,r))$ and $\tau(\mathcal{B}(x,r), \mathcal{B}(x,2r)^c)$ and between times $\tau(\mathcal{B}(x,r), \mathcal{B}(x,2r)^c, \mathcal{B}(x,r))$ and $\tau(\mathcal{B}(x,r), \mathcal{B}(x,2r)^c, \mathcal{B}(x,r), \mathcal{B}(x,2r)^c)$ the *upcrossing excursions*, see Figure 4.

From now on let T be the first exit time of Brownian motion from the unit ball. Recall the following fact from Theorem 8.23 and the discussion of the Martin kernels in Chapter 8.

LEMMA 9.25. *Let* $1 < m < n$ *be two integers and* B *be a ball of radius* 2^{-n} *with centre at distance at least* 2^{-m} *and at most* 2^{1-m} *from the origin. Then, for sufficiently large* m*, we have*

$$\frac{m}{2n} \leq \mathbb{P}\big\{\tau(B) < T\big\} \leq \frac{2m}{n}\,.$$

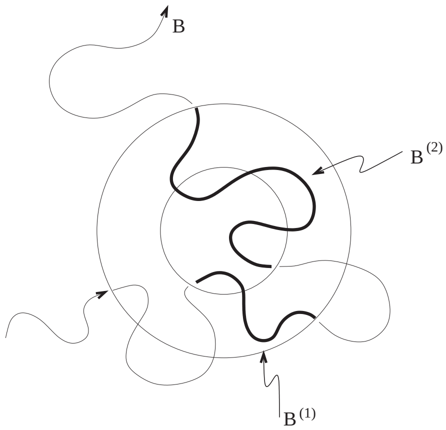

FIGURE 4. The path $B\colon [0,\infty) \to \mathbb{R}^2$ upcrosses the shell twice; the upcrossing excursions are bold and marked $B^{(1)}$, $B^{(2)}$.

Recall from Theorem 3.43 that the density of $B(T)$ under $\mathbb{P}_z$ is given by the Poisson kernel, which is

$$\mathcal{P}(z,w) = \frac{1-|z|^2}{|z-w|^2} \qquad \text{for any } z \in \mathcal{B}(0,1) \text{ and } w \in \partial\mathcal{B}(0,1).$$

LEMMA 9.26. *Consider Brownian motion started at $z \in \mathcal{B}(0,r)$ where $r<1$, and stopped at time T when it exits the unit ball. Let $\tau \le T$ be a stopping time, and let $A \in \mathcal{F}(\tau)$. Then we have*

(i) $\mathbb{P}_z\big(A \,\big|\, B(T)\big) = \mathbb{P}_z(A)\, \dfrac{\mathbb{E}_z\big[\mathcal{P}(B(\tau),B(T)) \,\big|\, A\big]}{\mathcal{P}(z,B(T))}$.

(ii) *If* $\mathbb{P}_z\big(\{B(\tau) \in \mathcal{B}(0,r)\} \,\big|\, A\big) = 1$, *then*

$$\Big(\frac{1-r}{1+r}\Big)^2 \mathbb{P}_z(A) \le \mathbb{P}_z\big(A \,\big|\, B(T)\big) \le \Big(\frac{1+r}{1-r}\Big)^2 \mathbb{P}_z(A) \quad \textit{almost surely.}$$

Proof. (i) Let $I \subset \partial\mathcal{B}(0,1)$ be a Borel set. Using the strong Markov property in the second step, we get

$$\begin{aligned}\mathbb{P}_z\big(A \,\big|\, \{B(T)\in I\}\big)\, \mathbb{P}_z\{B(T)\in I\} &= \mathbb{P}_z(A)\, \mathbb{P}_z\big(\{B(T)\in I\} \,\big|\, A\big)\\ &= \mathbb{P}_z(A)\, \mathbb{E}_z\big[\mathbb{P}_{B(\tau)}\big\{B(T)\in I\big\} \,\big|\, A\big].\end{aligned}$$

As a function of I, both sides of the equation define a finite measure with total mass $\mathbb{P}_z(A)$. Comparing the densities of the measures with respect to the surface measure on $\partial\mathcal{B}(0,1)$ gives

$$\mathbb{P}_z\big(A \,\big|\, B(T)\big)\, \mathcal{P}(z,B(T)) = \mathbb{P}_z(A)\, \mathbb{E}_z\big[\mathcal{P}(B(\tau),B(T)) \,\big|\, A\big].$$

(ii) The assumption of this part and (i) imply that the ratio $\mathbb{P}_z(A|B(T))/\mathbb{P}_z(A)$ can be written as an average of ratios $\mathcal{P}(u,w)/\mathcal{P}(z,w)$ where $w = B(T) \in \partial\mathcal{B}(0,1)$ and $u,z \in \mathcal{B}(0,r)$. The assertion follows by finding the minimum and maximum of $\mathcal{P}(u,w)$ as u ranges over $\mathcal{B}(0,r)$. ∎

The following lemma, concerning the common upcrossings of L Brownian excursions, will be the engine driving the proof of Theorem 9.24.

LEMMA 9.27. *Let $n > 5$ and let $\{x_1, \ldots, x_{4^{n-5}}\}$ be points such that the balls $\mathcal{B}(x_i, 2^{1-n})$ are disjoint and contained in the shell $\{z : \frac{1}{4} \leq |z| \leq \frac{3}{4}\}$. Consider L independent Brownian upcrossing excursions $W_1, \ldots, W_L$, started at prescribed points on $\partial\mathcal{B}(0,1)$ and stopped when they reach $\partial\mathcal{B}(0,2)$. Let S denote the number of centres x_i, $1 \leq i \leq 4^{n-5}$ such that the shell $\mathcal{B}(x_i, 2^{-n+1}) \setminus \mathcal{B}(x_i, 2^{-n})$ is upcrossed twice by each of $W_1, \ldots, W_L$. Then there exists constants $c, c_* > 0$ such that*

$$\mathbb{P}\big(\{S > 4^n (c/n)^L\} \,\big|\, A\big) \geq \frac{c_*^L}{L!}\,. \tag{3.1}$$

Moreover, the same estimate (with a suitable constant c_*) is valid if we condition on the end points of the excursions $W_1, \ldots, W_L$.

Proof of Lemma 9.27. By Lemma 9.25, for any $z \in \partial\mathcal{B}(0,1)$, the probability of Brownian motion starting at z hitting the ball $\mathcal{B}(x_i, 2^{-n})$ before reaching $\partial\mathcal{B}(0,2)$ is at least $\frac{1}{2n}$, and the probability of the second upcrossing excursion of $\mathcal{B}(x_i, 2^{-n+1}) \setminus \mathcal{B}(x_i, 2^{-n})$, when starting at $\partial\mathcal{B}(x_i, 2^{1-n})$ is at least $1/2$. Thus

$$\mathbb{E}S \geq 4^{n-5}(4n)^{-L}\,. \tag{3.2}$$

We now estimate the second moment of S. Consider a pair of centres x_i, x_j such that $2^{-m} \leq |x_i - x_j| \leq 2^{1-m}$ for some $m < n-1$. For each $k \leq L$, let $V_k = V_k(x_i, x_j)$ denote the event that the balls $\mathcal{B}(x_i, 2^{-n})$ and $\mathcal{B}(x_j, 2^{-n})$ are both visited by W_k. Given that $\mathcal{B}(x_i, 2^{-n})$ is reached first, the conditional probability that W_k will also visit $\mathcal{B}(x_j, 2^{-n})$ is at most $\frac{2m}{n}$. We conclude that $\mathbb{P}(V_k) \leq \frac{4m}{n^2}$ whence

$$\mathbb{P}\Big(\bigcap_{k=1}^{L} V_k\Big) \leq \Big(\frac{4m}{n^2}\Big)^L\,.$$

For each $m < n-1$ and $i \leq 4^{n-5}$, the number of centres x_j such that $2^{-m} \leq |x_i - x_j| \leq 2^{1-m}$ is at most 4^{n-m}. We deduce that there exists $C_1 > 0$ such that

$$\mathbb{E}S^2 \leq \frac{C_1^L 4^{2n}}{n^{2L}} \sum_{m=1}^{n} m^L 4^{-m} \leq \frac{(2C_1)^L 4^{2n} L!}{n^{2L}}\,, \tag{3.3}$$

where the last inequality follows, e.g., from taking $x = 1/4$ in the binomial identity

$$\sum_{m=0}^{\infty} \binom{m+L}{L} x^m = (1-x)^{-L-1}\,.$$

Now (3.2), (3.3) and the Paley-Zygmund inequality yield (3.1). The final statement of the lemma follows from Lemma 9.26. ∎

Proof of Theorem 9.24. Fix an increasing sequence $\{n_i \colon i \geq 1\}$ to be chosen later, and let $N_\ell = \sum_{i=1}^{\ell} n_i$. Denote $q_i = 4^{n_i} - 5$ and $Q_i = 4^{N_i - 5i}$. We begin by constructing a nested sequence of centres which we associate with a forest, i.e. a collection of trees, in the following manner. The first level of the forest consists of Q_1 centres, $\{x_1^{(1)}, \ldots, x_{Q_1}^{(1)}\}$, chosen such that the balls $\{\mathcal{B}(x_k^{(1)}, 2^{-N_1+1}) : 1 \leq k \leq Q_1\}$ are disjoint and contained in the annulus $\{z \colon \frac{1}{4} \leq |z| \leq \frac{3}{4}\}$.

Continue this construction recursively. For $\ell > 1$ suppose that level $\ell - 1$ of the forest has been constructed. Level ℓ consists of Q_ℓ vertices $\{x_1^{(\ell)}, \dots, x_{Q_\ell}^{(\ell)}\}$. Each vertex $x_i^{(\ell-1)}$, $1 \le i \le Q_{\ell-1}$, at level $\ell - 1$ has q_ℓ children $\{x_j^{(\ell)} : (i-1)q_\ell < j \le iq_\ell\}$ at level ℓ; the balls of radius $2^{-N_\ell+1}$ around these children are disjoint and contained in the annulus

$$\{z : \tfrac{1}{4}2^{-N_{\ell-1}} \le |z - x_i^{(\ell-1)}| \le \tfrac{3}{4}2^{-N_{\ell-1}}\}.$$

Recall that $T = \inf\{t > 0 : |B(t)| = 1\}$. We say that a level one vertex $x_k^{(1)}$ *survived* if Brownian motion upcrosses the shell $\mathcal{B}(x_k^{(1)}, 2^{-N_1+1}) \setminus \mathcal{B}(x_k^{(1)}, 2^{-N_1})$ twice before T. A vertex at the second level $x_k^{(2)}$ is said to have *survived* if its parent vertex survived, and in each upcrossing excursion of its parent, Brownian motion upcrosses the shell $\mathcal{B}(x_k^{(2)}, 2^{-N_2+1}) \setminus \mathcal{B}(x_k^{(2)}, 2^{-N_2})$ twice. Recursively, we say a vertex $x_k^{(\ell)}$, at level ℓ of the forest, *survived* if its parent vertex survived, and in each of the $2^{\ell-1}$ upcrossing excursions of its parent, Brownian motion upcrosses the shell $ball(x_k^{(\ell)}, 2^{-N_\ell+1}) \setminus \mathcal{B}(x_k^{(\ell)}, 2^{-N_\ell})$ twice. Also, for any $\ell \ge 1$, let S_ℓ denote the number of vertices at level ℓ of the forest that survived.

Using the notation of Lemma 9.27, denote $\Gamma_\ell = 4^{n_\ell}(c/n_\ell)^L$ and $p_\ell = \frac{c_*^L}{L!}$. where $L = L(\ell) = 2^{\ell-1}$. Lemma 9.27 with $n = n_1$ states that

$$\mathbb{P}\{S_1 > \Gamma_1\} \ge p_1 = c_* \tag{3.4}$$

For $\ell > 1$, the same lemma, and independence of excursions in disjoint shells given their endpoints, yield

$$\mathbb{P}\big(\{S_{\ell+1} \le \Gamma_{\ell+1}\} \,\big|\, \{S_\ell > \Gamma_\ell\}\big) \le (1 - p_{\ell+1})^{\Gamma_\ell} \le \exp(-p_{\ell+1}\Gamma_\ell)\,. \tag{3.5}$$

By picking n_ℓ large enough, we can ensure that $p_{\ell+1}\Gamma_\ell > \ell$, whence the right-hand side of (3.5) is summable in ℓ. Consequently

$$\alpha = \mathbb{P}\Big(\bigcap_{\ell=1}^{\infty}\{S_\ell > \Gamma_\ell\}\Big) = \mathbb{P}\{S_1 > \Gamma_1\}\prod_{\ell=1}^{\infty}\mathbb{P}\big(\{S_{\ell+1} > \Gamma_{\ell+1}\} \,\big|\, \{S_\ell > \Gamma_\ell\}\big) > 0\,. \tag{3.6}$$

Thus with probability at least α, there is a nested sequence of closed balls $\overline{\mathcal{B}}(x_{k(\ell)}^{(\ell)}, 2^{-N_\ell})$ for $\ell = 1, 2, \dots$ such that all their centres survive. The intersection of such a nested sequence yields a point visited by Brownian motion uncountably many times before it exits the unit disk.

Let H_r denote the event that Brownian motion, killed on exiting $\mathcal{B}(0, r)$, has a point of uncountable multiplicity. As explained above, (3.6) implies that $\mathbb{P}(H_1) \ge \alpha > 0$. By Brownian scaling, $\mathbb{P}(H_r)$ does not depend on r, whence

$$\mathbb{P}\Big(\bigcap_{n=1}^{\infty} H_{1/n}\Big) \ge \alpha\,.$$

The Blumenthal zero-one law implies that this intersection has probability 1, so there are points of uncountable multiplicity almost surely. ∎

4. Kaufman's dimension doubling theorem

In Theorem 4.37 we have seen that d-dimensional Brownian motion maps any set of dimension α almost surely into a set of dimension 2α. Surprisingly, by a famous result of Kaufman, the dimension doubling property holds almost surely *simultaneously* for all sets.

THEOREM 9.28 (Kaufman 1969). *Let $\{B(t) : t \geq 0\}$ be Brownian motion in dimension $d \geq 2$. Almost surely, for any $A \subset [0, \infty)$, we have*

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

The power of this result lies in the fact that the dimension doubling formula can now be applied to arbitrary random sets. As a first application we ask, how big the sets

$$T(x) = \big\{t \geq 0 \,:\, B(t) = x\big\}$$

of times mapped by d-dimensional Brownian motion onto the same point x can possibly be. We have seen so far in this chapter and Theorem 6.38 that, almost surely,

- in dimension $d \geq 4$ 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,
- in dimension $d = 1$ all sets $T(x)$ have at least Hausdorff dimension $\frac{1}{2}$.

We now use Kaufman's theorem to determine the Hausdorff dimension of the sets $T(x)$ in the case of planar and linear Brownian motion.

COROLLARY 9.29. *Suppose $\{B(t)\colon t \geq 0\}$ is a planar Brownian motion. Then, almost surely, for all $x \in \mathbb{R}^2$, we have* $\dim T(x) = 0$.

Proof. By Kaufman's theorem we have $\dim T(x) = \frac{1}{2} \dim\{x\} = 0$. ■

COROLLARY 9.30. *Suppose $\{B(t)\colon t \geq 0\}$ is a linear Brownian motion. Then, almost surely, for all $x \in \mathbb{R}$, we have* $\dim T(x) = \frac{1}{2}$.

Proof. The lower bound was shown in Theorem 6.38. For the upper bound let $\{W(t) : t \geq 0\}$ be a Brownian motion independent of $\{B(t) : t \geq 0\}$. Applying Kaufman's theorem for the planar Brownian motion given by $\widetilde{B}(t) = (B(t), W(t))$ we get, almost surely, $\dim T(x) = \dim \widetilde{B}^{-1}(\{x\} \times \mathbb{R}) \leq \frac{1}{2} \dim(\{x\} \times \mathbb{R}) = \frac{1}{2}$. ■

We now prove Kaufman's theorem, first in the case $d \geq 3$. Note first that $\dim B(A) \leq 2 \dim A$ holds in all dimensions and for all sets $A \subset [0, \infty)$ if $\{B(t) : t \geq 0\}$ is α-Hölder continuous for any $\alpha < \frac{1}{2}$. By Corollary 1.20 this holds almost surely. Hence only the lower bound $\dim B(A) \geq 2 \dim A$ requires proof. The crucial idea here is that one uses a standardised covering of $B(A)$ by dyadic cubes and ensures that, simultaneously for all possible covering cubes the preimages allow an efficient covering. An upper bound for $\dim A$ follows by selecting from the coverings of all preimages.

LEMMA 9.31. *Consider a cube $Q \subset \mathbb{R}^d$ centred at a point x and having diameter $2r$. Let $\{B(t) : t \geq 0\}$ be d-dimensional Brownian motion, with $d \geq 3$. Define recursively*

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

with the usual convention that $\inf \emptyset = \infty$. *Then there exists $0 < \theta < 1$ depending only on the dimension d, such that $\mathbb{P}_z\{\tau_{n+1}^Q < \infty\} \leq \theta^n$ for all $z \in \mathbb{R}^2$ and $n \in \mathbb{N}$.*

Proof. It is sufficient to show that for some θ as above,

$$\mathbb{P}_z \left\{\tau_{k+1}^Q = \infty \,\middle|\, \tau_k^Q < \infty\right\} > 1 - \theta.$$

But the quantity on the left can be bounded from below by

$$\mathbb{P}_z\big\{\tau_{k+1}^Q = \infty \,\big|\, |B(\tau_k^Q + r^2) - x| > 2r, \tau_k^Q < \infty\big\}\mathbb{P}_z\big\{|B(\tau_k^Q + r^2) - x| > 2r \,\big|\, \tau_k^Q < \infty\big\}$$

The second factor is clearly positive, by the strong Markov property, and the first is also positive since Brownian motion is transient in $d \geq 3$. Both probabilities are invariant under changing the scaling factor r. ∎

LEMMA 9.32. *Let $\mathfrak{C}_m$ denote the set of dyadic cubes of side length 2^{-m} inside $[-\frac{1}{2}, \frac{1}{2}]^d$. Almost surely there exists a random variable $C = C(\omega)$ so that for all m and for all cubes $Q \in \mathfrak{C}_m$ we have $\tau_{\lceil mC+1 \rceil}^Q = \infty$.*

Proof. From Lemma 9.31 we get that

$$\sum_{m=1}^{\infty} \sum_{Q \in \mathfrak{C}_m} \mathbb{P}\big\{\tau_{\lceil cm+1 \rceil}^Q < \infty\big\} \leq \sum_{m=1}^{\infty} 2^{dm} \theta^{cm}.$$

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 $\tau_{\lceil cm+1 \rceil + 1}^Q = \infty$ for all $Q \in \mathfrak{C}_m$. Finally, we can choose a random $C(\omega) > c$ to handle the finitely many exceptional cubes. ∎

Proof of Theorem 9.28 for $d \geq 3$. As mentioned before we can focus on the '$\geq$' direction. We fix L and show that, almost surely, for all subsets S of $[-L, L]^d$ we have

$$\dim B^{-1}(S) \leq \tfrac{1}{2} \dim S. \tag{4.1}$$

Applying this to $S = B(A) \cap [-L, L]^d$ successively for a countable unbounded set of L we get the desired conclusion. By scaling, it is sufficient to prove (4.1) for $L = 1/2$. The idea now is to verify (4.1) for all paths satisfying Lemma 9.32 using completely deterministic reasoning. As this set of paths has full measure, this verifies the statement.

Hence fix a path $\{B(t) : t \geq 0\}$ satisfying Lemma 9.32 for a constant $C > 0$. If $\beta > \dim S$ and $\varepsilon > 0$ there exists a covering of S by binary cubes $\{Q_j : j \in \mathbb{N}\} \subset \bigcup_{m=1}^{\infty} \mathfrak{C}_m$ such that $\sum |Q_j|^\beta < \varepsilon$. If N_m denotes the number of cubes from $\mathfrak{C}_m$ in such a covering, then

$$\sum_{m=1}^{\infty} N_m \, 2^{-m\beta} < \varepsilon.$$

Consider the inverse image of these cubes under $\{B(t) : t \geq 0\}$. Since we chose this path so that Corollary 9.32 is satisfied, this yields a covering of $B^{-1}(S)$, which for each $m \geq 1$ uses at most CmN_m intervals of length $r^2 = d2^{-2m}$.

For $\gamma > \beta$ we can bound the $\gamma/2$-dimensional Hausdorff content of $B^{-1}(S)$ from above by

$$\sum_{m=1}^{\infty} Cm\, N_m (d2^{-2m})^{\gamma/2} = C\, d^{\gamma/2} \sum_{m=1}^{\infty} m\, N_m\, 2^{-m\gamma}.$$

This can be made small by choosing a suitable $\varepsilon > 0$. Thus $B^{-1}(S)$ has Hausdorff dimension at most $\gamma/2$ for all $\gamma > \beta > \dim S$, and therefore $\dim B^{-1}(S) \leq \dim S/2$. ■

In $d = 2$ we cannot rely on transience of Brownian motion. To get around this problem, we can look at the Brownian path up to a stopping time. A convenient choice of stopping time for this purpose is $\tau_R^* = \min\{t : |B(t)| = R\}$. For the two dimensional version of Kaufman's theorem it is sufficient to show that, almost surely,

$$\dim B(A) \geq 2 \dim(A \cap [0, \tau_R^*]) \text{ for all } A \subset [0, \infty).$$

Lemma 9.31 has to be changed accordingly.

Lemma 9.33. *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 0. Let $\{B(t) : t \geq 0\}$ be planar Brownian motion. Define τ_k^Q as in (4.1), and assume that the cube Q is inside the ball of radius R about the origin. Then there exists $c = c(R) > 0$ such that, with $2^{-m-1} < r < 2^{-m}$, for any $z \in \mathbb{R}^2$,*

$$\mathbb{P}_z\{\tau_k^Q < \tau_R^*\} \leq \Big(1 - \frac{c}{m}\Big)^k \leq e^{-ck/m}. \tag{4.2}$$

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

$$\mathbb{P}_z\big\{\tau_{k+1}^Q \geq \tau_R^* \,\big|\, |B(\tau_k^Q + r^2) - x| > 2r, \tau_k^Q < \tau_R^*\big\}\, \mathbb{P}_z\big\{|B(\tau_k^Q + r^2) - x| > 2r \,\big|\, \tau_k^Q < \tau_R^*\big\}.$$

The second factor does not depend on r and R, and it can be bounded from below by a constant. 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 $\partial\mathcal{B}(0, r)$. Using Theorem 3.17 this probability is given by

$$\frac{\log 2r - \log r}{\log 2R - \log r} \geq \frac{1}{\log_2(R + 1 + m)}.$$

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

The bound (4.2) on $\mathbb{P}\{\tau_k^Q < \infty\}$ in two dimensions is worse by a linear factor than the bound in higher dimensions. This, however, does not make a significant difference in the proof of the two dimensional version of Theorem 9.28, which can now be completed in the same way.

There is also a version of Kaufman's theorem for Brownian motion in dimension one.

Theorem 9.34. *Suppose $\{B(t) : t \geq 0\}$ is a linear Brownian motion. Then, almost surely, for all nonempty closed sets $S \subset \mathbb{R}$, we have*

$$\dim B^{-1}(S) = \tfrac{1}{2} + \tfrac{1}{2} \dim S.$$

Remark 9.35. Note that here it is essential to run Brownian motion on an unbounded time interval. For example, for the point $x = \max_{0\le t\le 1} B(t)$ the set $\{t \in [0,1] : B(t) = x\}$ is a singleton almost surely. The restriction to closed sets comes from the Frostman lemma, which we have proved for closed sets only, and can be relaxed accordingly. ◇

Proof. For the proof of the upper bound let $\{W(t) : t \ge 0\}$ be a Brownian motion independent of $\{B(t) : t \ge 0\}$. Applying Kaufman's theorem for the planar Brownian motion given by $\widetilde{B}(t) = (B(t), W(t))$ we get almost surely, for all $S \subset \mathbb{R}$,

$$\dim B^{-1}(S) = \dim \widetilde{B}^{-1}(S \times \mathbb{R}) \le \tfrac{1}{2} \dim(S \times \mathbb{R}) = \tfrac{1}{2} + \tfrac{1}{2} \dim S,$$

where we have used the straightforward fact that $\dim(S \times \mathbb{R}) = 1 + \dim S$.

The lower bound requires a more complicated argument, based on Frostman's lemma. For this purpose we may suppose that $S \subset (-M, M)$ is closed and $\dim S > \alpha$. Then there exists a measure μ supported by S such that

$$\mu(\mathcal{B}(x,r)) \le r^\alpha \qquad \text{for all } x \in S,\ 0 < r < 1.$$

Let ℓ^a be the measure with cumulative distribution function given by the local time at level a. Let ν be the measure on $B^{-1}(S)$ given by

$$\nu(A) = \int \mu(da)\, \ell^a(A), \qquad \text{for } A \subset [0,\infty) \text{ Borel.}$$

Then, by Theorem 6.18, one can find a constant $C > 0$ such that

$$\ell^a(\mathcal{B}(x,r)) = L^a(x+r) - L^a(x-r) \le C r^{\frac{1}{2}-\varepsilon}$$

for all $a \in [-M, M]$ and $0 < r < 1$. By Hölder continuity of Brownian motion there exists, for given $\varepsilon > 0$, a constant $c > 0$ such that, for every $x \in [0,1]$,

$$|B(x+s) - B(x)| \le c r^{\frac{1}{2}-\varepsilon} \text{ for all } s \in [-r, r].$$

From this we get the estimate

$$\begin{aligned}\nu(\mathcal{B}(x,r)) = \int \mu(da)\big[L^a(x+r) - L^a(x-r)\big] &\le \int_{B(x)-cr^{\frac{1}{2}-\varepsilon}}^{B(x)+cr^{\frac{1}{2}-\varepsilon}} \mu(da)\big[L^a(x+r) - L^a(x-r)\big] \\ &\le c^\alpha r^{\frac{\alpha}{2}-\varepsilon\alpha} C r^{\frac{1}{2}-\varepsilon} \qquad \text{for all } x \in S,\ 0 < r < 1.\end{aligned}$$

Hence, by the mass distribution principle, we get the lower bound $\alpha/2 + 1/2 - \varepsilon(1+\alpha)$ for the dimension and the result follows when $\varepsilon \downarrow 0$ and $\alpha \uparrow \dim S$. ■

As briefly remarked in the discussion following Theorem 4.37, Brownian motion is also 'capacity-doubling'. This fact holds for a very general class of kernels, we give an elegant proof of this fact here.

Theorem 9.36. *Let $\{B(t) : t \in [0,1]\}$ be d-dimensional Brownian motion and $A \subset [0,1]$ a closed set. Suppose f is decreasing and there is a constant $C > 0$ with*

$$\int_0^1 \frac{f(r^2 x)}{f(x)}\, r^{d-1}\, dr \le C \text{ for all } x \in (0,1)\,, \tag{4.3}$$

and let $\phi(x) = x^2$. Then, almost surely,

$$\mathrm{Cap}_f\big(A\big) > 0 \qquad \textit{if and only if} \qquad \mathrm{Cap}_{f\circ\phi}\big(B(A)\big) > 0\,.$$

REMARK 9.37. Condition (4.3) is only used in the 'only if' part of the statement. Note that if $f(x) = x^\alpha$ is a power law, then (4.3) holds if and only if $2\alpha < d$. ◇

Proof. We start with the 'only if' direction, which is easier. Suppose $\mathrm{Cap}_f(A) > 0$. This implies that there is a mass distribution μ on A such that the f-energy of μ is finite. Then $\mu \circ B^{-1}$ is a mass distribution on $B(A)$ and we will show that it has finite $f \circ \phi$-energy. Indeed,

$$\begin{aligned} I_{f\circ\phi}\big(\mu \circ B^{-1}\big) &= \iint f \circ \phi\big(|x-y|\big)\, \mu \circ B^{-1}(dx)\, \mu \circ B^{-1}(dy) \\ &= \iint f\big(|B(s) - B(t)|^2\big)\, \mu(ds)\, \mu(dt)\,. \end{aligned}$$

Hence,

$$\mathbb{E} I_{f\circ\phi}\big(\mu \circ B^{-1}\big) = \iint \mathbb{E} f\big(|X|^2\, |s-t|\big)\, \mu(ds)\, \mu(dt)\,,$$

where X is a d-dimensional standard normal random variable. Using polar coordinates and the monotonicity of f we get, for a constant $\kappa(d)$ depending only on the dimension,

$$\begin{aligned} \mathbb{E}\big[f\big(|X|^2\,|s-t|\big)\big] &= \kappa(d) \int_0^\infty f(r^2\,|s-t|)\, e^{-r^2/2}\, r^{d-1}\, dr \\ &\le f(|s-t|)\, \kappa(d) \Big(\int_0^1 \frac{f(r^2\,|s-t|)}{f(|s-t|)\, r^{1-d}}\, dr + \int_1^\infty e^{-r^2/2}\, r^{d-1}\, dr \Big)\,. \end{aligned}$$

By (4.3) the bracket on the right hand side is bounded by a constant independent of $|s-t|$, and hence $\mathbb{E}[I_{f\circ\phi}(\mu \circ B^{-1})] < \infty$, which in particular implies $I_{f\circ\phi}(\mu) < \infty$ almost surely.

The difficulty in the 'if' direction is that a measure on $B(A)$ with finite $f\circ\phi$-energy cannot easily be transported backwards onto A. To circumvent this problem we use the characterisation of capacity in terms of polarity with respect to percolation limit sets, recall Theorem 9.18.

Fix a unit cube Cube such that $\mathrm{Cap}_{f\circ\phi}(B(A)\cap\mathsf{Cube}) > 0$ with positive probability, and let Γ be a percolation limit set with retention probabilities associated to the decreasing function $f(x^2/4)$ as in Theorem 9.18, which is independent of Brownian motion. Then, by Theorem 9.18, we have $B(A) \cap \Gamma \neq \emptyset$ with positive probability. Define a random variable

$$T = \inf\big\{t \in A \,:\, B(t) \in \Gamma\big\}\,,$$

which is finite with positive probability. Hence the measure μ given by

$$\mu(B) = \mathbb{P}\big\{B(T) \in B,\, T < \infty\big\}$$

is a mass distribution on A. We shall show that it has finite f-energy, which completes the proof. Again we use the polarity criterion of Theorem 9.18 to do this. Let $\mathsf{S}_n = \bigcup_{S\in\mathfrak{S}_n} S$ be the union of all cubes retained in the construction up to step n. Then, by looking at the retention probability of any fixed point in Cube, we have, for any $s \in A$,

$$\mathbb{P}\big\{B(s) \in \mathsf{S}_n\big\} \le p_1 \cdots p_n \le C\, \frac{1}{f \circ \phi(2^{-n-1})}\,. \tag{4.4}$$

Conversely, by a first entrance decomposition,

$$\begin{aligned}\mathbb{P}\big\{B(s) \in \mathsf{S}_n\big\} &\geq \mathbb{P}\big\{B(s) \in \mathsf{S}_n,\, B(T) \in \mathsf{S}_n,\, T < \infty\big\} \\ &= \int_0^s \mu(dt)\, \mathbb{P}\big\{B(s) \in \mathsf{S}_n \,\big|\, B(t) \in \mathsf{S}_n\big\}\end{aligned}$$

Given $B(t) \in \mathsf{S}_n$ and $\sqrt{s-t} \leq 2^{-n+k}$ for some $k \in \{0, \dots, n\}$, the probability that $B(s)$ and $B(t)$ are contained in the same dyadic cube $Q \in \mathfrak{C}_{n-k}$ is bounded from below by a constant. Given this event, we know that Q is retained in the percolation (otherwise we could not have $B(t) \in \mathsf{S}_n$) and the probability that the cube in $\mathfrak{C}_n$, that contains $B(s)$, is retained in the percolation is at least $p_{n-k+1} \cdots p_n$. Therefore

$$\begin{aligned}\int_0^s \mu(dt)\, \mathbb{P}\big\{B(s) \in \mathsf{S}_n \,\big|\, B(t) \in \mathsf{S}_n\big\} &\geq \sum_{k=0}^n \mu\big([s - 2^{-2n+2k}, s - 2^{-2n+2k-2})\big)\, p_{n-k+1} \cdots p_n \\ &\geq c \sum_{k=0}^n \mu\big([s - 2^{-2n+2k}, s - 2^{-2n+2k-2})\big)\, \frac{f \circ \phi(2^{-n+k-1})}{f \circ \phi(2^{-n-1})} \\ &\geq c\, \frac{1}{f \circ \phi(2^{-n-1})} \int_0^{s-2^{-2n-2}} \mu(dt)\, f\big(s-t\big)\,.\end{aligned}$$

Finiteness of the f-energy follows by comparing this with (4.4), cancelling the factor $1/f \circ \phi(2^{-n-1})$, integrating over $\mu(ds)$, and letting $n \to \infty$. This completes the proof. ■

Exercises

Exercise 9.1.

(a) Suppose that $\{B_1(t)\colon t \geq 0\}, \{B_2(t)\colon t \geq 0\}$ are independent standard Brownian motions in $\mathbb{R}^3$. Then, almost surely, $B_1[0,t] \cap B_2[0,t] \neq \emptyset$ for any $t > 0$.

(b) Suppose that $\{B_1(t)\colon t \geq 0\}, \ldots, \{B_p(t)\colon t \geq 0\}$ are p independent standard Brownian motions in $\mathbb{R}^d$, and $d > d(p-2)$. Then, almost surely,

$$\dim \big(B_1[0,t_1] \cap \cdots \cap B_p[0,t_p]\big) = d - p(d-2) \qquad \text{for any } t_1, \ldots, t_p > 0.$$

Exercise 9.2 (∗). For a d-dimensional Brownian motion $\{B(t)\colon t \geq 0\}$ we denote by

$$S(p) = \big\{x \in \mathbb{R}^d \,:\, \exists 0 < t_1 < \cdots < t_p < 1 \text{ with } x = B(t_1) = \cdots = B(t_p)\big\}$$

the set of p-fold multiple points. Show that, for $d > p\,(d-2)$,

(a) $\dim S(p) = d - p\,(d-2)$, almost surely.
(b) for any closed set Λ, we have

$$\mathbb{P}\big\{S(p) \cap \Lambda \neq \emptyset\big\} > 0 \qquad \text{if and only if} \qquad \mathrm{Cap}_{f^p}(\Lambda) > 0\,,$$

where the decreasing function f is the radial potential.

Exercise 9.3.

(a) Let A be a set of rooted trees. We say that A is *inherited* if every finite tree is in A, and if $T \in A$ and $v \in V$ is a vertex of the tree then the tree $T(v)$, consisting of all successors of v, is in A.

Prove the *Galton-Watson 0–1 law*: For a Galton-Watson tree, conditional on survival, every inherited set has probability zero or one.

(b) Show that for the percolation limit sets $\Gamma[\gamma] \subset \mathbb{R}^d$ with $0 < \gamma < d$ we have

$$\mathbb{P}\big\{\, \dim \Gamma[\gamma] = d - \gamma \mid \Gamma[\gamma] \neq \emptyset \big\} = 1.$$

Exercise 9.4 (∗). Consider a standard Brownian motion $\{B(t)\colon t \geq 0\}$ and let $A_1, A_2 \subset [0, \infty)$.

(a) Show that if $\dim(A_1 \times A_2) < 1/2$ then $\mathbb{P}\{B(A_1) \text{ intersects } B(A_2)\} = 0$.

(b) Derive the same conclusion under the weaker assumption that $A_1 \times A_2$ has vanishing 1/2-dimensional Hausdorff measure.

(c) Show that if $\mathrm{Cap}_{1/2}(A_1 \times A_2) > 0$, then $\mathbb{P}\{B(A_1) \text{ intersects } B(A_2)\} > 0$.

(d) Find a set $A \subset [0, \infty)$ such that the probability that $\{B(t) \colon t \geq 0\}$ is one-to-one on A is strictly between zero and one.

Exercise 9.5 (∗). Let $\{B(t) \colon 0 \leq t \leq 1\}$ be a planar Brownian motion. For every $a \in \mathbb{R}$ define the sets

$$S(a) = \{y \in \mathbb{R} \colon (a, y) \in B[0, t]\},$$

the vertical slices of the path. Show that, almost surely,

$$\dim S(a) = 1,$$

for every $a \in (\min\{x \colon (x, y) \in B[0, t]\}, \max\{x \colon (x, y) \in B[0, t]\})$.

Exercise 9.6. Let $\{B(t) : t \geq 0\}$ be Brownian motion in dimension $d \geq 2$.
Show that, almost surely, for any $A \subset [0, \infty)$, we have

$$\overline{\dim}_{\mathrm{M}} B(A) = 2\, \overline{\dim}_{\mathrm{M}} A \qquad \text{and} \qquad \underline{\dim}_{\mathrm{M}} B(A) = 2\, \underline{\dim}_{\mathrm{M}} A\,.$$

Notes and Comments

The question whether there exist p-multiple points of a d-dimensional Brownian motion was solved in various stages in the early 1950s. First, Lévy showed in [**Le40**] that almost all paths of a planar Brownian motion have double points, and Kakutani [**Ka44a**] showed that if $n \geq 5$ almost no paths have double points. The cases of $d = 3, 4$ where added by Dvoretzky, Erdős and Kakutani in [**DEK50**] and the same authors showed in [**DEK54**] that planar Brownian motion has points of arbitrary multiplicity. Finally, Dvoretzky, Erdős, Kakutani and Taylor showed in [**DEKT57**] that there are no triple points in $d = 3$. Clearly the existence of p-fold multiple points is essentially equivalent to the problem whether p independent Brownian motions have a common intersection.

The problem of finding the Hausdorff dimension of the set of p-fold multiple points in the plane, and of double points in $\mathbb{R}^3$, was still open when Itô and McKean wrote their influential book on the sample paths of diffusions, see [**IM74**, p. 261] in 1964, but was solved soon after by Taylor [**Ta66**] and Fristedt [**Fr67**]. Perkins and Taylor [**PT88**] provide fine results when Brownian paths in higher dimensions 'come close' to self-intersecting. The method of stochastic codimension, which we use to find these dimensions, is due to Taylor [**Ta66**], who used the range of stable processes as 'test sets'. The restriction of the stable indices to the range $\alpha \in (0, 2]$ leads to complications, which can be overcome by a projection method of Fristedt [**Fr67**] or by using multiparameter processes [**Kh02**]. The use of percolation limit sets as test sets is much more recent and due to Khoshnevisan, Peres and Xiao [**KPX00**], though similar ideas are used in the context of trees ar least since the pioneering work of Lyons [**Ly90**]. The latter paper is also the essential source for our proof of Hawkes' theorem.

Some very elegant proofs of these classical facts were given later: Rosen [**Ro83**] provides a local time approach, and Kahane [**Ka86**] proves a general formula for the intersection of independent random sets satisfying suitable conditions. The bottom line of Kahane's approach is that the formula 'codimension of the intersection is equal to the sum of codimensions of the intersected sets' which is well-known from linear subspaces in general position can be extended to the Hausdorff dimension of a large class of random sets, which includes the paths of Brownian motion, see also [**Fa97a, Ma95**].

The intersection equivalence approach we describe in Section 2 is taken from [**Pe96a**], [**Pe96b**]. The proof of Lyons' theorem we give is taken from [**BPP95**]. See [**Ly92**, Theorem 2.1] for the original proof.

Hendricks and Taylor conjectured in 1976 a characterisation of the polar sets for the multiple points of a Brownian motion or a more general Markov process, which included the statement of Theorem 9.21. Sufficiency of the capacity criterion in Theorem 9.21 was proved by Evans [**Ev87a, Ev87b**] and independently by Tongring [**To88**], see also Le Gall, Rosen and Shieh [**LRS89**]. The full equivalence was later proved in a much more general setting by Fitzsimmons and Salisbury [**FS89**]. A quantitative treatment of the question, which sets contain double points of Brownian motion is given in [**PP07**].

Points of multiplicity strictly n where identified by Adelman and Dvoretzky [**AD85**] and the result is also an immediate consequence of the exact Hausdorff gauge function identified by Le Gall [**LG86**].

The existence of points of infinite multiplicity in the planar case was first stated in [**DEK58**] though their proof seems to have a gap. Le Gall [**LG87**] proves a stronger result: Two sets $A, B \subset \mathbb{R}$ are said to be of the same *order type* if there exists an increasing homeomorphism ϕ of $\mathbb{R}$ such that $\phi(A) = B$. Le Gall shows that, for any totally disconnected, compact $A \subset \mathbb{R}$, almost surely there exists a point $x \in \mathbb{R}^2$ such that the set $\{t \geq 0 : B(t) = x\}$ has the same order type as A. In particular, there exist points of countably infinite and uncountable multiplicity. Le Gall's proof is based on the properties of natural measures on the intersection of Brownian paths. Our proof avoids this and seems to be new, though it uses arguments of Le Gall's proof as well as some techniques from [**KM05**].

Substantial generalisations of Exercise 9.4 can be found in papers by Khoshnevisan [**Kh99**] and Khoshnevisan and Xiao [**KX05**]. For example, in [**Kh99**, Theorem 8.2] it is shown that the condition in part (c) is an equivalence.

Kaufman proved his dimension doubling theorem in [**Ka69**]. The version for Brownian motion in dimension one is due to Serlet [**Se95**].

The capacity-doubling result in the given generality is new, but Khoshnevisan and Xiao [**KX05**, Question 1.1, Theorem 7.1] prove the special case when f is a power law using a different method. Their argument is based on the investigation of additive Lévy processes and works for a class of processes much more general than Brownian motion. Theorem 9.36 does not hold uniformly for all sets A. Examples can be constructed along the lines in [**PT87**].

In this book we do not construct a measure on the intersection of p Brownian paths. However this is possible and yields the *intersection local time* first studied by Geman, Horowitz and Rosen [**GHR84**], see also [**Ro83**]. This quantity plays a key role in the analysis of Brownian paths and [**LG91**] gives a very accessible account of the state of research in 1991, which is still worth reading. Recent research deals with fine Hausdorff dimension properties of the intersections, see for example [**KM02, KM05**].

CHAPTER 10

Exceptional sets for Brownian motion

The techniques developed in this book so far give a fairly satisfactory picture of the behaviour of a Brownian motion at a typical time, like a fixed time or a stopping time. In this chapter we explore exceptional times, for example times where the path moves slower or faster than in the law of the iterated logarithm, or does not wind as in Spitzer's law. Again Hausdorff dimension is the right tool to describe just how exceptional an exceptional behaviour is, but we shall see that another notion of dimension, the packing dimension, can provide additional insight.

1. The fast times of Brownian motion

In a famous paper from 1974, Orey and Taylor raise the question, how often, on a Brownian path, the law of the iterated logarithm fails. To understand this, recall that, by Corollary 5.3 and the Markov property, for a linear Brownian motion $\{B(t)\colon t \geq 0\}$ and for every $t \in [0,1]$, almost surely,

$$\limsup_{h \downarrow 0} \frac{|B(t+h) - B(t)|}{\sqrt{2h \log\log(1/h)}} = 1.$$

This contrasts sharply with the following result (note the absence of the iterated logarithm!).

THEOREM 10.1. *Almost surely, we have*

$$\max_{0 \leq t \leq 1} \limsup_{h \downarrow 0} \frac{|B(t+h) - B(t)|}{\sqrt{2h \log(1/h)}} = 1.$$

REMARK 10.2. At the time $t \in [0,1]$ where the maximum in Theorem 10.1 is attained, the law of the iterated logarithm fails and it is therefore an *exceptional* time. $\diamond$

Proof. The upper bound follows from Lévy's modulus of continuity, Theorem 1.14, as

$$\sup_{0 \leq t < 1} \limsup_{h \downarrow 0} \frac{|B(t+h) - B(t)|}{\sqrt{2h \log(1/h)}} \leq \limsup_{h \downarrow 0} \sup_{0 \leq t \leq 1-h} \frac{|B(t+h) - B(t)|}{\sqrt{2h \log(1/h)}} = 1.$$

Readers who have skipped the proof of Theorem 1.14 given in Chapter 1 will be able to infer the upper bound directly from Remark 10.5 below. It remains to show that there exists a time $t \in [0,1]$ such that

$$\limsup_{h \downarrow 0} \frac{|B(t+h) - B(t)|}{\sqrt{2h \log(1/h)}} \geq 1.$$

Recall from Proposition 1.13 that, almost surely, for every constant $c < \sqrt{2}$ and every $\varepsilon, \delta > 0$ there exists $t \in [0, \delta]$ and $0 < h < \varepsilon$ with

$$\big|B(t+h) - B(t)\big| > c\sqrt{h \log(1/h)}.$$

Using the Markov property this implies that, for $c < \sqrt{2}$, the sets

$$M(c, \varepsilon) = \big\{t \in [0,1] \,:\, \text{ there is } h \in (0, \varepsilon) \text{ such that } \big|B(t+h) - B(t)\big| > c\sqrt{h \log(1/h)}\big\}$$

are almost surely dense in $[0,1]$. By continuity of Brownian motion they are open, and clearly $M(c, \varepsilon) \subset M(d, \delta)$ whenever $c > d$ and $\varepsilon < \delta$. Hence, by Baire's (category) theorem, the intersection

$$\bigcap_{\substack{c<\sqrt{2} \\ \varepsilon>0}} M(c, \varepsilon) = \Big\{t \in [0,1] \,:\, \limsup_{h \downarrow 0} \frac{\big|B(t+h) - B(t)\big|}{\sqrt{2h \log(1/h)}} \geq 1\Big\}$$

is dense and hence nonempty almost surely. ∎

To explore how often we come close to the exceptional behaviour described in Theorem 10.1 we introduce a spectrum of exceptional points. Given $a > 0$ we call a time $t \in [0,1]$ an **a-fast time** if

$$\limsup_{h \downarrow 0} \frac{|B(t+h) - B(t)|}{\sqrt{2h \log(1/h)}} \geq a\,,$$

and $t \in [0,1]$ is a **fast time** if it is a-fast for some $a > 0$. By Theorem 10.1 fast times exist, in fact the proof even shows that the set of fast times is dense in $[0,1]$ and hence is infinite. Conversely it is immediate from the law of the iterated logarithm that the set has Lebesgue measure zero, recall Remark 1.28. The appropriate notion to measure the quantity of a-fast times is, again, Hausdorff dimension.

Theorem 10.3 (Orey and Taylor 1974). *Suppose $\{B(t) \colon t \geq 0\}$ is a linear Brownian motion. Then, for every $a \in [0,1]$, we have almost surely,*

$$\dim \Big\{t \in [0,1] \,:\, \limsup_{h \downarrow 0} \frac{|B(t+h) - B(t)|}{\sqrt{2h \log(1/h)}} \geq a\Big\} = 1 - a^2\,.$$

The rest of this section is devoted to the proof of this result. We start with a proof of the *upper bound*, which also shows that there are almost surely no a-fast points for $a > 1$.

So fix an arbitrary $a > 0$. Let $\varepsilon > 0$ and $\eta > 1$, having in mind that we later let $\eta \downarrow 1$ and $\varepsilon \downarrow 0$. The basic idea is to cover the interval $[0,1)$ by a collection of intervals of the form $[j\eta^{-k}, (j+1)\eta^{-k})$ with $j = 0, \ldots, \lceil \eta^k - 1 \rceil$ and $k \geq 1$. Any such interval of length $h := \eta^{-k}$ is included in the covering if, for $h' := k\eta^{-k}$,

$$|B(jh + h') - B(jh)| > a(1 - 4\varepsilon)\sqrt{2h' \log(1/h')}.$$

Let $\mathfrak{I}_k = \mathfrak{I}_k(\eta,\ \varepsilon)$ be the collection of intervals of length η^{-k} chosen in this procedure.

LEMMA 10.4. *Almost surely, for every $\varepsilon > 0$ and $\delta > 0$, there is an $\eta > 1$ and $m \in \mathbb{N}$ such that the collection $\mathfrak{I} = \mathfrak{I}(\varepsilon, \delta) = \{I \in \mathfrak{I}_k : k \geq m\}$ is a covering of the set of a-fast points consisting of intervals of diameter no bigger than δ.*

Proof. We first note that by Theorem 1.12 there exists a constant $C > 0$ such that, almost surely, there exists $\rho > 0$ such that, for all $s, t \in [0, 2]$ with $|s - t| \leq \rho$,

$$\big|B(s) - B(t)\big| \leq C\sqrt{|s-t|\log(1/|s-t|)}. \tag{1.1}$$

Choose $\eta > 1$ such that $\sqrt{\eta - 1} \leq a\varepsilon/C$. Let M be the minimal integer with $M\eta^{-M} \leq \rho$ and $m \geq M$ such that $m\eta^{-m} < \delta$ and $k\eta^{-k} < \ell\eta^{-\ell}$ for all $k > \ell \geq m$. Now suppose that $t \in [0, 1]$ is an a-fast point. By definition there exists an $0 < u < m\eta^{-m}$ such that

$$|B(t+u) - B(t)| \geq a(1-\varepsilon)\sqrt{2u\log(1/u)}.$$

We pick the unique $k \geq m$ such that $k\eta^{-k} < u \leq (k-1)\eta^{-k+1}$, and let $h' = k\eta^{-k}$. By (1.1), we have

$$\begin{aligned}|B(t+h') - B(t)| &\geq |B(t+u) - B(t)| - |B(t+u) - B(t+h')| \\ &\geq a(1-\varepsilon)\sqrt{2u\log(1/u)} - C\sqrt{(u-h')\log(1/(u-h'))}.\end{aligned}$$

As $0 \leq u - h' \leq (\eta - 1)k\eta^{-k}$, and by our choice of η, the subtracted term is smaller than $a\varepsilon\sqrt{2h'\log(1/h')}$ for sufficiently large m. Hence there exists $k \geq m$ such that

$$|B(t+h') - B(t)| \geq a(1-2\varepsilon)\sqrt{2h'\log(1/h')}.$$

Now let j be such that $t \in [j\eta^{-k}, (j+1)\eta^{-k})$. As before let $h = \eta^{-k}$. Then, by the triangle inequality and using (1.1) twice, we have

$$\begin{aligned}|B(jh+h') - B(jh)| &\geq |B(t+h') - B(t)| - |B(t) - B(jh)| - |B(jh+h') - B(t+h')| \\ &\geq a(1-2\varepsilon)\sqrt{2h'\log(1/h')} - 2C\sqrt{h\log(1/h)} \\ &> a(1-4\varepsilon)\sqrt{2h'\log(1/h')},\end{aligned}$$

using in the last step that, by choosing m sufficiently large, the subtracted term can be made smaller than $2a\varepsilon\sqrt{2h'\log(1/h')}$. ∎

Proof of the upper bound in Theorem 10.3. This involves only a first moment calculation. All there is to show is that, for any $\gamma > 1 - a^2$ there exists $\varepsilon > 0$ such that the random variable $\sum_{I \in \mathfrak{I}(\varepsilon,\delta)} |I|^\gamma$ is finite, almost surely. For this it suffices to verify that its expectation is finite. Note that

$$\mathbb{E}\Big[\sum_{I \in \mathfrak{I}(\varepsilon,\delta)} |I|^\gamma\Big] = \sum_{k=m}^{\infty} \sum_{j=0}^{\lceil \eta^k - 1\rceil} \eta^{-k\gamma}\,\mathbb{P}\Big\{\frac{|B(j\eta^{-k} + k\eta^{-k}) - B(j\eta^{-k})|}{\sqrt{2\,k\eta^{-k}\log(\eta^k/k)}} > a(1-4\varepsilon)\Big\}.$$

So it all boils down to an estimate of a single probability, which is very simple as it involves just one normal random variable, namely $B(j\eta^{-k} + k\eta^{-k}) - B(j\eta^{-k})$. More precisely, for X

standard normal,

$$
(1.2)\qquad
\begin{aligned}
&\mathbb{P}\Big\{\frac{|B(j\eta^{-k}+k\eta^{-k})-B(j\eta^{-k})|}{\sqrt{2k\eta^{-k}\log(\eta^k/k)}}>a(1-4\varepsilon)\Big\}=\mathbb{P}\big\{|X|>a(1-4\varepsilon)\sqrt{2\log(\eta^k/k)}\big\}\\
&\quad\leq \tfrac{1}{a(1-4\varepsilon)\sqrt{\log(\eta^k/k)\pi}}\exp\big\{-a^2(1-4\varepsilon)^2\log(\eta^k/k)\big\}\leq \eta^{-ka^2(1-4\varepsilon)^3},
\end{aligned}
$$

for all sufficiently large k and all $0\leq j<2^k$, using the estimate for normal random variables of Lemma II.3.1 in the penultimate step. Given $\gamma>1-a^2$ we can finally find $\varepsilon>0$ such that $\gamma+a^2(1-4\varepsilon)^3>1$, so that

$$
\sum_{k=m}^{\infty}\sum_{j=0}^{\eta^k-1}\eta^{-k\gamma}\,\mathbb{P}\Big\{\frac{|B(j\eta^{-k}+k\eta^{-k})-B(j\eta^{-k})|}{\sqrt{2k\eta^{-k}\log(\eta^k/k)}}>a(1-4\varepsilon)\Big\}\leq\sum_{k=1}^{m}\eta^k\eta^{-k\gamma}\,\eta^{-ka^2(1-4\varepsilon)^3}<\infty,
$$

completing the proof of the upper bound in Theorem 10.3. ■

Remark 10.5. If $a>1$ one can choose $\gamma<0$ in the previous proof, which shows that there are no a-fast times as the empty collection is suitable to cover the set of a-fast times. ◇

For the *lower bound* we have to work harder. We divide, for any positive integer k, the interval $[0,1]$ into nonoverlapping dyadic subintervals $[j2^{-k},(j+1)2^{-k}]$ for $j=0,\dots,2^k-1$. As before, we denote this collection of intervals by $\mathfrak{C}_k$ and by $\mathfrak{C}$ the union over all collections $\mathfrak{C}_k$ for $k\geq 1$. To each interval $I\in\mathfrak{C}$ we associate a $\{0,1\}$-valued random variable $Z(I)$ and then define sets

$$
A(k):=\bigcup_{\substack{I\in\mathfrak{C}_n\\ Z(I)=1}} I\qquad\text{and}\qquad A:=\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}A(k)\,.
$$

Because $\mathbb{1}_A=\limsup\mathbb{1}_{A(k)}$ the set A is often called the **limsup fractal** associated with the family $(Z(I)\colon I\in\mathfrak{C})$. We shall see below that the set of a-fast points contains a large limsup fractal and derive the lower bound from the following general result on limsup fractals.

Proposition 10.6. *Suppose that $(Z(I)\colon I\in\mathfrak{C})$ is a collection of random variables with values in $\{0,1\}$ such that $p_k:=\mathbb{P}\{Z(I)=1\}$ is the same for all $I\in\mathfrak{C}_k$. For $I\in\mathfrak{C}_m$, with $m<n$, define*

$$
M_n(I):=\sum_{\substack{J\in\mathfrak{C}_n\\ J\subset I}} Z(J)\,.
$$

Let $\zeta(n)\geq 1$ and $\gamma>0$ be such that

(1) $\mathrm{Var}(M_n(I))\leq\zeta(n)\,\mathbb{E}[M_n(I)]=\zeta(n)\,p_n2^{n-m}$,

(2) $\lim\limits_{n\uparrow\infty}2^{n(\gamma-1)}\,\zeta(n)\,p_n^{-1}=0$,

.

then $\dim A\geq\gamma$ almost surely for the limsup fractal A associated with $(Z(I)\colon I\in\mathfrak{C})$.

The *idea of the proof* of Proposition 10.6 is to construct a probability measure μ on A and then use the energy method. To this end, we choose an increasing sequence $(\ell_k \colon k \in \mathbb{N})$ such that $M_{\ell_k}(D) > 0$ for all $D \in \mathfrak{C}_{\ell_{k-1}}$. We then define a (random) probability measure μ in the following manner: Assign mass one to any interval $I \in \mathfrak{C}_{\ell_0}$. Proceed inductively: if $J \in \mathfrak{C}_m$ with $\ell_{k-1} < m \le \ell_k$ and $J \subset D$ for $D \in \mathfrak{C}_{\ell_{k-1}}$ define

$$\mu(J) = \frac{M_{\ell_k}(J)\mu(D)}{M_{\ell_k}(D)}. \tag{1.3}$$

Then μ is consistently defined on all intervals in $\mathfrak{C}$ and therefore can be extended to a probability measure on $[0,1]$ by the measure extension theorem. The crucial part of the proof is then to show that, for a suitable choice of $(\ell_k \colon k \in \mathbb{N})$ the measure μ has finite γ-energy.

For the proof of Proposition 10.6 we need two lemmas. The first one is a simple combination of two facts, which have been established at other places in the book: The bounds for the energy of a measure established in Lemma 9.20, and the lower bound of Hausdorff dimension in terms of capacity which follows from the potential theoretic method, see Theorem 4.27.

Lemma 10.7. *Suppose $B \subset [0,1]$ is a Borel set and μ is a probability measure on B. Then*

$$\sum_{m=1}^{\infty} \sum_{J \in \mathfrak{C}_m} \frac{\mu(J)^2}{2^{-\alpha m}} < \infty \quad \textit{implies} \quad \dim B \ge \alpha.$$

Proof. By Lemma 9.20 with $f(x) = x^\alpha$ and $h(n) = 2^{-n\alpha}(1 - 2^{-\alpha})$ we obtain, for a suitable constant $C > 0$ that

$$I_\alpha(\mu) \le C \sum_{m=1}^{\infty} \sum_{J \in \mathfrak{C}_m} \frac{\mu(J)^2}{2^{-\alpha m}}.$$

If the right hand side is finite, then so is the α-energy of the measure μ. We thus obtain $\dim B \ge \alpha$ by Theorem 4.27. ∎

For the formulation of the second lemma we use (2) to pick, for any $\ell \in \mathbb{N}$ an integer $n = n(\ell) \ge \ell$ such that

$$2^{n(\gamma-1)}\, \zeta(n) \le p_n\, 2^{-3\ell}.$$

Lemma 10.8. *There exists an almost surely finite random variable ℓ_0 such that, for all $\ell \ge \ell_0$ and $D \in \mathfrak{C}_\ell$, with $n = n(\ell)$,*

- *for all $D \in \mathfrak{C}_\ell$ we have*

  $$\big| M_n(D) - \mathbb{E} M_n(D) \big| < \tfrac{1}{2} \mathbb{E} M_n(D),$$

  and, in particular, $M_n(D) > 0$;

- *for a constant C depending only on γ,*

  $$\sum_{m=\ell}^{n} 2^{\gamma m} \sum_{\substack{J \in \mathfrak{C}_m \\ J \subset D}} \frac{M_n(J)^2}{(2^{n-\ell} p_n)^2} \le C 2^{\gamma \ell}.$$

REMARK 10.9. The first statement in the lemma says intuitively that the variance of the random variables $M_n(D)$ is small, i.e. they are always close to their mean. This is essentially what makes this proof work. $\diamond$

Proof of Lemma 10.8. For $m \le n$, $J \in \mathfrak{C}_m$ we denote $\Delta_n(J) := M_n(J) - \mathbb{E}M_n(J)$ and, for $\ell \le n$ and $D \in \mathfrak{C}_\ell$, set

$$\Upsilon_n(D) := \sum_{m=\ell}^{n} 2^{m\gamma} \sum_{\substack{J\in\mathfrak{C}_m \\ J\subset D}} \Delta_n(J)^2.$$

By assumption (1) we have $\mathbb{E}\big[\Delta_n(J)^2\big] \le \zeta(n)p_n 2^{n-m}$ and therefore, for all $D \in \mathfrak{C}_\ell$,

$$\mathbb{E}\Upsilon_n(D) \le \sum_{m=\ell}^{n} 2^{m\gamma} \sum_{\substack{J\in\mathfrak{C}_m \\ J\subset D}} \mathbb{E}[\Delta_n(J)^2] \le \sum_{m=\ell}^{n} 2^{m\gamma}\zeta(n)\, p_n\, 2^{n-\ell} \le \frac{2^{n\gamma}}{2^\gamma - 1}\,\zeta(n)\, p_n\, 2^{n-\ell}.$$

By our choice of $n = n(\ell)$ we thus obtain

$$\mathbb{E}\Big[\sum_{D\in\mathfrak{C}_\ell} \frac{\Upsilon_n(D)}{(2^{n-\ell}p_n)^2}\Big] \le \frac{1}{2^\gamma - 1}\,\zeta(n)\, 2^{2\ell-n+n\gamma} p_n^{-1} \le \frac{2^{-\ell}}{2^\gamma - 1}\,.$$

Since the right hand side is summable in ℓ we conclude that the summands inside the last expectation converge to zero as $\ell \uparrow \infty$. In particular, there exists $\ell_0 < \infty$ such that, for all $\ell \ge \ell_0$ we have $2^{-\ell\gamma} \le 1/4$ and, for $n = n(\ell)$ and $D \in \mathfrak{C}_\ell$,

$$\Upsilon_n(D) \le \big(2^{n-\ell}p_n\big)^2 = \big(\mathbb{E}M_n(D)\big)^2.$$

The first statement follows from this very easily: For any $\ell \ge \ell_0$ and $n = n(\ell)$ we have (recalling the definition of $\Upsilon_n(D)$),

$$\Delta_n(D)^2 \le 2^{-\ell\gamma}\Upsilon_n(D) \le 2^{-\ell\gamma}\big(\mathbb{E}M_n(D)\big)^2 \le \tfrac{1}{4}\big(\mathbb{E}M_n(D)\big)^2.$$

In order to get the second statement we calculate,

$$\sum_{\substack{J\in\mathfrak{C}_m \\ J\subset D}} \frac{(\mathbb{E}M_n(J))^2}{(2^{n-\ell}p_n)^2} = \sum_{\substack{J\in\mathfrak{C}_m \\ J\subset D}} 2^{2(\ell-m)} = 2^{\ell-m}.$$

Therefore

$$\sum_{m=\ell}^{n} 2^{m\gamma} \sum_{\substack{J\in\mathfrak{C}_m \\ J\subset D}} \frac{(\mathbb{E}M_n(J))^2}{(2^{n-\ell}p_n)^2} = 2^{\ell} \sum_{m=\ell}^{n} 2^{-m(1-\gamma)} \le \frac{2^{\ell\gamma}}{1-2^{-(1+\gamma)}}\,. \tag{1.4}$$

Now, recalling the choice of n,

$$\sum_{m=\ell}^{n} 2^{m\gamma} \sum_{\substack{J\in\mathfrak{C}_m \\ J\subset D}} \frac{\Delta_n(J)^2}{(2^{n-\ell}p_n)^2} = \frac{1}{(2^{n-\ell}p_n)^2} \sum_{m=\ell}^{n} 2^{m\gamma} \sum_{\substack{J\in\mathfrak{C}_m \\ J\subset D}} \Delta_n(J)^2 = \frac{\Upsilon_n(D)}{(2^{n-\ell}p_n)^2} \le 1. \tag{1.5}$$

Since $M_n(J)^2 = \big(\mathbb{E}M_n(J) + \Delta_n(J)\big)^2 \le 2\big(\mathbb{E}M_n(J)\big)^2 + 2\big(\Delta_n(J)\big)^2$, adding the inequalities (1.4) and (1.5) and setting $C := 2 + 2/(1-2^{-(1+\gamma)})$ proves the second statement. $\blacksquare$

We now define the sequence $(\ell_k\colon k \in \mathbb{N})$ by $\ell_{k+1} = n(\ell_k)$ for all integers $k \geq 0$. The first statement of Lemma 10.8 ensures that μ is well defined by (1.3), and together with the second statement will enable us to check that μ has finite γ-energy.

Proof of Proposition 10.6. We can now use Lemma 10.8 to verify the condition of Lemma 10.7 and finish the proof of Proposition 10.6. Indeed, by definition of μ,

$$\sum_{m=\ell_0+1}^{\infty} \sum_{J\in\mathfrak{C}_m} \frac{\mu(J)^2}{2^{-\gamma m}} = \sum_{k=0}^{\infty} \sum_{m=\ell_k+1}^{\ell_{k+1}} 2^{\gamma m} \sum_{D\in\mathfrak{C}_{\ell_k}} \frac{\mu(D)^2}{M_{\ell_{k+1}}(D)^2} \sum_{\substack{J\in\mathfrak{C}_m\\ J\subset D}} M_{\ell_{k+1}}(J)^2. \tag{1.6}$$

Recall that $q_{k+1} := \mathbb{E} M_{\ell_{k+1}}(D) = 2^{\ell_{k+1}-\ell_k} p_{\ell_{k+1}}$ and, by the first statement of Lemma 10.8, for every $k \in \mathbb{N}$ and $D \in \mathfrak{C}_{\ell_k}$,

$$\tfrac{1}{2}\, q_{k+1} \leq M_{\ell_{k+1}}(D) \leq 2q_{k+1}.$$

Now, from the definition of the measure μ we get, with $D \subset D' \in \mathfrak{C}_{k-1}$,

$$\mu(D) = \frac{M_{\ell_k}(D)\mu(D')}{M_{\ell_k}(D')} \leq 2Z(D)/q_k,$$

and therefore we can continue (1.6) with the upper bound

$$16 \sum_{k=0}^{\infty} \frac{1}{q_k^2} \sum_{D\in\mathfrak{C}_{\ell_k}} Z(D) \sum_{m=\ell_k+1}^{\ell_{k+1}} 2^{\gamma m} \sum_{\substack{J\in\mathfrak{C}_m\\ J\subset D}} \frac{M_{\ell_{k+1}}(J)^2}{q_{k+1}^2} \leq 16\, C \sum_{k=0}^{\infty} \frac{1}{q_k^2} \sum_{D\in\mathfrak{C}_{\ell_k}} Z(D)\, 2^{\gamma \ell_k},$$

using the second statement of Lemma 10.8 and the definition of q_{k+1}. Finally, using the definition of $M_{\ell_{k+1}}$ and the definition of $(\ell_k : k \in \mathbb{N})$ we note that,

$$\sum_{k=1}^{\infty} \frac{1}{q_k^2} M_{\ell_k}([0,1])\, 2^{\gamma\ell_k} \leq \sum_{k=1}^{\infty} \frac{2^{\ell_{k-1}}}{q_k}\, 2^{\gamma\ell_k} \leq \sum_{k=1}^{\infty} 2^{2\ell_{k-1}-\ell_k} \frac{2^{\gamma\ell_k}}{p_{\ell_k}} \leq \sum_{k=1}^{\infty} 2^{-\ell_{k-1}} < \infty.$$

This ensures convergence of the sequence (1.6) and thus completes the proof. ∎

Coming back to the lower bound in Theorem 10.3 we fix $\varepsilon > 0$. Given $I = [jh, (j+1)h]$ with $h := 2^{-k}$ we let $Z(I) = 1$ if and only if

$$|B(jh + h') - B(jh)| \geq a(1+\varepsilon)\sqrt{2h' \log(1/h')}, \qquad \text{for } h' := k2^{-k}.$$

LEMMA 10.10. *Almost surely, the set A associated with this family $(Z(I)\colon I \in \mathfrak{C})$ of random variables is contained in the set of a-fast times.*

Proof. Recall that by Theorem 1.12 there exists a constant $C > 0$ such that, almost surely,

$$|B(s) - B(t)| \leq C\sqrt{|t-s| \log(1/|t-s|)}, \qquad \text{for all } s,t \in [0,2].$$

Now assume that k is large enough that $\big(\frac{2C}{a\varepsilon}\big)^2 \, k \log 2 + k \log k \leq \; k^2 \log 2$. Let $t \in A$ and suppose that $t \in I \in \mathfrak{C}_k$ with $Z(I) = 1$. Then, by the triangle equality,

$$\begin{aligned} |B(t+h') - B(t)| &\geq |B(jh+h') - B(jh)| - |B(t+h') - B(jh+h')| - |B(jh) - B(t)| \\ &\geq a(1+\varepsilon) \sqrt{h' \log(1/h')} - 2C \sqrt{h \log h} \\ &\geq \; a \sqrt{h' \log(1/h')}. \end{aligned}$$

As this happens for infinitely many k, this proves that t is an a-fast point. ■

The next lemma singles out the crucial estimates of expectation and variance needed to apply Proposition 10.6. The first is based on the upper tail estimate for a standard normal distribution, the second on the 'short range' of the dependence of the family $(Z(I) \,:\, I \in \mathfrak{C})$.

LEMMA 10.11. *Define* $p_n = \mathbb{E}[Z(I)]$ *for* $I \in \mathfrak{C}_n$, *and* $\eta(n) := 2n+1$. *Then,*

(a) *for* $I \in \mathfrak{C}_k$ *we have* $\mathbb{E}[Z(I)] \geq 2^{-k\, a^2\, (1+\varepsilon)^3}$;

(b) *for* $m \leq n$ *and* $J \in \mathfrak{C}_m$, *we have* $\mathbb{E}[(M_n(J) - \mathbb{E}M_n(J))^2] \leq p_n \, 2^{n-m} \, \eta(n)$.

Proof. For part (a), denoting by X a standard normal random variable,

$$\begin{aligned} &\mathbb{P}\big\{|B(jh+h') - B(jh)| \geq a(1+\varepsilon)\sqrt{2h' \log(1/h')}\big\} = \mathbb{P}\big\{|X| > a(1+\varepsilon)\sqrt{2\log(1/h')}\big\} \\ &\quad\geq \tfrac{a(1+\varepsilon)\sqrt{2\log(1/h')}}{1+a^2(1+\varepsilon)^2 \log(1/h')} \tfrac{1}{\sqrt{2\pi}} \exp\big\{ - a^2\,(1+\varepsilon)^2\, \log(1/h')\big\} \geq 2^{-k\,a^2\,(1+\varepsilon)^3}, \end{aligned} \tag{1.7}$$

for all sufficiently large k and all $0 \leq j < 2^k$, using the lower estimate for normal random variables of Lemma 3.1 in the penultimate step.

For part (b) note that for two intervals $J_1, J_2 \in \mathfrak{C}_n$ the associated random variables $Z(J_1)$ and $Z(J_2)$ are independent if their distance is at least than $n2^{-n}$. Using this whenever possible and the trivial estimate $\mathbb{E}[Z(J_1)Z(J_2)] \leq \mathbb{E}Z(J_1)$ otherwise, we get

$$\mathbb{E}M_n(J)^2 = \sum_{\substack{J_1,J_2 \in \mathfrak{C}_n \\ J_1,J_2 \subset I}} \mathbb{E}\big[Z(J_1)Z(J_2)\big] \leq \sum_{\substack{J_1 \in \mathfrak{C}_n \\ J_1 \subset I}} (2n+1)\mathbb{E}Z(J_1) + \mathbb{E}Z(J_1) \sum_{\substack{J_2 \in \mathfrak{C}_n \\ J_2 \subset I}} \mathbb{E}Z(J_2),$$

Hence we obtain

$$\mathbb{E}\big[(M_n(J) - \mathbb{E}M_n(J))^2\big] \leq \sum_{\substack{J_1 \in \mathfrak{C}_n \\ J_1 \subset I}} (2n+1)p_n = p_n 2^{n-m}(2n+1),$$

which settles the lemma. ■

Proof of the lower bound in Theorem 10.3. By Lemma 10.11 the conditions of Proposition 10.6 hold for any $\gamma < 1 - a^2\,(1+\varepsilon)^3$. As, for any $\varepsilon > 0$, the set A associated to $(Z(I)\colon I \in \mathfrak{C})$ is contained in the set of a-fast times, the latter has at least dimension $1 - a^2$. ■

2. Packing dimension and limsup fractals

In this section we ask for a precise criterion, whether a set E contains a-fast times for various values of a. It turns out that such a criterion depends not on the Hausdorff, but on the packing dimension of the set E. We therefore begin by introducing the concept of *packing dimension*, which was briefly mentioned in the beginning of Chapter 4, in some detail.

We choose to define packing dimension in a way, which indicates its nature as a concept, which is a natural *dual* to the notion of Hausdorff dimension. The natural *dual* operation to covering a set with balls, as we have done in the case of Hausdorff dimension, is the operation of *packing* balls into the set.

DEFINITION 10.12. *Suppose E is a metric space. For every $\delta > 0$, a* **δ-packing** *of $A \subset E$ is a countable collection of* disjoint *balls*

$$\mathcal{B}(x_1, r_1), \mathcal{B}(x_2, r_2), \mathcal{B}(x_3, r_3), \ldots$$

with centres $x_i \in A$ and radii $0 \le r_i \le \delta$. For every $s \ge 0$ we introduce the **s-value** *of the packing as $\sum_{i=1}^{\infty} r_i^s$. The* **s-packing number** *of A is defined as*

$$P^s(A) = \lim_{\delta \downarrow 0} P^s_\delta \quad \text{for } P^s_\delta = \sup\Big\{ \sum_{i=1}^{\infty} r_i^s \,:\, (\mathcal{B}(x_i, r_i)) \text{ a } \delta\text{-packing of } A \Big\}.$$ ◇

Note that the packing number is defined in the same way as the Hausdorff measure with efficient (small) coverings replaced by efficient (large) packings. A difference is that the packing numbers do *not* define a reasonable measure. However a small modification gives the so-called packing measure,

$$\mathcal{P}^s(A) = \inf\Big\{ \sum_{i=1}^{\infty} P^s(A_i) \,:\, A = \bigcup_{i=1}^{\infty} A_i \Big\}.$$

The packing dimension has a definition analogous to the definition of Hausdorff dimension with Hausdorff measures replaced by packing measures.

DEFINITION 10.13. *The* **packing dimension** *of E is $\dim_P E = \inf\{s : \mathcal{P}^s(E) = 0\}$.* ◇

REMARK 10.14. It is not hard to see that

$$\dim_P E = \inf\{s : \mathcal{P}^s(E) < \infty\} = \sup\{s : \mathcal{P}^s(E) > 0\} = \sup\{s : \mathcal{P}^s(E) = \infty\},$$

a proof of this fact is suggested as Exercise 10.1. ◇

An alternative approach to packing dimension is to use a suitable *regularisation* of the upper Minkowski dimension, recall Remark 4.4.4 where we have hinted at this possibility.

THEOREM 10.15. *For every metric space E we have*

$$\dim_P E = \inf\Big\{ \sup_{i=1}^{\infty} \overline{\dim}_M E_i \,:\, E = \bigcup_{i=1}^{\infty} E_i \,, E_i \text{ bounded} \Big\}.$$

Remark 10.16. We have, for all bounded sets E, that $\dim_P E \le \overline{\dim}_M E$ and, of course, strict inequality may hold. Obviously, every countable set has packing dimension 0, compare with the example in Exercise 4.2. For this definition it is not hard to see that the countable stability property is satisfied. $\diamond$

Proof. Define, for every $A \subset E$ and $\varepsilon > 0$,

$$P(A,\varepsilon) = \max\big\{k : \text{ there are disjoint balls } \mathcal{B}(x_1,\varepsilon),\dots,\mathcal{B}(x_k,\varepsilon) \text{ with } x_i \in A\big\}.$$

Recall the definition of the numbers $M(A,\varepsilon)$ from (1.1) in Chapter 4. We first show that

$$P(A,4\varepsilon) \le M(A,2\varepsilon) \le P(A,\varepsilon)\,.$$

Indeed, if $k = P(A,\varepsilon)$ let $\mathcal{B}(x_1,\varepsilon),\dots,\mathcal{B}(x_k,\varepsilon)$ be disjoint balls with $x_i \in A$. Suppose $x \in A \setminus \bigcup_{i=1}^k \mathcal{B}(x_i,2\varepsilon)$, then $\mathcal{B}(x,\varepsilon)$ is disjoint from all balls $\mathcal{B}(x_i,\varepsilon)$ contradicting the choice of k. Hence $\mathcal{B}(x_1,2\varepsilon),\dots,\mathcal{B}(x_k,2\varepsilon)$ is a covering of A and we have shown $M(A,2\varepsilon) \le P(A,\varepsilon)$.

For the other inequality let $m = M(A,2\varepsilon)$ and $k = P(A,4\varepsilon)$ and choose $x_1,\dots,x_m \in A$ and $y_1,\dots,y_k \in A$ such that

$$A \subset \bigcup_{i=1}^m \mathcal{B}(x_i,2\varepsilon) \text{ and } \mathcal{B}(y_1,4\varepsilon),\dots,\mathcal{B}(y_k,4\varepsilon) \text{ disjoint.}$$

Then each y_j belongs to some $\mathcal{B}(x_i,2\varepsilon)$ and no such ball contains more than one such point. Thus $k \le m$, which proves $P(A,4\varepsilon) \le M(A,2\varepsilon)$.

Suppose now that $\inf\{t : \mathcal{P}^t(E) = 0\} < s$. Then there is $t < s$ and $E = \bigcup_{i=1}^\infty A_i$ such that, for every set $A = A_i$, we have $P^t(A) < 1$. Obviously, $P^t_\varepsilon(A) \ge P(A,\varepsilon)\varepsilon^t$. Letting $\varepsilon \downarrow 0$ gives

$$\lim_{\varepsilon\downarrow 0} M(A,\varepsilon)\varepsilon^t \le \lim_{\varepsilon\downarrow 0} P(A,\varepsilon/2)\varepsilon^t \le 2^t P^t(A) < 2^t.$$

Hence $\overline{\dim}_M A \le t$ and by definition $\sup_{i=1}^\infty \overline{\dim}_M A_i \le t < s$.

To prove the opposite inequality, let

$$0 < t < s < \inf\{r : \mathcal{P}^r(E) = 0\}\,,$$

and $A_i \subset E$ bounded with $E = \bigcup_{i=1}^\infty A_i$. It suffices to show that $\overline{\dim}_M(A_i) \ge t$ for some i. Since $\mathcal{P}^s(E) > 0$ there is i such that $P^s(A_i) > 0$. Let $0 < \alpha < P^s(A_i)$, then for all $\delta \in (0,1)$ we have $P^s_\delta(A_i) > \alpha$ and there exist disjoint balls $\mathcal{B}(x_1,r_1),\mathcal{B}(x_2,r_2),\mathcal{B}(x_3,r_3),\dots$ with centres $x_j \in A_i$ and radii r_j smaller than δ with

$$\sum_{j=1}^\infty r_j^s \ge \alpha\,.$$

For every m let k_m be the number of balls with radius $2^{-m-1} < r_j \le 2^{-m}$. Then,

$$\sum_{m=0}^\infty k_m 2^{-ms} \ge \sum_{j=1}^\infty r_j^s \ge \alpha\,.$$

This yields, for some integer $N \geq 0$, $2^{Nt}(1-2^{t-s})\alpha \leq k_N$, since otherwise

$$\sum_{m=0}^{\infty} k_m 2^{-ms} < \sum_{m=0}^{\infty} 2^{mt}(1-2^{t-s})2^{-ms}\alpha = \alpha\,.$$

Since $r_j \leq \delta$ for all j, we have $2^{-N-1} < \delta$. Moreover,

$$P(A_i, 2^{-N-1}) \geq k_N \geq 2^{Nt}(1-2^{t-s})\alpha\,,$$

which gives

$$\sup_{0\leq\varepsilon\leq\delta} P(A_i,\varepsilon)\varepsilon^t \geq P(A_i, 2^{-N-1})2^{-Nt-t} \geq 2^{-t}(1-2^{t-s})\alpha\,.$$

Letting $\delta \downarrow 0$, and recalling the relation of $M(A,\varepsilon)$ and $P(A,\varepsilon)$ established at the beginning of the proof, we obtain

$$\limsup_{\varepsilon\downarrow 0} M(A_i,\varepsilon)\varepsilon^t \geq \limsup_{\varepsilon\downarrow 0} P(A_i, 2\varepsilon)\varepsilon^t > 0\,,$$

and thus $\overline{\dim}_M A_i \geq t$, as required. ■

Remark 10.17. It is easy to see that, for every metric space, $\dim_P E \geq \dim E$. This is suggested as Exercise 10.2. ◇

The following result shows that every closed subset of $\mathbb{R}^d$ has a large subset, which is 'regular' in a suitable sense. It will be used in the proof of Theorem 10.28 below.

Lemma 10.18. *Let $A \subset \mathbb{R}^d$ be closed.*

(i) *If any open set V which intersects A satisfies* $\overline{\dim}_M(A\cap V) \geq \alpha$ *, then* $\dim_P(A) \geq \alpha$.
(ii) *If $\dim_P(A) > \alpha$, then there is a (relatively closed) nonempty subset $\widetilde{A}$ of A, such that, for any open set V which intersects $\widetilde{A}$, we have* $\dim_P(\widetilde{A}\cap V) > \alpha$.

Proof. Let $A \subset \bigcup_{j=1}^{\infty} A_j$, where the A_j are closed. We are going to show that there exist an open set V and an index j such that $V \cap A \subset A_j$. For this V and j we have,

$$\overline{\dim}_M(A_j) \geq \overline{\dim}_M(A_j\cap V) \geq \overline{\dim}_M(A\cap V) \geq \alpha.$$

This in turn implies that $\dim_P(A) \geq \alpha$.

Suppose now that for any V open such that $V\cap A \neq \emptyset$, it holds that $V \cap A \not\subset A_j$. Then A_j^c is a dense open set relative to A. By Baire's (category) theorem $A \cap \bigcap_j A_j^c \neq \emptyset$, which means that $A \not\subset \bigcup_j A_j$, contradicting our assumption and proving (i).

Now choose a countable basis $\mathcal{B}$ of the topology of $\mathbb{R}^d$ and define

$$\widetilde{A} = A \setminus \bigcup \big\{B \in \mathcal{B} \,:\, \overline{\dim}_M(B\cap A) \leq \alpha\big\}.$$

Then, $\dim_P(A\setminus\widetilde{A}) \leq \alpha$ using $\dim_P \leq \overline{\dim}_M$ and countable stability of packing dimension. From this we conclude that

$$\dim_P \widetilde{A} = \dim_P A > \alpha.$$

If for some V open, $V \cap \widetilde{A} \neq \emptyset$ and $\dim_P(\widetilde{A} \cap V) \leq \alpha$ then V contains some set $B \in \mathcal{B}$ such that $\widetilde{A} \cap B \neq \emptyset$. For that set we have $\dim_P(A \cap B) \leq \dim_P(A \setminus \widetilde{A}) \wedge \dim_P(\widetilde{A} \cap B)) \leq \alpha$, contradicting the construction of $\widetilde{A}$. ■

EXAMPLE 10.19. As an example of a result demonstrating the duality between Hausdorff and packing dimension is the *product formula*, see [**BP96**]. In the dimension theory of smooth sets (manifolds, linear spaces) we have the following formula for product sets

$$\dim(E \times F) = \dim E + \dim F\,.$$

The example discussed in Exercise 2.3 shows that this formula fails for Hausdorff dimension, a reasonable formula for the Hausdorff dimension of product sets necessarily involves information about the packing dimension of one of the factor sets. In [**BP96**] it is shown that, for every Borel set $A \subset \mathbb{R}^d$,

$$\dim_P(A) = \sup_B \Big\{ \dim(A \times B) - \dim(B) \Big\}$$

where the supremum is over all compact sets $B \subset \mathbb{R}^d$. One can also show that, if A satisfies $\dim A = \dim_P A$, then the product formula $\dim(A \times B) = \dim A + \dim B$ holds. ◇

Before moving back to our study of Brownian paths we study the packing dimension of the 'test sets' we have used in the stochastic codimension method, see Section 9.1.

THEOREM 10.20. *Let $\gamma \in [0, d]$ and $\Gamma[\gamma]$ be a percolation limit set in $\mathbb{R}^d$ with retention parameter $2^{-\gamma}$. Then*

- $\dim_P \Gamma[\gamma] \leq d - \gamma$ *almost surely,*
- $\dim_P \Gamma[\gamma] = d - \gamma$ *almost surely on* $\Gamma[\gamma] \neq \emptyset$.

Proof. For the first item, as packing dimension is bounded from above by the upper Minkowski dimension, it suffices to show that $\underline{\dim}_{\mathrm{M}} \Gamma[\gamma] \leq d - \gamma$ almost surely. For this purpose we use the formula for the upper Minkowski dimension given in Remark 4.2. For a given n, we cover the percolation limit set by $\mathfrak{S}_n$, the collection of cubes retained in the nth construction step. The probability that a given cube of sidelength 2^{-n} is in $\mathfrak{S}_n$ is $2^{-n\gamma}$ and hence the expected number of cubes in $\mathfrak{S}_n$ is $2^{n(d-\gamma)}$. Hence, for any $\varepsilon > 0$,

$$\mathbb{P}\Big\{2^{n(\gamma-d-\varepsilon)}\, \#\mathfrak{S}_n > 1\Big\} \leq 2^{n(\gamma-d-\varepsilon)}\mathbb{E}\#\mathfrak{S}_n \leq 2^{-n\varepsilon},$$

which is summable. Hence, almost surely, $2^{n(\gamma-d-\varepsilon)}\, \#\mathfrak{S}_n \leq 1$ for all but finitely many n. Thus, almost surely,

$$\overline{\dim}_{\mathrm{M}} \leq \limsup_{n\uparrow\infty} \frac{\log \#\mathfrak{S}_n}{n\, \log 2} \leq d - \gamma + \varepsilon \qquad \text{for every } \varepsilon > 0.$$

For the second item recall the corresponding statement for Hausdorff dimension from Exercise 9.3. The result follows, as packing dimension is bounded from below by the Hausdorff dimension, see Remark 10.17. ■

REMARK 10.21. At a first glance the concept of packing dimension does not seem to add substantial news to the discussion of fine properties of d-dimensional Brownian motion. However, a first sign that something interesting might be going on can be found in Exercise 10.5, where we show that the Hausdorff and packing dimension of the sets of a-fast points differ. This is indicative of the fact that optimal coverings of these sets uses coverings sets of widely differing size, and that optimal packings use sets of quite different scale. $\diamond$

Given a set $E \subset [0,1]$ we now ask for the maximal value of a such that E contains an a-fast time with positive probability. This notion of size is most intimately linked to packing dimension as the following theorem shows. We denote by $F(a) \subset [0,1]$ the set of a-fast times.

THEOREM 10.22 (Khoshnevisan, Peres and Xiao). *For any compact set $E \subset [0,1]$, almost surely,*

$$\sup_{t\in E} \limsup_{h\downarrow 0} \frac{|B(t+h)-B(t)|}{\sqrt{2h\log(1/h)}} = \sqrt{\dim_P(E)}.$$

Moreover, if $\dim_P(E) > a^2$, then $\dim_P(F(a)\cap E) = \dim_P(E)$.

REMARK 10.23. The result can be extended from compact sets E to more general classes of sets, more precisely the *analytic* sets, see [**KPX00**]. $\diamond$

REMARK 10.24. An equivalent formulation of the theorem is that, for any compact $E \subset [0,1]$, almost surely,

$$\mathbb{P}\big\{F(a)\cap E \neq \emptyset\big\} = \begin{cases} 1, & \text{if } \dim_P(E) > a^2, \\ 0, & \text{if } \dim_P(E) < a^2. \end{cases}$$

Using the compact percolation limit sets $E = \Gamma[\gamma]$ in this result and Hawkes' theorem, Theorem 9.5, one can obtain an alternative proof of the Orey-Taylor theorem. Indeed, by Theorem 10.20, if $\gamma < 1-a^2$ we have $\dim_P(E) > a^2$ with positive probability, and therefore, $\mathbb{P}\big\{F(a)\cap E \neq \emptyset\big\} > 0$. Hence, by Hawkes' theorem, $\dim\, F(a) \geq 1-a^2$ with positive probability. Brownian scaling maps a-fast points onto a-fast points. Therefore there exists $\varepsilon > 0$ such that, for any $n \in \mathbb{N}$ and $0 \leq j \leq n-1$,

$$\mathbb{P}\big\{\dim(F(a)\cap[j/n,(j+1)/n]) \geq 1-a^2\big\} \geq \varepsilon,$$

and hence

$$\mathbb{P}\big\{\dim F(a) \geq 1-a^2\big\} \geq 1-(1-\varepsilon)^n \to 1.$$

Conversely, by Theorem 10.20, if $\gamma > 1-a^2$ we have $\dim_P(E) < a^2$ almost surely, and therefore, $\mathbb{P}\big\{F(a)\cap E \neq \emptyset\big\} = 0$. Hence, by Hawkes' theorem, Theorem 9.5, we have $\dim F(a) \leq 1-a^2$ almost surely. $\diamond$

Theorem 10.22 can be seen as a probabilistic interpretation of packing dimension. The upper and lower Minkowski dimensions allow a similar definition when the order of sup and lim are interchanged.

THEOREM 10.25. *For any compact $E \subset [0,1]$ almost surely,*

$$\limsup_{h\downarrow 0} \sup_{t\in E} \frac{|B(t+h)-B(t)|}{\sqrt{2h\log(1/h)}} = \sqrt{\overline{\dim}_M(E)}. \tag{2.1}$$

Proof of the upper bounds in Theorem 10.22 and Theorem 10.25. Suppose $E \subset [0,1]$ is compact. We assume that $\overline{\dim}_M(E) < \lambda$ for some $\lambda < a^2$ and show that

$$\limsup_{h\downarrow 0} \sup_{t\in E} \frac{|B(t+h)-B(t)|}{\sqrt{2h\log(1/h)}} \le \lambda \qquad \text{almost surely.} \tag{2.2}$$

Note that this is the upper bound in Theorem 10.25. Once this is shown it immediately implies

$$\sup_{t\in E} \limsup_{h\downarrow 0} \frac{|B(t+h)-B(t)|}{\sqrt{2h\log(1/h)}} \le \overline{\dim}_M(E) \qquad \text{almost surely.}$$

Now, for any decomposition $E = \bigcup_{i=1}^\infty E_i$, we have

$$\sup_{t\in E} \limsup_{h\downarrow 0} \frac{|B(t+h)-B(t)|}{\sqrt{2h\log(1/h)}} = \sup_{i=1}^{\infty} \sup_{t\in \mathrm{cl}(E_i)} \limsup_{h\downarrow 0} \frac{|B(t+h)-B(t)|}{\sqrt{2h\log(1/h)}} \le \sup_{i=1}^{\infty} \overline{\dim}_M(E_i),$$

where we have made use of the fact that the upper Minkowski dimension is insensitive under taking the closure of a set. Theorem 10.15 now implies that

$$\sup_{t\in E} \limsup_{h\downarrow 0} \frac{|B(t+h)-B(t)|}{\sqrt{2h\log(1/h)}} \le \dim_P(E) \qquad \text{almost surely,}$$

which is the upper bound in Theorem 10.22.

For the proof of (2.2) cover E by disjoint subintervals $I = [(j/k)\eta^{-k}, ((j+1)/k)\eta^{-k})$ for $j = 0, \ldots, \lceil k\eta^k - 1\rceil$, of equal length $h = \eta^{-k}/k$ such that $I \cap E \neq \emptyset$. By definition of the upper Minkowski dimension there exists an m such that, for all $k \ge m$, no more than $\eta^{\lambda k}$ different such intervals of length $h = \eta^{-k}/k$ intersect E.

Now fix $\varepsilon > 0$ such that $\lambda < \frac{a^2}{2}(1-\varepsilon)^3$, which is possible by our condition on λ. Let $Z(I) = 1$ if , for $h' = \eta^{-k}$,

$$|B(jh+h') - B(jh)| > a(1-3\varepsilon)\sqrt{h'\log(1/h')}.$$

Recall from the proof of Lemma 10.4 that there is an $\eta > 1$ such that, for any $m \in \mathbb{N}$, the collection

$$\big\{ I = [(j/k)\eta^{-k}, ((j+1)/k)\eta^{-k}) \,:\, Z(I) = 1, I\cap E \neq \emptyset, k \ge m \big\}$$

is a covering of the set

$$M(m) := \Big\{ t \in E \,:\, \sup_{\eta^{-k} < u \le \eta^{-k+1}} \frac{|B(t+u)-B(t)|}{\sqrt{u\log(1/u)}} \ge a(1-\varepsilon) \text{ for some } k \ge m \Big\}.$$

Moreover, we recall from (1.2), that

$$\mathbb{P}\{Z(I) = 1\} \le \eta^{-k\frac{a^2}{2}(1-\varepsilon)^3},$$

and, sticking to our notation $I = [(j/k)\eta^{-k}, ((j+1)/k)\eta^{-k})$ for a little while longer,

$$\sum_{k=0}^{\infty} \sum_{j=0}^{\lceil k\eta^k - 1\rceil} \mathbb{P}\big\{Z(I) = 1\big\}\, \mathbb{1}\{I \cap E \neq \emptyset\} \leq \sum_{k=0}^{\infty} \eta^{\lambda k} \eta^{-k\frac{a^2}{2}(1-\varepsilon)^3} < \infty,$$

and hence by the Borel-Cantelli lemma there exists an m such that $Z(I) = 0$ whenever $I = [(j/k)\eta^{-k}, ((j+1)/k)\eta^{-k})$ for some $k \geq m$. This means that the set $M(m)$ can be covered by the empty covering, so it must itself be empty. This shows (2.2) and completes the proof. ∎

We are going to embed the proof of the lower bound now into a more general framework, including the discussion of limsup fractals in a d-dimensional cube.

Definition 10.26. *Fix an open unit cube* $\mathsf{Cube} = x_0 + (0,1)^d \subset \mathbb{R}^d$. *For any nonnegative integer* k, *denote by* $\mathfrak{C}_k$ *the collection of dyadic cubes*

$$x_0 + \prod_{i=1}^{d} [j_i 2^{-k}, (j_i+1)2^{-k}] \qquad \text{with } j_i \in \{0, \ldots, 2^k - 1\} \text{ for all } i \in \{1, \ldots, d\},$$

and $\mathfrak{C} = \bigcup_{k \geq 0} \mathfrak{C}_k$. *Denote by* $(Z(I) \colon I \in \mathfrak{C})$ *a collection of random variables each taking values in* $\{0,1\}$. *The* **limsup fractal** *associated to this collection is the random set*

$$A := \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \bigcup_{\substack{I \in \mathfrak{C}_k \\ Z(I)=1}} \operatorname{int}(I),$$

where $\operatorname{int}(I)$ *is the interior of the cube* I. ◇

Remark 10.27. Compared with the setup of the previous section we have switched to the use of *open* cubes in the definition of limsup fractals. This choice is more convenient when we prove hitting estimates, whereas in Proposition 10.6 the choice of closed cubes was more convenient when constructing random measures on A. ◇

The key to our result are the hitting probabilities for the discrete limsup fractal A under some conditions on the random variables $(Z(I) \colon I \in \mathfrak{C})$.

Theorem 10.28. *Suppose that*

(i) *the means* $p_n = \mathbb{E}[Z(I)]$ *are independent of the choice of* $I \in \mathfrak{C}_n$ *and satisfy*

$$\liminf_{n\uparrow\infty} \frac{\log p_n}{n \log 2} \geq -\gamma, \qquad \text{for some } \gamma > 0;$$

(ii) *there exists* $c > 0$ *such that the random variables* $Z(I)$ *and* $Z(J)$ *are independent whenever* $I, J \in \mathfrak{C}_n$ *and the distance of* I *and* J *exceeds* $cn2^n$.

Then, for any compact $E \subset \mathsf{Cube}$ *with* $\dim_P(E) > \gamma$, *we have*

$$\mathbb{P}\big\{A \cap E \neq \emptyset\big\} = 1.$$

Remark 10.29. The second assumption, which gives us the necessary independence for the lower bound, can be weakened, see [**KPX00**]. Note that no assumption is made concerning the dependence of random variables $Z(I)$ for intervals I of different size. $\diamond$

Proof of Theorem 10.28 Let $E \subset \mathsf{Cube}$ be compact with $\dim_P E > \gamma$. Let $\widetilde{E}$ be defined as in Lemma 10.18 for example as

$$\widetilde{E} = E \setminus \bigcup_{\substack{a_i<b_i \\ \text{rational}}} \Big\{ \prod_{i=1}^{d} (a_i, b_i) \,:\, \overline{\dim}_M \big(E \cap \prod_{i=1}^{d} (a_i, b_i) \big) < \gamma \Big\}.$$

From the proof of Lemma 10.18 we have $\dim_P E = \dim_P \widetilde{E}$. Define open sets

$$A_n = \bigcup_{I \in \mathfrak{C}_n} \big\{ \mathrm{int}(I) \,:\, Z(I) = 1 \big\},$$

and

$$A_n^* = \bigcup_{m \geq n} A_n = \bigcup_{m \geq n} \bigcup_{I \in \mathfrak{C}_m} \big\{ \mathrm{int}(I) \,:\, Z(I) = 1 \big\}.$$

By definition $A_n^* \cap \widetilde{E}$ is open in $\widetilde{E}$. We will show that it is also dense in $\widetilde{E}$ with probability one. This, by Baire's category theorem, will imply that

$$A \cap \widetilde{E} = \bigcap_{n=1}^{\infty} A_n^* \cap \widetilde{E} \neq \emptyset, \qquad \text{almost surely,}$$

as required. To show that $A_n^* \cap \widetilde{E}$ is dense in $\widetilde{E}$, we need to show that for any open binary cube J which intersects $\widetilde{E}$, the set $A_n^* \cap \widetilde{E} \cap J$ is almost surely non-empty.

For the rest of the proof, take $\varepsilon > 0$ small and n large enough so that $\widetilde{E} \cap J$ intersects more than $2^{n(\gamma+2\varepsilon)}$ binary cubes of sidelength 2^{-n}, and so that $(\log p_n)/n > -(\log 2)(\gamma+\varepsilon)$. Let $\mathcal{S}_n$ be the set of cubes in $\mathfrak{C}_n$ that intersect $\widetilde{E}$. Define

$$T_n = \sum_{I \in \mathcal{S}_n} Z(I),$$

so that $\mathbb{P}\big\{ A_n \cap \widetilde{E} \cap J = \emptyset \big\} = \mathbb{P}\{T_n = 0\}$. To show that this probability converges to zero, by the Paley-Zygmund inequality, it suffices to prove that $(\operatorname{Var} T_n)/(\mathbb{E} T_n)^2$ does. The first moment of T_n is given by

$$\mathbb{E} T_n = s_n \, p_n > 2^{(\gamma+2\varepsilon)n} 2^{-\gamma-\varepsilon n} = 2^{\varepsilon n},$$

where s_n denotes the cardinality of $\mathcal{S}_n$. The variance can be written as

$$\operatorname{Var} T_n = \operatorname{Var} \sum_{I \in \mathcal{S}_n} Z(I) = \sum_{I \in \mathcal{S}_n} \sum_{J \in \mathcal{S}_n} \operatorname{Cov}(Z(I), Z(J)).$$

Here each summand is at most p_n, and the summands for which I and J have distance at least $cn2^{-n}$ vanish by assumption. Thus

$$\begin{aligned} \sum_{I \in \mathcal{S}_n} \sum_{J \in \mathcal{S}_n} \operatorname{Cov}(Z(I), Z(J)) &\leq p_n \, \#\big\{ (I, J) \in \mathcal{S}_n \times \mathcal{S}_n \,:\, \operatorname{dist}(I, J) \leq cn\, 2^{-n} \big\} \\ &\leq p_n s_n \, c(2n+1) = c(2n+1)\, \mathbb{E} T_n. \end{aligned}$$

This implies that $(\mathrm{Var}\ T_n)/(\mathbb{E}T_n)^2 \to 0$. Hence, almost surely, A_n^* is an open dense set, concluding the proof. ■

We now show how the main statement of Theorem 10.22 follows from this, and how the ideas in the proof also lead to the lower bound in Theorem 10.25.

Proof the lower bound in Theorem 10.22 and Theorem 10.25. For the lower bound we look at a compact set $E \subset (0,1)$ with $\dim_P(E) > a^2/2$ and first go for the result in Theorem 10.22. Choose $\varepsilon > 0$ such that $\dim_P(E) > \frac{a^2}{2}(1+\varepsilon)^3$. Associate to every dyadic interval $I = [jh, (j+1)h] \in \mathfrak{C}_k$ with $h = 2^{-k}$ the random variable $Z(I)$, which takes the value one if and only if, for $h' = k2^{-k}$,

$$|B(jh+h') - B(jh)| \geq a(1+\varepsilon)\sqrt{h'\log(1/h')},$$

and note that by Lemma 10.10 the limsup fractal associated to these random variables is contained in the set of a-fast times. It remains to note that the collection $\{Z(I) : I \in \mathfrak{C}_k, k \geq 0\}$ satisfies the condition (i) with $\gamma = \frac{a^2}{2}(1+\varepsilon)^3$ by (1.7) and condition (ii) with $c = 1$. Theorem 10.28 now gives that

$$\mathbb{P}\big\{A \cap E \neq \emptyset\big\} = 1,$$

and therefore

$$\sup_{t\in E}\limsup_{h\downarrow 0}\frac{|B(t+h)-B(t)|}{\sqrt{2h\log(1/h)}} \geq \sqrt{\dim_P(E)}.$$

For the lower bound in Theorem 10.25 we look at a compact set $E \subset (0,1)$ with $\overline{\dim}_M(E) > a^2/2$ and fix $\varepsilon > 0$ such that $\overline{\dim}_M(E)\,(1-\varepsilon)^2 \geq \frac{a^2}{2}$. Hence there exists a sequence $(n_k : k \in \mathbb{N})$ such that

$$\#\big\{I \in \mathfrak{C}_{n_k} : I \cap E \neq \emptyset\big\} \geq 2^{n_k \frac{a^2}{2}(1-\varepsilon)}.$$

With $Z(I)$ defined as above we obtain, using notation and proof of Theorem 10.28, that

$$\mathbb{P}\{Z(I) = 1\} \leq 2^{\gamma}, \qquad \text{with } \gamma = \frac{a^2}{2}(1+\varepsilon)^4,$$

and

$$\mathrm{Var}\, T_{n_k} \leq (2n_k+1)\,\mathbb{E}T_{n_k}, \qquad \text{for } T_n = \sum_{I\in\mathfrak{C}_n} Z(I)\,\mathbb{1}\{I \cap E \neq \emptyset\}.$$

By Chebyshev's inequality we get, for $1/2 < \eta < 1$,

$$\mathbb{P}\big\{|T_{n_k} - \mathbb{E}T_{n_k}| \geq (\mathbb{E}T_{n_k})^{\eta}\big\} \leq (2n_k+1)\,(\mathbb{E}T_{n_k})^{1-2\eta}.$$

As $\mathbb{E}T_n$ is exponentially increasing we can infer, using the Borel-Cantelli lemma, that

$$\lim_{k\uparrow\infty}\frac{T_{n_k}}{\mathbb{E}T_{n_k}} = 1 \qquad \text{almost surely.}$$

This implies that $T_{n_k} \neq 0$ for all sufficiently large k. Hence, as $Z(I) = 1$ and $I \cap E \neq \emptyset$ imply that there exists $t \in I \cap E$ with with $|B(t+h') - B(t)| \geq a\sqrt{h'\log(1/h')}$ for $h' = n_k 2^{n_k}$, by the proof of Lemma 10.10, we have completed the proof of Theorem 10.25. ■

3. Slow times of Brownian motion

At the fast times Brownian motion has, in infinitely many small scales, an unusually large growth. Conversely, one may ask whether there are times where a Brownian path has, at *all* small scales, unusually small growth.

This question of *slow times* for the Brownian motion is related to nondifferentiability of the Brownian path. Indeed, in our proof of non-differentiability, we showed that almost surely,

$$\limsup_{h\downarrow 0} \frac{|B(t+h)-B(t)|}{h} = \infty, \qquad \text{for all } t\in[0,1],$$

and in 1963 Dvoretzky showed that there exists a constant $\delta>0$ such that almost surely,

$$\limsup_{h\downarrow 0} \frac{|B(t+h)-B(t)|}{\sqrt{h}} > \delta, \qquad \text{for all } t\in[0,1].$$

In 1983 Davis and, independently, Perkins and Greenwood, found that the optimal constant in this result is equal to one.

Theorem 10.30 (Davis, Perkins and Greenwood). *Almost surely,*

$$\inf_{t\in[0,1]} \limsup_{h\downarrow 0} \frac{|B(t+h)-B(t)|}{\sqrt{h}} = 1.$$

Remark 10.31. We call a time $t\in[0,1]$ an a-**slow time** if

$$\limsup_{h\downarrow 0} \frac{|B(t+h)-B(t)|}{\sqrt{h}} \le a. \tag{3.1}$$

The result shows that a-slow times exist for $a>1$ but not for $a<1$. The Hausdorff dimension of the set of a-slow times is studied in Perkins [**Pe83**]. ◇

For the proof of Theorem 10.30 we need to investigate the probability that the graph of a Brownian motion stays within a parabola open to the right. The following lemma is what we need for an upper bound.

Lemma 10.32. *Let* $M := \max_{0\le t\le 1}|B(t)|$ *and, for* $r<1$, *define the stopping time*

$$T = \inf\{t\ge 1 \,:\, |B(t)| = M + r\sqrt{t}\}.$$

Then $\mathbb{E}T<\infty$.

Proof. By Theorem 2.44, for every $t\ge 1$, we have

$$\begin{aligned}\mathbb{E}[T\wedge t] &= \mathbb{E}[B(T\wedge t)^2] \le \mathbb{E}\big[(M+r\sqrt{T\wedge t})^2\big]\\ &= \mathbb{E}M^2 + 2r\mathbb{E}\big[M\sqrt{T\wedge t}\,\big] + r^2\mathbb{E}[T\wedge t]\\ &\le \mathbb{E}M^2 + 2r(\mathbb{E}M^2)^{1/2}(\mathbb{E}[T\wedge t])^{1/2} + r^2\mathbb{E}[T\wedge t],\end{aligned}$$

where Hölder's inequality was used in the last step. This gives

$$(1-r^2)\mathbb{E}[T\wedge t] \le \mathbb{E}[M^2] + 2r\big(\mathbb{E}M^2\big)^{1/2}\big(\mathbb{E}[T\wedge t]\big)^{1/2},$$

and as $\mathbb{E}[M^2] < \infty$ we get that $\mathbb{E}[T \wedge t]$ is bounded and hence $\mathbb{E}T < \infty$. ■

Proof of the upper bound in Theorem 10.30. We show that, for $r < 1$ the set

$$A = \big\{t \in [0,1] \,:\, \limsup_{h \downarrow 0} |B(t+h) - B(t)| < r\sqrt{h}\big\}$$

is empty almost surely. The crucial input is that, for any interval $I = [a,b] \subset [0,1]$, we have by the triangle inequality and Brownian scaling, for $M = \max\{|B(t)|\colon 0 \le t \le b-a\}$,

$$\begin{aligned}\mathbb{P}\big\{\exists t \in I \,:\, & |B(t+h) - B(t)| < r\sqrt{h} \text{ for all } 0 < h \le 1\big\}\\ &\le \mathbb{P}\big\{|B(a+h) - B(a)| < M + r\sqrt{h} \text{ for all } b-a < h \le 1\big\}\\ &\le \mathbb{P}\big\{T \ge \tfrac{1}{b-a}\big\}.\end{aligned}$$

Now, dividing $[0,1]$ into n intervals of length $1/n$ we get

$$\mathbb{P}\{A \ne \emptyset\} \le \liminf_{n \to \infty} \sum_{k=0}^{n-1} \mathbb{P}\{A \cap [k/n, (k+1)/n] \ne \emptyset\} \le \liminf_{n \to \infty} n\mathbb{P}\big\{T \ge n\big\} = 0,$$

using in the final step that $\infty > \mathbb{E}T \ge \sum_n \mathbb{P}\big\{T \ge n\big\}$ and divergence of the harmonic series. ■

We turn to the proof of the upper bound. Again we start by studying exit times from a parabola. For $0 < r < \infty$ and $a > 0$ let

$$T(r,a) := \inf\{t \ge 0 \,:\, |B(t)| = r\sqrt{t+a}\}.$$

For the moment it suffices to note the following property of $T(1,a)$.

Lemma 10.33. *We have $\mathbb{E}T(1,a) = \infty$.*

Proof. Suppose that $\mathbb{E}T(1,a) < \infty$. Then, by Theorem 2.44, we have that $\mathbb{E}T(1,a) = \mathbb{E}B(T(1,a))^2 = \mathbb{E}T(1,a) + a$, which is a contradiction. Hence $\mathbb{E}T(1,a) = \infty$. ■

For $0 < r < \infty$ and $a > 0$ we now define further stopping times

$$S(r,a) := \inf\{t \ge a \,:\, |B(t)| \ge r\sqrt{t}\}.$$

Lemma 10.34. *If $r > 1$ there is a $p = p(r) < 1$ such that $\mathbb{E}[S(r,1)^p] = \infty$. In particular,*

$$\limsup_{n \uparrow \infty} n \frac{\mathbb{P}\{S(r,1) > n\}}{\mathbb{E}[S(r,1) \wedge n]} > 0.$$

The proof uses the following general lemma.

Lemma 10.35. *Suppose X is a nonnegative random variable and $\mathbb{E}X^p = \infty$ for some $p < 1$. Then*

$$\limsup_{n \uparrow \infty} n\, \mathbb{P}\{X > n\}/\mathbb{E}[X \wedge n] > 0.$$

Proof. Let $p<1$ and suppose for contradiction that, for some $\varepsilon < 1/(2+\frac{2}{1-p})$,

$$n\,\mathbb{P}\{X>n\} < \varepsilon\,\mathbb{E}[X\wedge n] \qquad \text{for all integers } n > y_0 \ge 2.$$

We have that

$$\mathbb{E}[(X\wedge N)^p] = \int_0^{N^p} \mathbb{P}\{X^p > x\}\,dx = p\int_0^N y^{p-1}\,\mathbb{P}\{X>y\}\,dy\,,$$

and hence, for all N,

$$\begin{aligned}
\mathbb{E}[(X\wedge N)^p] &\le p\int_0^{y_0} y^{p-1}\,dy + \varepsilon\,\tfrac{y_0}{y_0-1}\,p\int_{y_0}^N y^{p-2}\,\mathbb{E}[X\wedge y]\,dy\\
&\le y_0^p + 2\varepsilon p\int_0^N y^{p-1}\,\mathbb{P}\{X>y\}\,dy + 2\varepsilon p\int_{y_0}^N y^{p-2}\int_0^y \mathbb{P}\{X>z\}\,dz\,dy\\
&\le y_0^p + 2\varepsilon\,\mathbb{E}[(X\wedge N)^p] + 2\varepsilon p\int_0^N \mathbb{P}\{X>z\}\int_z^\infty y^{p-2}\,dy\,dz\\
&\le y_0^p + \varepsilon\,2\left(1+\tfrac{1}{1-p}\right)\mathbb{E}[(X\wedge N)^p]\,.
\end{aligned}$$

This implies that $\mathbb{E}[X^p] = \sup \mathbb{E}[(X\wedge N)^p] < \infty$, and so the statement follows. ∎

Proof of Lemma 10.34. Define a sequence of stopping times by $\tau_0 = 1$ and, for $k\ge 1$,

$$\tau_k = \begin{cases} \inf\{t\ge \tau_{k-1} \,:\, B(t)=0 \text{ or } |B(t)|\ge r\sqrt{t}\} & \text{if } k \text{ odd},\\ \inf\{t\ge\tau_{k-1} \,:\, |B(t)|\ge\sqrt{t}\} & \text{if } k \text{ even}.\end{cases}$$

By the strong Markov property and Brownian scaling we get, for any $\lambda>0$,

$$\mathbb{P}\{\tau_{2k}-\tau_{2k-1} > \lambda\tau_{2k-1}\,|\,B(\tau_{2k-1})=0\} = \mathbb{P}\{T(1,1)>\lambda\}.$$

Define, with $\mathbb{P}_1$ referring to a Brownian motion started in $B(0)=1$,

$$c := \mathbb{P}_1\big\{\inf\{t\ge 0 \,:\, B(t)=0\} < \inf\{t\ge 0 \,:\, |B(t)| = r\sqrt{t+1}\}\big\}.$$

Now, for $k\ge 2$ and $\lambda>0$, on $\{\tau_{2k-2} < S(r,1)\}$,

$$\begin{aligned}
&\mathbb{P}\big\{\tau_{2k}-\tau_{2k-1} > \lambda\tau_{2k-2}\,\big|\,\mathcal{F}(\tau_{2k-2})\big\}\\
&\quad\ge \mathbb{P}\big\{\tau_{2k}-\tau_{2k-1}>\lambda\tau_{2k-1}\,\big|\,\mathcal{F}(\tau_{2k-2}), B(\tau_{2k-1})=0\big\}\,\mathbb{P}\big\{B(\tau_{2k-1})=0\,\big|\,\mathcal{F}(\tau_{2k-2})\big\}\\
&\quad= c\,\mathbb{P}\{T(1,1)>\lambda\}.
\end{aligned}$$

To pass from this estimate to the p^{th} moments we use that, for any nonnegative random variable X, we have $\mathbb{E}X^p = \int_0^\infty \mathbb{P}\{X^p>\lambda\}\,d\lambda$. This is an easy consequence of Fubini's theorem and gives

$$\begin{aligned}
\mathbb{E}\big[(\tau_{2k}-\tau_{2k-1})^p\big] &= \mathbb{E}\int_0^\infty \tau_{2k-2}^p\,\mathbb{P}\big\{(\tau_{2k}-\tau_{2k-1})^p > \lambda\tau_{2k-2}^p\,\big|\,\mathcal{F}(\tau_{2k-2})\big\}\,d\lambda\\
&\ge \mathbb{E}\int_0^\infty \tau_{2k-2}^p\,\mathbb{P}\big\{\tau_{2k}-\tau_{2k-1}>\lambda^{1/p}\tau_{2k-2}\,\big|\,\mathcal{F}(\tau_{2k-2})\big\}\,\mathbb{1}\{\tau_{2k-2}<S(r,1)\}\,d\lambda\\
&\ge c\,\mathbb{E}\int_0^\infty \tau_{2k-2}^p\,\mathbb{P}\big\{T(1,1)>\lambda^{1/p}\big\}\,\mathbb{1}\{\tau_{2k-2}<S(r,1)\}\,d\lambda.
\end{aligned}$$

Now, using the formula for $\mathbb{E}X^p$ again, but for $X = T(1,1)$ and noting that $\{\tau_{2k-2} < S(r,1)\} = \{\tau_{2k-3} < \tau_{2k-2}\}$, we obtain

$$\begin{aligned}\mathbb{E}\big[(\tau_{2k} - \tau_{2k-1})^p\big] &\geq c\mathbb{E}[T(1,1)^p]\,\mathbb{E}\big[\tau_{2k-2}^p \mathbb{1}\{\tau_{2k-2} < S(r,1)\}\big] \\ &\geq c\mathbb{E}[T(1,1)^p]\,\mathbb{E}\big[(\tau_{2k-2} - \tau_{2k-3})^p\big],\end{aligned}$$

and by iterating this,

$$\mathbb{E}\big[(\tau_{2k} - \tau_{2k-1})^p\big] \geq \Big(c\mathbb{E}[T(1,1)^p]\Big)^{k-1}\,\mathbb{E}\big[(\tau_2 - \tau_1)^p\big].$$

Note that, by Fatou's lemma and by Lemma 10.33,

$$\liminf_{p\uparrow 1} \mathbb{E}[T(1,1)^p] \geq \mathbb{E}T(1,1) = \infty.$$

Hence we may pick $p < 1$ such that $\mathbb{E}[T(1,1)^p] > 1/c$. Then

$$\mathbb{E}[S(r,1)^p] \geq \mathbb{E}[\tau_{2k}^p] \geq \mathbb{E}\big[(\tau_{2k} - \tau_{2k-1})^p\big] \longrightarrow \infty,$$

as $k \uparrow \infty$, which is the first statement we wanted to prove. The second statement follows directly from the general fact stated as Lemma 10.35. ■

Proof of the lower bound in Theorem 10.30. Fix $r > 1$ and let

$$A(n) = \big\{t \in [0,1] \,:\, |B(t+h) - B(t)| < r\sqrt{h}, \text{ for all } \tfrac{1}{n} \leq h \leq 1\big\}.$$

Note that $n \geq m$ implies $A(n) \subset A(m)$. We show that

$$\mathbb{P}\Big\{\bigcap_{n=1}^{\infty} A(n) \neq \emptyset\Big\} = \lim_{n\to\infty} \mathbb{P}\big\{A(n) \neq \emptyset\big\} > 0. \tag{3.2}$$

For $n \in \mathbb{N}$ and let $v(0,n) = 0$ and, for $i \geq 1$,

$$\begin{aligned} v(i,n) &:= (v(i-1,n) + 1) \\ &\qquad \wedge \inf\big\{t \geq v(i-1,n) + \tfrac{1}{n} \,:\, |B(t) - B(v(i-1,n))| \geq r\sqrt{t - v(i-1,n)}\big\}.\end{aligned}$$

Then $\mathbb{P}\{v(i+1,n) - v(i,n) = 1 \,|\, \mathcal{F}(v(i,n))\} = \mathbb{P}\{S(r,1) \geq n\}$, and by Brownian scaling,

$$\mathbb{E}[v(i+1,n) - v(i,n) \,|\, \mathcal{F}(v(i,n))] = \tfrac{1}{n}\mathbb{E}[S(r,1) \wedge n]. \tag{3.3}$$

Of course $v(k,n) \geq 1$ if $v(i,n) - v(i-1,n) = 1$ for some $i \leq k$. Thus, for any m,

$$\begin{aligned}\mathbb{P}\{v(i+1,n) &- v(i,n) = 1 \text{ for some } i \leq m \text{ such that } v(i,n) \leq 1\big\} \\ &= \sum_{i=1}^{m} \mathbb{P}\{S(r,1) \geq n\}\mathbb{P}\{v(i,n) \leq 1\} \\ &\geq m\mathbb{P}\{S(r,1) \geq n\}\mathbb{P}\{v(m,n) \leq 1\}.\end{aligned}$$

Let $(n_k \,:\, k \in \mathbb{N})$ be an increasing sequence of integers such that

$$n_k \frac{\mathbb{P}\{S(r,1) \geq n_k\}}{\mathbb{E}[S(r,1) \wedge n_k]} \geq \varepsilon > 0,$$

and $\mathbb{E}[S(r,1) \wedge n_k] \leq n_k/6$ for all k, which is possible by Lemma 10.34.

Choose the integers m_k so that they satisfy

$$\frac{1}{3} \le \frac{m_k}{n_k}\, \mathbb{E}[S(r,1) \wedge n_k] \le \frac{1}{2}.$$

Summing (3.3) over all $i = 1, \dots, m_k - 1$,

$$\mathbb{E}v(m_k, n_k) = \frac{m_k}{n_k}\, \mathbb{E}[S(r,1) \wedge n_k],$$

hence $\mathbb{P}\{v(m_k, n_k) \ge 1\} \le 1/2$. Now we get, putting all our ingredients together,

$$\begin{aligned}
\mathbb{P}\{A(n_k) \ne \emptyset\} &\ge \mathbb{P}\big\{v(i+1, n_k) - v(i, n_k) = 1 \text{ for some } i \le m_k \text{ such that } v(i, n_k) \le 1\big\} \\
&\ge m_k \mathbb{P}\{S(r,1) \ge n_k\} \mathbb{P}\{v(m_k, n_k) \le 1\} \\
&\ge m_k \mathbb{P}\{S(r,1) \ge n_k\}/2 \ge \frac{m_k}{2n_k}\, \varepsilon \mathbb{E}[S(r,1) \wedge n_k] \ge \frac{\varepsilon}{6}.
\end{aligned}$$

This proves (3.2). It remains to observe that, by Brownian scaling, there exists $\delta > 0$ such that, for all $n \in \mathbb{N}$,

$$\mathbb{P}\big\{\exists t \in [0, 1/n] \,:\, \limsup_{h \downarrow 0} |B(t+h) - B(t)|/\sqrt{h} \le r\big\} \ge \delta.$$

Hence, by independence,

$$\mathbb{P}\big\{\exists t \in [0, 1] \,:\, \limsup_{h \downarrow 0} |B(t+h) - B(t)|/\sqrt{h} \le r\big\} \ge 1 - (1-\delta)^n \longrightarrow 1.$$

This completes the proof of the lower bound, and hence completes the proof of Theorem 10.30. ■

4. Cone points of planar Brownian motion

We now focus on a planar Brownian motion $\{B(t) \,:\, t \ge 0\}$. Recall from Section 7.2 that around a *typical point* on the path this motion performs an infinite number of windings in both directions. It is easy to see that there are exceptional points for this behaviour: Let

$$x = \min\{x \colon (x, 0) \in B[0, 1]\}.$$

Then Brownian motion does not perform *any* windings around $(x, 0)$, as this would necessarily imply that it crosses the half-line $\{(y, 0) \colon y < x\}$ contradicting the minimality of x. More generally, each point $(x, y) \in \mathbb{R}^2$ with $x = \min\{x \colon (x, y) \in B[0, 1]\}$ has this property, if the set is nonempty. Hence, the set of such points has dimension at least one, as the projection onto the y-axis gives a nontrivial interval. We shall see below that this set has indeed Hausdorff dimension one.

We now look at points where a cone-shaped area with the tip of the cone placed in the point is avoided by the Brownian motion. These points are called cone points.

Definition 10.36. *Let $\{B(t) \,:\, t \ge 0\}$ be a planar Brownian motion. For any angle $\alpha \in (0, 2\pi)$ and direction $\xi \in [0, 2\pi)$, define the closed* **cone**

$$W[\alpha, \xi] := \big\{x = re^{\mathbf{i}(\theta - \xi)} \colon |\theta| \le \alpha/2, r \ge 0\big\} \subset \mathbb{R}^2.$$

Given a cone $x + W[\alpha, \xi]$ we call its **dual** *the reflection of its complement about the tip, i.e. the cone $x + W[2\pi - \alpha, \xi + \pi]$. A point $x = B(t)$, $0 < t < 1$, is an α-***cone point** *if there exists $\varepsilon > 0$ and $\xi \in [0, 2\pi)$ such that*

$$B(0,1) \cap \mathcal{B}(x, \varepsilon) \subset x + W[\alpha, \xi]\,. \qquad \diamond$$

REMARK 10.37. Clearly, if $x = B(t)$ is a cone point, then there exists a small $\delta > 0$ such that $B(t - \delta, t + \delta) \subset x + W[\alpha, \xi]$. Hence the path $\{B(t) \colon 0 \le t \le 1\}$ performs only a finite number of windings around x. $\diamond$

We now identify the angles α for which there exist α-cone points, and, if they exist, determine the Hausdorff dimension of the set of α-cone points.

THEOREM 10.38 (Evans 1985). *Let $\{B(t) \colon 0 \le t \le 1\}$ be a planar Brownian motion. Then, almost surely, α-cone points exist for any $\alpha \ge \pi$ but not for $\alpha < \pi$. Moreover, if $\alpha \in [\pi, 2\pi)$, then*

$$\dim \big\{ x \in \mathbb{R}^2 \,:\, x \text{ is an } \alpha\text{-cone point} \big\} = 2 - \tfrac{2\pi}{\alpha}\,.$$

In the proof of Theorem 10.38 we identify $\mathbb{R}^2$ with the complex plane and use complex notation wherever convenient. Suppose that $\{B(t) \colon t \ge 0\}$ is a planar Brownian motion defined for all positive times. We first fix an angle $\alpha \in (0, 2\pi)$ and a direction $\xi \in [0, 2\pi)$ and define the notion of an approximate cone point as follows: For any $0 < \delta < \varepsilon$ we let

$$T_\delta(z) := \inf \big\{ s \ge 0 \colon B(s) \in \mathcal{B}(z, \delta) \big\}$$

and

$$S_{\delta,\varepsilon}(z) := \inf \big\{ s \ge T_{\delta/2}(z) \colon B(s) \notin \mathcal{B}(z, \varepsilon) \big\}\,.$$

We say that $z \in \mathbb{R}^2$ is a (δ, ε)-*approximate cone point* if

$$B(0, T_\delta(z)) \subset z + W[\alpha, \xi], \quad \text{and} \quad B(T_{\delta/2}(z), S_{\delta,\varepsilon}(z)) \subset z + W[\alpha, \xi]\,.$$

Note that we do not require (δ, ε)-approximate cone points to belong to the Brownian path. The relation between cone points and approximate cone points will become clear later, we first collect the necessary information about the probability that a given point is a (δ, ε)-approximate cone point. The strong Markov property allows us to consider the events happening during the intervals $[0, T_\delta(z)]$ and $[T_{\delta/2}(z), 1]$ separately.

LEMMA 10.39. *There exist constants $C > c > 0$ such that, for every $\delta > 0$,*

(a) *for all $z \in \mathbb{R}^2$,*

$$\mathbb{P}\big\{ B(0, T_\delta(z)) \subset z + W[\alpha, \xi] \big\} \le C \left(\tfrac{\delta}{|z|} \right)^{\frac{\pi}{\alpha}},$$

(b) *for all $z \in \mathbb{R}^2$ with $0 \in z + W[\alpha/2, \xi]$,*

$$\mathbb{P}\big\{ B(0, T_\delta(z)) \subset z + W[\alpha, \xi] \big\} \ge c \left(\tfrac{\delta}{|z|} \right)^{\frac{\pi}{\alpha}}.$$

Proof. We write $z = |z|\, e^{\mathbf{i}\theta}$ and apply the skew-product representation, Theorem 7.25, to the Brownian motion $\{z - B(t) \colon t \geq 0\}$ and obtain

$$B(t) = z - R(t)\, \exp\big(\mathbf{i}\, \theta(t)\big), \text{ for all } t \geq 0,$$

for $R(t) = \exp(W_1(H(t))$ and $\theta(t) = W_2(H(t))$, where $\{W_1(t) \colon t \geq 0\}$ and $\{W_2(t) \colon t \geq 0\}$ are independent linear Brownian motions started in $\log |z|$, resp. in θ, and a strictly increasing time-change $\{H(t) \colon t \geq 0\}$ which depends only on the first of these motions. This implies that $T_\delta(z) = \inf\{s \geq 0 \colon R(s) \leq \delta\}$ and therefore

$$H(T_\delta(z)) = \inf\big\{u \geq 0 \colon W_1(u) \leq \log \delta\big\} =: \tau_{\log \delta}\,.$$

We infer that

$$\big\{B(0, T_\delta(z)) \subset z + W[\alpha, \xi]\big\} = \big\{|W_2(u) + \pi - \xi| \leq \tfrac{\alpha}{2} \text{ for all } u \in [0, \tau_{\log \delta}]\big\}.$$

The latter event means that a linear Brownian motion started in θ stays inside the interval $[\xi - \pi - \alpha/2, \xi - \pi + \alpha/2]$ up to the independent random time $\tau_{\log \delta}$. For the probability of such events we have found two formulas, (4.2) and (4.3) in Chapter 7. The latter formula gives

$$\begin{aligned}
&\mathbb{P}\big\{|W_2(u) + \pi - \xi| \leq \tfrac{\alpha}{2} \text{ for all } u \in [0, \tau_{\log \delta}]\big\} \\
&\quad = \sum_{k=0}^{\infty} \tfrac{4}{(2k+1)\,\pi}\, \sin\big(\tfrac{(2k+1)\pi(\alpha/2 + \xi - \pi - \theta)}{\alpha}\big)\, \mathbb{E}\Big[\exp\big(-\tfrac{(2k+1)^2\pi^2}{2\alpha^2}\, \tau_{\log \delta}\big)\Big] \\
&\quad = \sum_{k=0}^{\infty} \tfrac{4}{(2k+1)\,\pi}\, \sin\big(\tfrac{(2k+1)\pi(\alpha/2 + \xi - \pi - \theta)}{\alpha}\big)\, \big(\tfrac{\delta}{|z|}\big)^{(2k+1)\,\frac{\pi}{\alpha}},
\end{aligned}$$

using Exercise 2.16 (a) to evaluate the Laplace transform of the first hitting times of a point by linear Brownian motion. Now note that the upper bound, part (a) of the lemma, is trivial if $|z| \leq 2\delta$, and otherwise one can bound the exact formula from above by

$$\big(\tfrac{\delta}{|z|}\big)^{\frac{\pi}{\alpha}} \sum_{k=0}^{\infty} \tfrac{4}{(2k+1)\,\pi}\, 2^{-2k\,\frac{\pi}{\alpha}}.$$

The lower bound, part (b) of the lemma, follows from Brownian scaling if $\delta/|z|$ is bounded from below. Otherwise note that, under our assumption on z we have $|\theta + \pi - \xi| \leq \frac{\alpha}{4}$ and thus the sine term corresponding to $k = 0$ is bounded from below by $\sin(\pi/4) > 0$. Thus we get a lower bound of

$$\big(\tfrac{\delta}{|z|}\big)^{\frac{\pi}{\alpha}} \Big(\tfrac{4}{\pi}\, \sin(\pi/4) - \sum_{k=1}^{\infty} \tfrac{4}{(2k+1)\,\pi}\, \big(\tfrac{\delta}{|z|}\big)^{2k\,\frac{\pi}{\alpha}}\Big),$$

and the bracket is bounded from below by a positive constant if $\delta/|z|$ is sufficiently small. ■

An entirely analogous argument also provides the estimates needed for the events imposed *after* the Brownian motion has hit the ball $\mathcal{B}(z, \delta/2)$.

LEMMA 10.40. *There exist constants $C > c > 0$ such that, for every $0 < \delta < \varepsilon$,*

(a) *for all $x, z \in \mathbb{R}^2$ with $|x - z| = \delta/2$,*

$$\mathbb{P}_x\big\{B(0, T_\varepsilon(z)) \subset z + W[\alpha, \xi]\big\} \le C \left(\tfrac{\delta}{\varepsilon}\right)^{\frac{\pi}{\alpha}}.$$

(b) *for all $x, z \in \mathbb{R}^2$ with $|x - z| = \delta/2$ and $x - z \in W[\alpha/2, \xi]$,*

$$\mathbb{P}_x\big\{B(0, T_\varepsilon(z)) \subset z + W[\alpha, \xi]\big\} \ge c \left(\tfrac{\delta}{\varepsilon}\right)^{\frac{\pi}{\alpha}}.$$

We now focus on the *upper bound.* Using the strong Markov property we may combine Lemmas 10.39 (a) and 10.40 (a) to obtain the following lemma.

LEMMA 10.41. *There exists a constant $C_0 > 0$ such that, for any $z \in \mathbb{R}^2$,*

$$\mathbb{P}\big\{z \text{ is a } (\delta, \varepsilon)\text{-approximate cone point } \big\} \le C_0\, |z|^{-\frac{\pi}{\alpha}}\, \varepsilon^{-\frac{\pi}{\alpha}}\, \delta^{\frac{2\pi}{\alpha}}.$$

Proof. By the strong Markov property applied at the stopping time $T_{\delta/2}(z)$ we get

$$\begin{aligned}
&\mathbb{P}\big\{z \text{ is a } (\delta, \varepsilon)\text{-approximate cone point } \big\} \\
&\quad \le \mathbb{E}\big[\mathbb{1}\{B(0, T_\delta(z)) \subset z + W[\alpha, \xi]\}\, \mathbb{P}_{B(T_{\delta/2}(z))}\,\{B(0, S^{(0)}_\varepsilon(z)) \subset z + W[\alpha, \xi]\}\big] \\
&\quad \le C^2 \left(\tfrac{\delta}{|z|}\right)^{\frac{\pi}{\alpha}} \left(\tfrac{\delta}{\varepsilon}\right)^{\frac{\pi}{\alpha}},
\end{aligned}$$

where we have used Lemmas 10.39 (a) and 10.40 (a). The result follows with $C_0 := C^2$. ■

Let $M(\alpha, \xi, \varepsilon)$ be the set of all points in the plane, which are (δ, ε)-approximate cone points for all $\delta > 0$. Obviously $z \in M(\alpha, \xi, \varepsilon)$ if and only if there exist $t > 0$ such that $z = B(t)$ and

$$B(0, t) \subset z + W[\alpha, \xi], \qquad \text{and} \qquad B(t, S^{(t)}_\varepsilon(z)) \subset z + W[\alpha, \xi]\,,$$

where

$$S^{(t)}_\varepsilon(z) := \inf\big\{s > t \colon B(s) \notin \mathcal{B}(z, \varepsilon)\big\}.$$

LEMMA 10.42. *Almost surely,*

- *if $\alpha \in (0, \pi)$ then $M(\alpha, \xi, \varepsilon) = \emptyset$,*
- *if $\alpha \in [\pi, 2\pi)$ then $\dim M(\alpha, \xi, \varepsilon) \le 2 - \frac{2\pi}{\alpha}$.*

Proof. Take a compact cube Cube of unit sidelength not containing the origin. It suffices to show that $M(\alpha, \xi, \varepsilon) \cap \mathsf{Cube} = \emptyset$ if $\alpha \in (0, \pi)$ and $\dim M(\alpha, \xi, \varepsilon) \cap \mathsf{Cube} \le 2 - \frac{2\pi}{\alpha}$ if $\alpha \in (\pi, 2\pi)$. Given a dyadic subcube $D \in \mathfrak{D}_k$ of sidelength 2^{-k} let $D^* \supset D$ be a concentric ball around D with radius $(1 + \sqrt{2})2^{-k}$. Define the *focal point* $x = x(D)$ of D to be

- if $\alpha < \pi$ the tip of the cone $x + W[\alpha, \xi]$ whose boundary halflines are tangent to D^*,
- if $\alpha > \pi$ the tip of the cone whose dual has boundary halflines tangent to D^*.

The following properties are easy to check:

- for every $y \in D$ we have $y + W[\alpha, \xi] \subset x + W[\alpha, \xi]$,

- for some constant $C_1 > 0$ depending only on α, and every $y \in D$ we have

$$\mathcal{B}(y, 2^{-k}) \subset \mathcal{B}(x, C_1 2^{-k}) \quad \text{ and } \mathcal{B}(y, \tfrac{1}{2}\, 2^{-k}) \subset \mathcal{B}(x, C_1 \tfrac{1}{2}\, 2^{-k}),$$

- for some $k_0 \in \mathbb{N}$ depending only on α and ε, every $y \in D$ and $k \geq k_0$,

$$\mathcal{B}(y, \varepsilon) \supset \mathcal{B}(x, \varepsilon/2).$$

This implies that, if $k \geq k_0$ and the cube $D \in \mathfrak{D}_k$ contains a $(2^{-k}, \varepsilon)$-approximate cone point, then its focal point x is a $(C_1 2^{-k}, \varepsilon/2)$-approximate cone point. Hence, by Lemma 10.41, for a constant $C_2 > 0$,

$$\mathbb{P}\big\{D \text{ contains a } (2^{-k}, \varepsilon)\text{-approximate cone point } \big\} \leq C_2\, |x(D)|^{-\frac{\pi}{\alpha}}\, \tilde{\varepsilon}^{-\frac{\pi}{\alpha}}\, 2^{-k\frac{2\pi}{\alpha}}.$$

Note that, given Cube and $\varepsilon > 0$ we can find $k_1 \geq k_0$ such that $|x(D)|$ is bounded away from zero over all $D \in \mathfrak{D}_k$ and $k \geq k_1$. Hence we obtain $C_3 > 0$ such that, for all $k \geq k_1$,

$$\mathbb{P}\big\{D \text{ contains a } (2^{-k}, \varepsilon)\text{-approximate cone point } \big\} \leq C_3\, 2^{-k\frac{2\pi}{\alpha}}.$$

Then, if $\alpha \in (0, \pi)$,

$$\begin{aligned} \mathbb{P}\{M(\alpha, \xi, \varepsilon) \neq \emptyset\} &\leq \sum_{D \in \mathfrak{D}_k} \mathbb{P}\big\{D \text{ contains a } (2^{-k}, \varepsilon)\text{-approximate cone point } \big\} \\ &\leq C_3\, 2^{2k}\, 2^{-k\frac{2\pi}{\alpha}} \quad \stackrel{k\to\infty}{\longrightarrow} \quad 0\,, \end{aligned}$$

proving part (a). Moreover, if $\alpha \in (\pi, 2\pi)$ and $k \geq k_1$, we may cover $M(\alpha, \xi, \varepsilon) \cap \mathsf{Cube}$ by the collection of cubes $D \in \mathfrak{D}_k$ which contain a $(2^{-k}, \varepsilon)$-approximate cone point. Then, for any $\gamma > 2 - \frac{2\pi}{\alpha}$ the expected γ-value of this covering is

$$\begin{aligned} &\mathbb{E} \sum_{D \in \mathfrak{D}_k} 2^{-k\gamma + \frac{3}{2}\gamma} \mathbb{1}\{D \text{ contains a } (2^{-k}, \varepsilon)\text{-approximate cone point } \} \\ &\qquad \leq 2^{\frac{3}{2}\gamma} \sum_{D \in \mathfrak{D}_k} 2^{-k\gamma}\, \mathbb{P}\big\{D \text{ contains a } (2^{-k}, \varepsilon)\text{-approximate cone point } \big\} \\ &\qquad \leq C_3\, 2^{k(2 - \frac{2\pi}{\alpha} - \gamma)} \quad \stackrel{k\to\infty}{\longrightarrow} \quad 0\,, \end{aligned}$$

and this proves that, almost surely, $\dim M(\alpha, \xi, \varepsilon) \leq \gamma$. ■

Proof of the upper bound in Theorem 10.38. Suppose $\delta > 0$ is arbitrary and $z \in \mathbb{R}^2$ is an α-cone point. Then there exist a rational number $q \in [0, 1)$, a rational direction $\xi \in [0, 2\pi)$, and a rational $\varepsilon > 0$, such that $z = B(t)$ for some $t \in (q, 1)$ and

$$B(q, t) \subset z + W[\alpha + \delta, \xi], \quad \text{ and } \quad B(t, S_\varepsilon^{(t)}(z)) \subset z + W[\alpha + \delta, \xi]\,.$$

By Lemma 10.42 for every fixed choice of rational parameters this set is empty almost surely if $\alpha + \delta < \pi$. For any $\alpha < \pi$ we can pick $\delta > 0$ with $\alpha + \delta < \pi$ and hence there are no α-cone points almost surely. Similarly, if $\alpha \geq \pi$ we can use Lemma 10.42 and the countable stability of Hausdorff dimension to obtain an almost sure upper bound of $2 - 2\pi/(\alpha + \delta)$ for the set of α-cone points. The result follows as $\delta > 0$ was arbitrary. ■

We now establish the framework to prove the lower bound in Theorem 10.38. Again we fix $x_0 \in \mathbb{R}^d$ and a cube $\mathsf{Cube} = x_0 + [0,1)^d$. Recall the definition of the collection $\mathfrak{D}_k$ of dyadic subcubes of sidelength 2^{-k} and let $\mathfrak{D} = \bigcup_{k=1}^{\infty} \mathfrak{D}_k$. Suppose that $\{Z(I) : I \in \mathfrak{D}\}$ is a collection of random variables each taking values in $\{0,1\}$. With this collection we associate the random set

$$A := \bigcap_{k=1}^{\infty} \bigcup_{\substack{I \in \mathfrak{D}_k \\ Z(I)=1}} I \,.$$

Theorem 10.43. *Suppose that the random variables $\{Z(I) : I \in \mathfrak{D}\}$ satisfy the monotonicity condition*

$$I \subset J \text{ and } Z(I) = 1 \qquad \Rightarrow \qquad Z(J) = 1.$$

Assume that, for some positive constants γ, c_1 and C_1,

(i) $c_1 \, |I|^{\gamma} \le \mathbb{E} Z(I) \le C_1 \, |I|^{\gamma}$ *for all* $I \in \mathfrak{D}$,

(ii) $\mathbb{E}\big[Z(I)Z(J)\big] \le C_1 \, |I|^{2\gamma} \operatorname{dist}(I,J)^{-\gamma}$ *for all* $I, J \in \mathfrak{D}_k$ *and* $k \ge 1$.

Then, for $\lambda > \gamma$ and $\Lambda \subset \mathsf{Cube}$ closed with $\mathcal{H}^{\lambda}(\Lambda) > 0$, there exists a $p > 0$, such that

$$\mathbb{P}\big\{ \dim(A \cap \Lambda) \ge \lambda - \gamma \big\} \ge p \,.$$

Remark 10.44. Though formally, if the monotonicity condition holds, A is a limsup fractal, the monotonicity establishes a strong dependence of the random variables $\{Z(I) : I \in \mathfrak{D}_k\}$ which in general invalidates the second assumption of Theorem 10.28. We therefore need a result which deals specifically with this situation. $\diamond$

We prepare the proof with a little lemma, based on Fubini's theorem.

Lemma 10.45. *Suppose ν is a probability measure on $\mathbb{R}^d$ such that $\nu\mathcal{B}(x,r) \le Cr^{\lambda}$ for all $x \in \mathbb{R}^d, r > 0$. Then, for all $0 < \beta < \lambda$ there exists $C_2 > 0$ such that,*

$$\int_{\mathcal{B}(x,r)} |x-y|^{-\beta} \, \nu(dy) \le C_2 \, r^{\lambda-\beta}, \qquad \textit{for every } x \in \mathbb{R}^d \textit{ and } r > 0.$$

This implies, in particular, that

$$\iint |x-y|^{-\beta} \, d\nu(x) \, d\nu(y) < \infty \,.$$

Proof. Fubini's theorem gives

$$\begin{aligned} \int_{\mathcal{B}(x,r)} |x-y|^{-\beta} \, \nu(dy) &= \int_0^{\infty} \nu\big\{ y \in \mathcal{B}(x,r) \colon |x-y|^{-\beta} > s \big\} \, ds \\ &\le \int_{r^{-\beta}}^{\infty} \nu\mathcal{B}(x, s^{-1/\beta}) \, ds + C \, r^{\lambda-\beta} \le C \int_{r^{-\beta}}^{\infty} s^{-\lambda/\beta} \, ds + C \, r^{\lambda-\beta}, \end{aligned}$$

which implies the first statement. Moreover,

$$\iint |x-y|^{-\beta} \, d\nu(x) \, d\nu(y) \le \int d\nu(x) \int_{\mathcal{B}(x,1)} |x-y|^{-\beta} \, d\nu(y) + 1 \le C_2 + 1 < \infty \,.$$

■

Proof of Theorem 10.43. We show that there exists $p > 0$ such that, for every $0 < \beta < \lambda - \gamma$, with probability at least p, there exists a positive measure μ on $\Lambda \cap A$ such that its β-energy $I_\beta(\mu)$ is finite. This implies $\dim(A \cap \Lambda) \geq \beta$ by the energy method, see Theorem 4.27.

To begin with, given $\Lambda \subset \mathsf{Cube}$ with $\mathcal{H}^\lambda(\Lambda) > 0$, we use Frostman's lemma to find a Borel probability measure ν on Λ and positive constants $0 < c < C$ such that $\nu(D) \leq C|D|^\lambda$ for all Borel sets $D \subset \mathbb{R}^d$. Writing

$$A_n := \bigcup_{\substack{I \in \mathfrak{D}_n \\ Z(I)=1}} I\,,$$

we define μ_n to be the measure supported on Λ given by

$$\mu_n(B) = 2^{n\gamma}\,\nu(B \cap A_n) \qquad \text{for any Borel set } B \subset \mathbb{R}^d.$$

Then, using (i), we get

$$\mathbb{E}\big[\mu_n(A_n)\big] \geq 2^{n\gamma} \sum_{I \in \mathfrak{D}_n} \nu(I)\,\mathbb{E}Z(I) \geq c_1 \sum_{I \in \mathfrak{D}_n} \nu(I) = c_1\,.$$

Moreover, using (ii), we obtain

$$\begin{aligned}
\mathbb{E}\big[\mu_n(A_n)^2\big] &\leq 2^{2n\gamma} \sum_{I \in \mathfrak{D}_n} \sum_{J \in \mathfrak{D}_n} \mathbb{E}[Z(I)Z(J)]\nu(I)\nu(J) \\
&\leq C_1 \sum_{I \in \mathfrak{D}_n} \sum_{\substack{J \in \mathfrak{D}_n \\ \operatorname{dist}(I,J)>0}} \operatorname{dist}(I,J)^{-\gamma}\nu(I)\nu(J) + C(3\sqrt{d})^\lambda\, 2^{2n\gamma}\, 2^{-n\lambda} \sum_{I \in \mathfrak{D}_n} \mathbb{E}[Z(I)]\nu(I) \\
&\leq C_1 3^\gamma \iint |x-y|^{-\gamma}\, d\nu(x)\, d\nu(y) + C_1\, C\, (3\sqrt{d})^\lambda =: C_3 < \infty,
\end{aligned}$$

where finiteness of C_3 follows from the second statement of Lemma 10.45.

We now show that, for every $\beta < \lambda - \gamma$ we find $k(\beta)$ such that $\mathbb{E}I_\beta(\mu_n) \leq k(\beta)$. Indeed,

$$\begin{aligned}
\mathbb{E}I_\beta(\mu_n) &= 2^{2n\gamma} \sum_{I,J \in \mathfrak{D}_n} \mathbb{E}[Z(I)Z(J)] \int_I d\nu(x) \int_J d\nu(y)\, |x-y|^{-\beta} \\
&\leq C_1 \sum_{I \in \mathfrak{D}_n} \sum_{\substack{J \in \mathfrak{D}_n \\ \operatorname{dist}(I,J)>0}} \operatorname{dist}(I,J)^{-\gamma} \int_I d\nu(x) \int_J d\nu(y)\, |x-y|^{-\beta} \\
&\qquad + C_1\, 2^{n\gamma} \sum_{I \in \mathfrak{D}_n} \sum_{\substack{J \in \mathfrak{D}_n \\ \operatorname{dist}(I,J)=0}} \int_I d\nu(x) \int_J d\nu(y)\, |x-y|^{-\beta}\,.
\end{aligned}$$

For the first summand, we use that $\operatorname{dist}(I,J)^{-\gamma} \leq (3\sqrt{d})^\gamma\, |x-y|^{-\gamma}$ whenever $x \in I$ and $y \in J$, and infer boundedness from the second statement of Lemma 10.45. For the second summand the first statement of Lemma 10.45 gives a bound of

$$C_1 C_2\, 2^{n\gamma}\, (3\sqrt{d}2^{-n})^{\lambda-\beta} \sum_{I \in \mathfrak{D}_n} \nu(I) \leq C_1\, C_2\, (3\sqrt{d})^{\lambda-\beta}\,.$$

Hence, $\mathbb{E}I_\beta(\mu_n)$ is bounded uniformly in n, as claimed. We therefore find $\ell(\beta)>0$ such that

$$\mathbb{P}\big\{I_\beta(\mu_n)\ge \ell(\beta)\big\}\le \frac{k(\beta)}{\ell(\beta)}\le \frac{c_1}{8C_3}\,.$$

Now, by the Paley-Zygmund inequality, see Lemma 3.22,

$$\mathbb{P}\big\{\mu_n(A_n)>\tfrac{c_1}{2}\big\}\ge \mathbb{P}\big\{\mu_n(A_n)>\tfrac12\,\mathbb{E}[\mu_n(A_n)]\big\}\ge \frac14\,\frac{\mathbb{E}[\mu_n(A_n)]^2}{\mathbb{E}[\mu_n(A_n)^2]}\ge \frac{c_1}{4C_3}\,.$$

Hence we obtain that

$$\mathbb{P}\big\{\mu_n(A_n)>\tfrac{c_1}{2},\, I_\beta(\mu_n)<\ell(\beta)\big\}\ge p:=\frac{c_1}{8C_3}\,.$$

Using Fatou's lemma we infer that

$$\mathbb{P}\big\{\mu_n(A_n)>\tfrac{c_1}{2},\, I_\beta(\mu_n)<\ell(\beta) \text{ infinitely often }\big\}\ge p\,.$$

On this event we can pick a subsequence along which μ_n converges to some measure μ. Then μ is supported by A and $\mu(\mathsf{Cube})\ge \liminf \mu_n(\mathsf{Cube})=\liminf \mu_n(A_n)\ge c_1/2$.

Finally, for each $\varepsilon>0$, where the limit is taken along the chosen subsequence,

$$\iint_{|x-y|>\varepsilon}|x-y|^{-\beta}d\mu(x)\,d\mu(y)=\lim\iint_{|x-y|>\varepsilon}|x-y|^{-\beta}d\mu_n(x)\,d\mu_n(y)=I_\beta(\mu_n)\le\ell(\beta),$$

and for $\varepsilon\downarrow 0$ we get $I_\beta(\mu)\le \ell(\beta)$. ■

We now use Theorem 10.43 to give a *lower bound* for the dimension of the set of cone points. Fix $\alpha\in(\pi,2\pi)$ and a unit cube

$$\mathsf{Cube}=x_0+[0,1]^2\subset W[\alpha/2,0]\,.$$

Choose a large radius $R>2$ such that $\mathsf{Cube}\subset\mathcal{B}(0,R/2)$ and define

$$r_k:=R-\sum_{j=1}^k 2^{-j}>R/2\,.$$

Given a cube $I\in\mathfrak{C}_k(x_0)$ we denote by z its centre and let $Z(I)=1$ if z is a $(2^{-k},r_k)$-approximate cone point with direction $\xi=\pi$, i.e. if

$$B(0,T_{2^{-k}}(z))\subset z+W[\alpha,\pi],\qquad\text{and}\qquad B(T_{2^{-k-1}}(z),S_{2^{-k},r_k}(z))\subset z+W[\alpha,\pi]\,,$$

and otherwise let $Z(I)=0$. By our choice of the sequence (r_k) we have

$$I\subset J \text{ and } Z(I)=1\qquad\Rightarrow\qquad Z(J)=1.$$

Lemma 10.46. *There are constants $0<c_1<C_1<\infty$ such that, for any cube $I\in\mathfrak{C}$, we have*

$$c_1\,|I|^{\frac{2\pi}{\alpha}}\le\mathbb{P}\big\{Z(I)=1\big\}\le C_1\,|I|^{\frac{2\pi}{\alpha}}\,.$$

Proof. The upper bound is immediate from Lemma 10.41. For the lower bound we use that, for any $z \in \mathsf{Cube}$ and $\delta > 0$,

$$\inf_{|x-z|=\delta} \mathbb{P}_x\big\{B(T_{\delta/2}(z)) \in z + W[\alpha/2,\pi]\big\} = \inf_{|x|=1} \mathbb{P}_x\big\{B(T_{1/2}(0)) \in W[\alpha/2,\pi]\big\} =: c_0 > 0\,,$$

and hence, if z is the centre of $I \in \mathfrak{C}_k$ and $\delta = 2^{-k}$, using Lemmas 10.39 (b) and 10.40 (b),

$$\begin{aligned}\mathbb{P}\big\{Z(I) = 1\big\} \geq \mathbb{E}\Big[&\mathbb{1}\{B(0,T_\delta(z)) \subset z + W[\alpha,\pi]\}\, \mathbb{E}_{B(T_\delta(z))}\big[\mathbb{1}\{B(T_{\delta/2}(z)) \in z + W[\alpha/2,\pi]\big\}\\ &\times \mathbb{P}_{B(T_{\delta/2}(z))}\{B(0,S^{(0)}_{r_k}(z)) \subset z + W[\alpha,\pi]\}\big]\Big]\\ \geq\ & c_0\, c^2\, \delta^{\frac{2\pi}{\alpha}}\, \big(R|z|\big)^{-\frac{\pi}{\alpha}},\end{aligned}$$

which gives the desired statement, as $|z|$ is bounded away from infinity. ∎

Lemma 10.47. *There is a constant $0 < C_1 < \infty$ such that, for any cubes $I, J \in \mathfrak{C}_k$, $k \geq 1$, we have*

$$\mathbb{E}\big[Z(I)Z(J)\big] \leq C_1\, |I|^{\frac{4\pi}{\alpha}}\ \mathrm{dist}(I,J)^{-\frac{2\pi}{\alpha}}.$$

Proof. Let z_I, z_J be the centres of I, resp. J, and abbreviate $\eta := |z_I - z_J|$ and $\delta := 2^{-k}$. Then, for $\eta > 2\delta$, using the strong Markov property and Lemmas 10.39 (a) and 10.40 (a),

$$\begin{aligned}&\mathbb{E}\big[Z(I)\, Z(J)\, \mathbb{1}\{T_{\delta/2}(z_I) < T_{\delta/2}(z_J)\big]\\ &\leq \mathbb{E}\Big[\mathbb{1}\{B(0,T_\delta(z_I)) \subset z_I + W[\alpha,\pi]\}\, \mathbb{E}_{B(T_{\delta/2}(z_I))}\big[\mathbb{1}\{B(0,S^{(0)}_{\eta/2}(z_I)) \subset z_I + W[\alpha,\pi]\}\\ &\times \mathbb{E}_{B(T_{\eta/2}(z_J))}\big[\mathbb{1}\{B(0,T_\delta(z_J)) \in z_J + W[\alpha,\pi]\big\}\, \mathbb{P}_{B(T_{\delta/2}(z_J))}\{B(0,S^{(0)}_{r_k}(z_J)) \subset z_J + W[\alpha,\pi]\}\big]\big]\Big]\\ &\leq C^4\, \big(\tfrac{\delta}{|z_I|}\big)^{\frac{\pi}{\alpha}}\, \big(\tfrac{\delta}{\eta}\big)^{\frac{2\pi}{\alpha}}\, \big(\tfrac{2\delta}{R}\big)^{\frac{\pi}{\alpha}} \leq (C_1/2)\, |I|^{\frac{4\pi}{\alpha}}\ \mathrm{dist}(I,J)^{-\frac{2\pi}{\alpha}},\end{aligned}$$

where we define $C_1 := 2^{\frac{\pi}{\alpha}+1}(C^4 \vee 1)$. Suppose now that $\eta \leq 2\delta$. Then, by a simpler argument,

$$\begin{aligned}&\mathbb{E}\big[Z(I)\, Z(J)\, \mathbb{1}\{T_{\delta/2}(z_I) < T_{\delta/2}(z_J)\big]\\ &\quad\leq \mathbb{E}\big[\mathbb{1}\{B(0,T_\delta(z_I)) \subset z_I + W[\alpha,\pi]\}\ \mathbb{P}_{B(T_{\delta/2}(z_J))}\{B(0,S^{(0)}_{r_k}(z_J)) \subset z_J + W[\alpha,\pi]\}\big]\\ &\quad\leq C^2\, \big(\tfrac{\delta}{|z_I|}\big)^{\frac{\pi}{\alpha}}\, \big(\tfrac{2\delta}{R}\big)^{\frac{\pi}{\alpha}} \leq (C_1/2)\, |I|^{\frac{4\pi}{\alpha}}\ \mathrm{dist}(I,J)^{-\frac{2\pi}{\alpha}}.\end{aligned}$$

Exchanging the rôles of I and J gives the corresponding estimate

$$\mathbb{E}[Z(I)Z(J)\mathbb{1}\{T_{\delta/2}(z_I) > T_{\delta/2}(z_J)] \leq (C_1/2)\, |I|^{\frac{4\pi}{\alpha}}\ \mathrm{dist}(I,J)^{-\frac{2\pi}{\alpha}},$$

and the proof is completed by adding the two estimates. ∎

Proof of the lower bound in Theorem 10.38. The set A which we obtain from our choice of $\{Z(I) \colon I \in \mathfrak{C}\}$ is contained in the set

$$\tilde{A} := \big\{B(t) : t > 0 \text{ and } B(0,S^{(t)}_{R/2}(B(t))) \subset B(t) + W[\alpha,\pi]\big\}\,.$$

Therefore, by Theorem 10.43, we have $\dim \tilde{A} \geq 2 - 2\pi/\alpha$ with positive probability. Given any $0 < \delta < 1/2$, we define a sequence $\tau_1^{(\delta)} \leq \tau_2^{(\delta)} \leq \dots$ of stopping times by $\tau_1^{(\delta)} = 0$ and, for $k \geq 1$,

$$\tau_k^{(\delta)} := S_{\delta R}^{(\tau_{k-1}^{(\delta)})}\big(B(\tau_{k-1}^{(\delta)})\big).$$

Denoting

$$A_k^{(\delta)} := \big\{B(t) : \tau_{k-1}^{(\delta)} \leq t \leq \tau_k^{(\delta)} \text{ and } B(\tau_{k-1}^{(\delta)}, S_{R/2}^{(t)}(B(t))) \subset B(t) + W[\alpha, \pi]\big\}$$

we have that

$$\tilde{A} \subset \bigcup_{k=1}^{\infty} A_k^{(\delta)}.$$

Now fix $\beta < 2 - 2\pi/\alpha$. The events $\{\dim A_k^{(\delta)} \geq \beta\}$ all have the same probability, which cannot be zero as this would contradict the lower bound on the dimension of $\tilde{A}$. In particular, there exists $p_R^{(\delta)} > 0$ such that

$$\mathbb{P}\big\{ \dim \big\{B(t) \colon 0 \leq t \leq S_{\delta R}^{(0)}(0) \text{ and } B(0, S_{R/2}^{(0)}(0)) \subset B(t) + W[\alpha, \pi]\big\} \geq \beta \big\} \geq p_R^{(\delta)}.$$

By scaling we get that $p_R^{(\delta)}$ does not depend on R. Hence, by Blumenthal's zero-one law, we have that $p_R^{(\delta)} = 1$ for all $\delta > 0, R > 0$. Letting $\varepsilon \downarrow 0$ we get, almost surely,

$$\dim \big\{B(t) \colon t \leq S_{\delta R}^{(0)}(0),\ B(0, S_{R/2}^{(0)}(0)) \subset B(t) + W[\alpha, \pi]\big\} \geq 2 - \tfrac{2\pi}{\alpha}$$

for every $\delta > 0$, $R > 0$. On this event we may choose first $R > 1$ and then $\delta > 0$ such that

$$S_{R/2}^{(0)}(0) > 1 \qquad \text{and} \qquad S_{\delta R}^{(0)}(0) < 1,$$

and get that the Hausdorff dimension of the set of α-cone points is at least $2 - \frac{2\pi}{\alpha}$. ■

A surprising consequence of the non-existence of cone points for angles smaller then π is that the convex hull of the planar Brownian curve is a fairly smooth set.

Theorem 10.48 (Adelman (1982)). *Almost surely, the convex hull of* $\{B(s) \colon 0 \leq s \leq 1\}$ *has a differentiable boundary.*

Proof. A compact, convex subset $H \subset \mathbb{R}^2$ is said to have a *corner* at $x \in \partial H$ if there exists a cone with vertex x and opening angle $\alpha < \pi$ which contains H. If H does not have corners, the supporting hyperplanes are unique at each point $x \in \partial H$ and thus ∂H is a differentiable boundary.

So all we have to show is that the convex hull H of $\{B(s) : 0 \leq s \leq 1\}$ has no corners. Clearly, by Spitzer's theorem, $B(0)$ and $B(1)$ are no corners almost surely. Suppose any other point $x \in \partial H$ is a corner, then obviously it is contained in the path, and therefore it is an α-cone point for some $\alpha > \pi$. By Theorem 10.38, almost surely, such points do not exist and this is a contradiction. ■

Exercises

Exercise 10.1. Show that, for every metric space E,

$$\dim_P E = \inf\{s : \mathcal{P}^s(E) < \infty\} = \sup\{s : \mathcal{P}^s(E) > 0\} = \sup\{s : \mathcal{P}^s(E) = \infty\}.$$

Exercise 10.2 (∗). Show that, for every metric space E, we have

$$\dim_P E \geq \dim E.$$

Exercise 10.3. Let $\{m_k : k \geq 1\}$ be a rapidly increasing sequence of positive integers such that

$$\lim_{k\to\infty} \frac{m_k}{m_{k+1}} = 0.$$

Define two subsets of $[0,1]$ by

$$E = \Big\{ \sum_{i=1}^{\infty} \frac{x_i}{2^i} : x_i \in \{0,1\} \text{ and } x_i = 0 \text{ if } m_k + 1 \leq i \leq m_{k+1} \text{ for some even } k \Big\}$$

and

$$F = \Big\{ \sum_{i=1}^{\infty} \frac{x_i}{2^i} : x_i \in \{0,1\} \text{ and } x_i = 0 \text{ if } m_k + 1 \leq i \leq m_{k+1} \text{ for some odd } k \Big\}.$$

Show that

(1) $\dim E = \underline{\dim}_M E = 0$ and $\dim F = \underline{\dim}_M F = 0$,

(2) $\dim_P E = \overline{\dim}_M E = 1$ and $\dim_P F = \overline{\dim}_M F = 1$,

(3) $\dim(E \times F) \geq 1$.

Exercise 10.4. Show that, almost surely,

- $\dim_P \mathsf{Range} = 2$, for Brownian motion in $d \geq 2$,
- $\dim_P \mathsf{Graph} = \frac{3}{2}$, for Brownian motion in $d = 1$,
- $\dim_P \mathsf{Zero} = \frac{1}{2}$, for Brownian motion in $d = 1$.

Exercise 10.5. Show that, for every $a \in [0, \sqrt{2}]$, we have almost surely,

$$\dim_P \Big\{ t \in [0,1] \, : \, \limsup_{h \downarrow 0} \frac{|B(t+h) - B(t)|}{\sqrt{h \log(1/h)}} \geq a \Big\} = 1.$$

Hint. This can be done directly, but it can also be derived from more general ideas, as formulated for example in Exercise 10.9.

Exercise 10.6. Show that

$$\liminf_{h \downarrow 0} \sup_{t \in E} \frac{|B(t+h) - B(t)|}{\sqrt{2h \log(1/h)}} = \sqrt{\underline{\dim}_M(E)}.$$

Exercise 10.7 (∗). Use Theorem 10.43 to prove once more that the zero set Zero of linear Brownian motion has Hausdorff dimension $\frac{1}{2}$ almost surely.

Exercise 10.8. Show that, if

$$\limsup_{n \uparrow \infty} \frac{\log p_n}{n \, \log 2} \leq -\gamma, \qquad \text{for some } \gamma > 0,$$

then, for any compact $E \subset [0,1]$ with $\dim_P(E) < \gamma$, we have

$$\mathbb{P}\big\{ A \cap E \neq \emptyset \big\} = 0.$$

Note that no independence assumption is needed for this statement.

Exercise 10.9 (∗).

(a) Suppose A is a discrete limsup fractal associated to random variables $\{Z(I) \, : \, I \in \mathfrak{C}_k, k \geq 1\}$ satisfying the conditions of Theorem 10.28. Then, if $\dim_P(E) > \gamma$, we have almost surely, $\dim_P(A \cap E) = \dim_P(E)$.

(b) Show that, if $\dim_P(E) > a^2$, then almost surely

$$\dim_P(F(a) \cap E) = \dim_P(E),$$

where $F(a)$ is the set of a-fast points.

Exercise 10.10 (∗). Give a proof of Lemma 10.40 (a) based on Theorem 7.24.

Exercise 10.11. Suppose $K \subset \mathbb{R}^2$ is a compact set and $x \in \mathbb{R}^2 \setminus K$ a point outside the set. Imagine K as a solid body, and x as the position of an observer. This observer can only see a part of the body, which can be formally described as

$$K(x) = \big\{ y \in K \,:\, [x,y] \cap K = \{y\} \big\},$$

where $[x,y]$ denotes the compact line segment connecting x and y. It is natural to ask for the Hausdorff dimension of the visible part of a set K. Assuming that $\dim K \geq 1$, an unresolved conjecture in geometric measure theory claims that, for Lebesgue-almost every $x \notin K$, the Hausdorff dimension of $K(x)$ is one.

Show that this conjecture holds for the path of planar Brownian motion, $K = B[0,1]$, in other words, almost surely, for Lebesgue almost every $x \in \mathbb{R}^2$, the Hausdorff dimension of the visible part $B[0,1](x)$ is one.

Exercise 10.12. Let $\{B(t) \colon t \geq 0\}$ be a planar Brownian motion and $\alpha \in [\pi, 2\pi)$. Show that, almost surely, no double points are α-cone points.

Exercise 10.13 (∗). Let $\{B(t) \colon t \geq 0\}$ be a planar Brownian motion and $\alpha \in (0, \pi]$. A point $x = B(t)$, $0 < t < 1$, is a *one-sided* α-cone point if there exists $\xi \in [0, 2\pi)$ such that

$$B(0,t) \subset x + W[\alpha, \xi]\,.$$

(a) Show that for $\alpha \leq \frac{\pi}{2}$, almost surely, there are no one-sided α-cone points.
(b) Show that for $\alpha \in (\frac{\pi}{2}, \pi]$, almost surely, the set of one-sided α-cone points has Hausdorff dimension $2 - \frac{\pi}{\alpha}$.

Notes and Comments

The paper [**OT74**] by Orey and Taylor is a seminal work in the study of dimension spectra for exceptional points of Brownian motion. It contains a proof of Theorem 10.3 using the mass distribution principle and direct construction of the Frostman measure. This approach can be extended to other limsup fractals, but this method requires quite strong independence assumptions which make this method difficult in many more general situations. In [**OT74**] the question how often on a Brownian path the law of the iterated logarithm fails is also answered in the sense that, for $\theta > 1$, almost surely, the set

$$\Big\{ t > 0 \colon \limsup_{h \downarrow 0} \frac{B(t+h) - B(t)}{\sqrt{2h \log\log(1/h)}} \geq \theta \Big\}$$

has zero or infinite Hausdorff measure for the gauge function $\phi(r) = r \log(1/r)^{\gamma}$ depending whether $\gamma < \theta^2 - 1$ or $\gamma > \theta^2 - 1$.

Our proof of Theorem 10.3 is based on estimates of energy integrals. This method was used by Hu and Taylor [**HT97**] and Shieh and Taylor [**ST99**], and our exposition follows closely [**DPRZ00**]. In the latter paper an interesting class of exceptional times for the Brownian motion is treated, the *thick times* of Brownian motions in dimension $d \geq 3$. For any time $t \in (0,1)$ we let $U(t,\varepsilon) = \mathcal{L}\{s \in (0,1) \,:\, |B(s) - B(t)| \leq \varepsilon\}$ the set of times where the Brownian is up to ε near to its position at time t. It is shown that, for all $0 \leq a \leq \frac{16}{\pi^2}$, almost surely,

$$\dim \Big\{ t \in [0,1] \,:\, \limsup_{\varepsilon \downarrow 0} \tfrac{U(t,\varepsilon)}{\varepsilon^2 \log(1/\varepsilon)} \geq a \Big\} = 1 - a\, \tfrac{\pi^2}{16}\,.$$

This paper should be very accessible to anyone who followed the arguments of Section 10.1. The method of [**DPRZ00**] can be extended to limsup fractals with somewhat weaker independence properties and also extends to the study of dimension spectra with strict equality.

The third way to prove Theorem 10.3 is the method of stochastic codimension explored in Section 10.2. An early reference for this method is Taylor [**Ta66**] who suggested to use the range of stable processes as test sets, and made use of the potential theory of stable processes to obtain lower bounds for Hausdorff dimension. This class of test sets is not big enough for all problems: the Hausdorff dimension of a stable processes is bounded above by its index, hence cannot exceed 2, and therefore these test sets can only test dimensions in the range $[d-2, d]$. A possible remedy is to pass to multiparameter processes, see the recent book of Khoshnevisan [**Kh02**] for a survey. Later, initiated by seminal papers of Hawkes [**Ha81**] and R. Lyons [**Ly90**], it was discovered that percolation limit sets are a very suitable class of test functions, see [**KPX00**]. Our exposition closely follows the latter reference.

Kaufman [**Ka75**] showed that every compact set $E \subset [0,1]$ with $\dim(E) > a^2$ almost surely contains an a-fast point, but the more precise result involving the packing dimension is due to [**KPX00**]. The concept of packing dimension was introduced surprisingly late by Tricot in [**Tr82**] and in [**TT85**] it was investigated together with the packing measure and applied to the Brownian path by Taylor and Tricot. Lemma 10.18(i) is from [**Tr82**], Lemma 10.18(ii) for trees can be found in [**BP94**], see Proposition 4.2(b), the general version given is in Falconer and Howroyd [**FH96**] and in Mattila and Mauldin [**MM97**].

Several people contributed to the investigation of slow points, for example Dvoretzky [**Dv63**], Kahane [**Ka76**], Davis [**Da83**], Greenwood and Perkins [**GP83**] and Perkins [**Pe83**]. There are a number of variants, for example one can allow $h < 0$ in (3.1) or omit the modulus signs. The Hausdorff dimension of a-slow points is discussed in [**Pe83**], this class of exceptional sets is not tractable with the limsup-methods: note that an exceptional behaviour is required at all small scales. The crucial ingredient, the finiteness criterion for moments of the stopping times $T(r, a)$ is due to Shepp [**Sh67**].

Cone points were discussed by Evans in [**Ev85**], an alternative discussion can be found in Lawler's survey paper [**La99**]. Our argument essentially follows the latter paper. The correlation condition in Theorem 10.43 appears in the strongly related context of quasi-Bernoulli percolation on trees, see Lyons [**Ly92**].

An alternative notion of *global* cone points requires that the entire path of the Brownian motion $\{B(t) \colon t \geq 0\}$ stays inside the cone with tip in the cone point. The same dimension formula holds for this concept. The upper bound follows of course from our consideration of local cone points, and our proof gives the lower bound with positive probability. The difficult part is to show that the lower bound holds with probability one. A solution to this problem is contained in Burdzy and San Martín [**BSM89**], and this technique has also been successfully used in the study of the Brownian frontier [**La96b, BJPP97**].

A discussion of the smoothness of the boundary of the convex hull can be found in Cranston, Hsu and March [**CHM89**], but our Theorem 10.48 is older. The result was stated by Lévy [**Le48**] and was probably first proved by Adelman in 1982, though this does not seem to be published.

It is conjectured in geometric measure theory that for any set of Hausdorff dimension dim $K \geq 1$, for Lebesgue-almost every $x \notin K$, the Hausdorff dimension of the visible part $K(x)$ is one. For upper bounds on the dimension and the state of the art on this conjecture, see [**ON04**]. It is natural to compare this to Makarov's theorem on the support of harmonic measure: if the rays of light were following Brownian paths rather than straight lines, the conjecture would hold by Makarov's theorem, see [**Ma85**].

Appendix I: Hints and solutions for selected exercises

In this section we give hints, solutions or additional references for the exercises marked with either of the symbols $(*)$ or $(**)$ in the main body of the text.

Exercise 1.2. Fix times $0 < t_1 < \ldots < t_n$. Let

$$M := \begin{pmatrix} 1 & 0 & \ldots & 0 \\ -1 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}, \qquad D := \begin{pmatrix} \frac{1}{\sqrt{t_1}} & 0 & \ldots & 0 \\ 0 & \frac{1}{\sqrt{t_2-t_1}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \frac{1}{\sqrt{t_n-t_{n-1}}} \end{pmatrix}.$$

Then, for a Brownian motion $\{B(t) : t \geq 0\}$ with start in x, by definition, the vector

$$X := D\,M\,\big(B(t_1)-x, \ldots, B(t_n)-x\big)^{\mathrm{T}}$$

has independent standard normal entries. As both D and M are nonsingular, the matrix $A := M^{-1}D^{-1}$ is well-defined and, denoting also $b = (x, \ldots, x)$, we have that

$$\big(B(t_1), \ldots, B(t_n)\big)^{\mathrm{T}} = AX + b\,.$$

By definition, this means that $(B(t_1), \ldots, B(t_n))$ is a Gaussian random vector.

Exercise 1.4. Note that $\{X(t)\colon 0 \leq t \leq 1\}$ is a Gaussian process, while the distributions given in (a) determine Gaussian random vectors. Hence it suffices to identify the means and covariances of $(X(t_1), \ldots, X(t_n))$ and compare them with those given in (a). Starting with the mean, on the one hand we obviously have $\mathbb{E}X(t) = x(1-t) + ty$, on the other hand

$$\begin{aligned} &\int z\, \frac{\mathfrak{p}(t,x,z)\mathfrak{p}(1-t,z,y)}{\mathfrak{p}(1,x,y)} \\ &\quad = \frac{1}{\mathfrak{p}(1,x,y)} \int \big(z - x(1-t) - ty\big)\mathfrak{p}(t,x,z)\mathfrak{p}(1-t,z,y)\,dz + x(1-t) + ty, \end{aligned}$$

and the integral can be seen to vanish by completing the square in the exponent of the integrand. To perform the covariance calculation one may assume that $x = y = 0$, which reduces the complexity of expressions significantly, see [**Du96**, (8.5) in Chapter 7] for more details.

Exercise 1.5. $B(t)$ does not oscillate too much between n and $n+1$ if

$$\limsup_{n\to\infty} \frac{1}{n} \big[\max_{n\le t\le n+1} B(t) - B(n) \big] = 0\,.$$

Use the Borel-Cantelli lemma and a suitable estimate for

$$\mathbb{P}\big\{ \max_{0\le t\le 1} B(t) \ge \varepsilon n \big\}$$

to obtain this result.

Exercise 1.6. One has to improve the lower bound, and show that, for every constant $c < \sqrt{2}$, almost surely, there exists $\varepsilon > 0$ such that, for all $0 < h < \varepsilon$, there exists $t \in [0, 1-h]$ with

$$\big|B(t+h) - B(t)\big| \ge c\sqrt{h \log(1/h)}.$$

To this end, given $\delta > 0$, let $c < \sqrt{2} - \delta$ and define, for integers $k, n \ge 0$, the events

$$A_{k,n} = \Big\{ B\big((k+1)e^{-\sqrt{n}}\big) - B\big(ke^{-\sqrt{n}}\big) > c\,\big(\sqrt{n}\,e^{-\sqrt{n}}\big)^{\frac12} \Big\}.$$

Then, using Lemma II.3.1, for any $k \ge 0$,

$$\mathbb{P}(A_{k,n}) = \mathbb{P}\Big\{ B\big(e^{-\sqrt{n}}\big) > c\,\big(\sqrt{n}e^{-\sqrt{n}}\big)^{\frac12} \Big\} = \mathbb{P}\big\{ B(1) > c\,n^{\frac14} \big\} \ge \frac{cn^{\frac14}}{c^2\sqrt{n}+1}\, e^{-c^2\sqrt{n}/2}\,.$$

Therefore, by our assumption on c, and using that $1 - x \le e^{-x}$ for all $x \ge 0$,

$$\sum_{n=0}^{\infty} \mathbb{P}\Big(\bigcap_{k=0}^{\lfloor e^{\sqrt{n}}-1 \rfloor} A_{k,n}^{\mathrm{c}} \Big) \le \sum_{n=0}^{\infty} \big(1 - \mathbb{P}(A_{0,n})\big)^{e^{\sqrt{n}}-1} \le \sum_{n=0}^{\infty} \exp\big(-(e^{\sqrt{n}}-1)\mathbb{P}(A_{0,n})\big) < \infty\,.$$

From the Borel-Cantelli lemma we thus obtain that, almost surely, there exists $n_0 \in \mathbb{N}$ such that, for all $n \ge n_0$, there exists $t \in [0, 1-e^{-\sqrt{n}}]$ of the form $t = ke^{-\sqrt{n}}$ such that

$$\big|B\big(t + e^{-\sqrt{n}}\big) - B(t)\big| > c\big(\sqrt{n}e^{-\sqrt{n}}\big)^{\frac12}\,.$$

In addition, we may choose n_0 big enough to ensure that $e^{-\sqrt{n_0}}$ is sufficiently small in the sense of Theorem 1.12. Then we pick $\varepsilon = e^{-\sqrt{n_0}}$ and, given $0 < h < \varepsilon$, choose n such that $e^{-\sqrt{n+1}} < h \le e^{-\sqrt{n}}$. Then, for t as above,

$$\begin{aligned} \big|B\big(t+h\big) - B(t)\big| &\ge \big|B\big(t+e^{-\sqrt{n}}\big) - B(t)\big| - \big|B\big(t+h\big) - B\big(t+e^{-\sqrt{n}}\big)\big| \\ &> c\big(\sqrt{n}e^{-\sqrt{n}}\big)^{\frac12} - C\,\Big(\big(e^{-\sqrt{n}} - e^{-\sqrt{n+1}}\big) \log\big(1/(e^{-\sqrt{n}} - e^{-\sqrt{n+1}}\big)\Big). \end{aligned}$$

It is not hard to see that the second (subtracted) term decays much more rapidly than the first, so that modifying n_0 to ensure that it is below $\delta\,(\sqrt{n}e^{-\sqrt{n}})^{\frac12}$ gives the result.

Exercise 1.7. It suffices to show that, for fixed $\varepsilon > 0$ and $c > 0$, almost surely, for all $t \geq 0$, there exists $0 < h < \varepsilon$ with $|B(t+h) - B(t)| > ch^\alpha$. By Brownian scaling we may further assume $\varepsilon = 1$. Note that, after this simplification, the complementary event means that there is a $t_0 \geq 0$ such that

$$\sup_{h \in (0,1)} \frac{B(t_0+h) - B(t_0)}{h^\alpha} \leq c \quad \text{and} \quad \inf_{h \in (0,1)} \frac{B(t_0+h) - B(t_0)}{h^\alpha} \geq -c\,.$$

We may assume that $t_0 \in [0,1)$. Fix $l \geq 1/(\alpha - \frac{1}{2})$. Then $t_0 \in \left[\frac{k-1}{2^n}, \frac{k}{2^n}\right)$ for any large n and some $0 \leq k < 2^n - l$. Then, by the triangle inequality, for all $j \in \{1, \ldots, 2^n - k\}$,

$$\left| B\left(\tfrac{k+j}{2^n}\right) - B\left(\tfrac{k+j-1}{2^n}\right) \right| \leq 2c \left(\tfrac{j+1}{2^n}\right)^\alpha.$$

Now, for any $0 \leq k < 2^n - l$, let $\Omega_{n,k}$ be the event

$$\Big\{ \big| B\left(\tfrac{k+j}{2^n}\right) - B\left(\tfrac{k+j-1}{2^n}\right) \big| \leq 2c \left(\tfrac{j+1}{2^n}\right)^\alpha \ \text{ for } \ j = 1, 2, \ldots, l \Big\}.$$

It suffices to show that, almost surely for all sufficiently large n and all $k \in \{0, \ldots, 2^n - l\}$ the event $\Omega_{n,k}$ does not occur. Observe that

$$\mathbb{P}(\Omega_{n,k}) \leq \Big[\mathbb{P}\big\{ |B(1)| \leq 2^{n/2}\, 2c \left(\tfrac{l+1}{2^n}\right)^\alpha \big\} \Big]^l \leq \big[2^{n/2}\, 2c \left(\tfrac{l+1}{2^n}\right)^\alpha \big]^l,$$

since the normal density is bounded by $1/2$. Hence, for a suitable constant C,

$$\mathbb{P}\Big(\bigcup_{k=0}^{2^n - l} \Omega_{n,k} \Big) \leq 2^n \left[2^{n/2}\, 2c \left(\tfrac{l+1}{2^n}\right)^\alpha \right]^l = C \big[2^{(1 - l(\alpha - 1/2))} \big]^n,$$

which is summable. Thus

$$\mathbb{P}\Big(\limsup_{n \to \infty} \bigcup_{k=0}^{2^n - l} \Omega_{n,k} \Big) = 0\,.$$

This is the required statement and hence the proof is complete.

Exercise 1.8. The proof can be found in [**Du95**, Chapter 3] or [**Ka02**, Theorem 3.15].

Exercise 1.10. Argue as in the proof of Theorem 1.30 with B replaced by $B + f$. The resulting term

$$\mathbb{P}\left\{ \big| B(1) + \sqrt{2^n}\, f\left((k+j)/2^n\right) - \sqrt{2^n}\, f\left((k+j-1)/2^n\right) \big| \leq 7M/\sqrt{2^n} \right\}$$

can be estimated in exactly the same manner as for the unshifted Brownian motion.

Exercise 1.11. This can be found, together with stronger and more general results, in [**BP84**]. Put $I = \big[B(1), \sup_{0\le s\le 1} B(s)\big]$, and define a function $g\colon I \to [0,1]$ by setting

$$g(x) = \sup\{s \in [0,1] \,:\, B(s) = x\}.$$

First check that almost surely the interval I is non-degenerate, g is strictly decreasing, left continuous and satisfies $B(g(x)) = x$. Then show that almost surely the set of discontinuities of g is dense in I. We restrict our attention to the event of probability 1 on which these assertions hold. Let

$$V_n = \big\{x \in I\colon g(x-h) - g(x) > nh \text{ for some } h \in (0, n^{-1})\big\}.$$

Now show that V_n is open and dense in I. By the Baire category theorem, $V := \bigcap_n V_n$ is uncountable and dense in I. Now if $x \in V$ then there is a sequence $x_n \uparrow x$ such that $g(x_n) - g(x) > n(x - x_n)$. Setting $t = g(x)$ and $t_n = g(x_n)$ we have $t_n \downarrow t$ and $t_n - t > n(B(t) - B(t_n))$, from which it follows that $D^*B(t) \ge 0$. On the other hand $D^*B(t) \le 0$ since $B(s) \le B(t)$ for all $s \in (t,1)$, by definition of $t = g(x)$.

Exercise 1.12. We first fix some positive ε and positive a. For some small h and an interval $I \subset [\varepsilon, 1-\varepsilon]$ with length h, we consider the event A that $t_0 \in I$ and we have

$$B(t_0 + \tilde h) - B(t_0) > -2ah^{1/4} \qquad \text{for some } h^{1/4} < \tilde h \le 2h^{1/4}.$$

We now denote by t_L the left endpoint of I. Using Theorem 1.12 we see there exists some positive C so that

$$B(t_0) - B(t_L) \le C\sqrt{h\log(1/h)}$$

Hence the event A implies the following events

$$A_1 = \big\{B(t_L - s) - B(t_L) \le C\sqrt{h\log(1/h)} \text{ for all } s \in [0,\varepsilon]\big\},$$

$$A_2 = \big\{B(t_L + s) - B(t_L) \le C\sqrt{h\log(1/h)} \text{ for all } s \in [0,h^{1/4}]\big\}.$$

We now define $T := \inf(s > t_L + h^{1/4} : B(s) > B(t) - 2ah^{1/4})$. Then by definition we have that $T \le t_L + 2h^{1/4}$ and this implies the event

$$A_3 = \big\{B(T+s) - B(T) \le 2ah^{1/4} + C\sqrt{h\log(1/h)} \text{ for all } s \in [0,\varepsilon]\big\}.$$

Now by the strong Markov property, these three events are independent and we obtain

$$\mathbb{P}(A) \le \mathbb{P}(A_1)\,\mathbb{P}(A_2)\,\mathbb{P}(A_3).$$

We estimate the probabilities of these three events and obtain

$$\begin{aligned}
\mathbb{P}(A_1) &= \mathbb{P}\big\{B(\varepsilon) \le C\sqrt{h\log(1/h)}\big\} \le \tfrac{1}{\sqrt{2\pi\varepsilon}}\, 2C\sqrt{h\log(1/h)},\\
\mathbb{P}(A_2) &= \mathbb{P}\big\{B(h^{1/4}) \le C\sqrt{h\log(1/h)}\big\} \le \tfrac{1}{\sqrt{2\pi h^{1/4}}}\, 2C\sqrt{h\log(1/h)},\\
\mathbb{P}(A_3) &= \mathbb{P}\big\{B(\varepsilon) \le 2ah^{1/4} + C\sqrt{h\log(1/h)}\big\} \le \tfrac{1}{\sqrt{2\pi\varepsilon}}\, 2\big(C\,h^{1/4} + 2ah^{1/4}\big).
\end{aligned}$$

Hence we obtain, for a suitable constant $K > 0$, depending on a and ε, that

$$\mathbb{P}(A) \le K\, h^{9/8}\log(1/h).$$

Summing over a covering collection of $1/h$ intervals of length h gives the bound

$$\begin{aligned}&\mathbb{P}\big\{t_0 \in [\varepsilon, 1-\varepsilon] \text{ and } B(t_0+\tilde{h}) - B(t_0) > -2ah^{1/4} \text{ for some } h^{1/4} < \tilde{h} \le 2h^{1/4}\big\}\\ &\quad \le K \log(1/h) h^{1/8}.\end{aligned}$$

Taking $h = 2^{-4n-4}$ in this bound and summing over n, we see that

$$\sum_{n=1}^{\infty} \mathbb{P}\Big\{t_0 \in [\varepsilon, 1-\varepsilon] \text{ and } \sup_{2^{-n-1}<h\le 2^{-n}} \frac{B(t_0+h)-B(t_0)}{h} > -a\Big\} < \infty,$$

and from the Borel-Cantelli lemma we obtain that, almost surely, either $t_0 \notin [\varepsilon, 1-\varepsilon]$, or

$$\limsup_{h\downarrow 0} \frac{B(t_0+h)-B(t_0)}{h} > -a.$$

Now recall that a and ε are arbitrary positive numbers, so taking a countable union over a and ε gives that, almost surely, $D^*B(t_0) = -\infty$, as required.

Exercise 1.13 By Brownian scaling it suffices to consider the case $t = 1$.

(a) We first show that, given $M > 0$ large, for any fixed point $s \in [0,1]$, almost surely there exists $n \in \mathbb{N}$ such that the dyadic interval $I(n,s) := [k2^{-n}, (k+1)2^{-n}]$ containing s satisfies

$$(*) \qquad \big|B\big((k+1)2^{-n}\big) - B\big(k2^{-n}\big)\big| \ge M\, 2^{-n/2}.$$

To see this, it is best to consider the construction of Brownian motion, see Theorem 1.4. Using the notation of that proof, let $d_0 = 1$ and $d_{n+1} \in \mathcal{D}_{n+1} \setminus \mathcal{D}_n$ be the dyadic point that splits the interval $[k2^{-n}, (k+1)2^{-n})$ containing s. This defines a sequence $Z_{d_n}, n = 0, 1, \ldots$ of independent, normally distributed random variables. Now let

$$n = \min\big\{k \in \{0,1,\ldots\} \colon |Z_{d_k}| \ge 3M\big\},$$

which is almost surely well-defined. Moreover,

$$\begin{aligned}3M \le \big|Z_{d_n}\big| &= 2^{\frac{n-1}{2}}\, \big|2B(d_n) - B(d-2^{-n}) - B(d+2^{-n})\big|\\ &\le 2^{\frac{n+1}{2}} \big|B(d) - B(d \pm 2^{-n})\big| + 2^{\frac{n-1}{2}} \big|B(d+2^{-n}) - B(d-2^{-n})\big|,\end{aligned}$$

where $\pm$ indicates that the inequality holds with either choice of sign. This implies that either $I(n,s)$ or $I(n-1,s)$ satisfies $(*)$. We denote by $N(s)$ be the smallest nonnegative integer n, for which $(*)$ holds.

By Fubini's theorem, almost surely, we have $N(s) < \infty$ for almost every $s \in [0,1]$. On this event, we can pick a finite collection of disjoint dyadic intervals $[t_{2j}, t_{2j+1}]$, $j = 0, \ldots, k-1$, with summed lengths exceeding $1/2$, say, such that the partition $0 = t_0 < \cdots < t_{2k} = 1$ given by their endpoints satisfies

$$\sum_{j=1}^{2k} \big(B(t_j) - B(t_{j-1})\big)^2 \ge M^2 \sum_{j=1}^{k} (t_{2j+1} - t_{2j}) \ge \frac{M}{2}\,,$$

from which (a) follows, as M was arbitrary.

(b) Note that the number of (finite) partitions of $[0,1]$ consisting of dyadic points is countable. Hence, by (a), given $n \in \mathbb{N}$, we can find a finite set P_n of partitions such that the probability that there exists a partition $0 = t_0 < \cdots < t_k = 1$ in P_n with the property that

$$\sum_{j=1}^{k} \big(B(t_j) - B(t_{j-1})\big)^2 \geq n$$

is bigger than $1 - \frac{1}{n}$. Successively enumerating the partitions in $P_1, P_2, \ldots$ yields a sequence satisfying the requirement of (b).

Exercise 1.14 To see convergence in the L^2-sense one can use the independence of the increments of a Brownian motion,

$$\mathbb{E}\Big[\sum_{j=1}^{k(n)} \big(B\big(t_{j+1}^{(n)}\big) - B\big(t_j^{(n)}\big)\big)^2 - t\Big]^2 = \sum_{j=1}^{k(n)} \mathbb{E}\Big[\big(B\big(t_{j+1}^{(n)}\big) - B\big(t_j^{(n)}\big)\big)^2 - \big(t_{j+1}^{(n)} - t_j^{(n)}\big)\Big]^2$$

$$\leq \sum_{j=1}^{k(n)} \mathbb{E}\Big[\big(B\big(t_{j+1}^{(n)}\big) - B\big(t_j^{(n)}\big)\big)^4 + \big(t_{j+1}^{(n)} - t_j^{(n)}\big)^2\Big].$$

Now, using that the fourth moments of a centred normal distribution with variance σ^2 is $3\sigma^4$, this can be estimated by a constant multiple of

$$\sum_{j=1}^{k(n)} \big(t_{j+1}^{(n)} - t_j^{(n)}\big)^2,$$

which goes to zero. Moreover, by the Markov inequality

$$\mathbb{P}\Big\{\Big|\sum_{j=1}^{k(n)} \big(B\big(t_{j+1}^{(n)}\big) - B\big(t_j^{(n)}\big)\big)^2 - t\Big| > \varepsilon\Big\} \leq \varepsilon^{-2}\, \mathbb{E}\Big[\sum_{j=1}^{n} \big(B\big(t_{j+1}^{(n)}\big) - B\big(t_j^{(n)}\big)\big)^2 - t\Big]^2,$$

and summability of the right hand side together with the Borel-Cantelli lemma ensures almost sure convergence.

Exercise 2.3. We first show that, given two disjoint closed time intervals, the maxima of Brownian motion on them are different almost surely. For this purpose, let $[a_1, b_1]$ and $[a_2, b_2]$ be two fixed intervals with $b_1 < a_2$. Denote by m_1 and m_2, the maxima of Brownian motion on these two intervals. Applying the Markov property at time b_1 we see that the random variable $B(a_2) - B(b_1)$ is independent of $m_1 - B(b_1)$. Using the Markov property at time a_2 we see that $m_2 - B(a_2)$ is also independent of both these variables. Conditioning on the values of the random variables $m_1 - B(b_1)$ and $m_2 - B(a_2)$, the event $m_1 = m_2$ can be written as

$$B(a_2) - B(b_1) = m_1 - B(b_1) - (m_2 - B(a_2)).$$

The left hand side being a continuous random variable, and the right hand side a constant, we see that this event has probability 0.

Now the statement just proved holds jointly for all disjoint pairs of intervals with rational endpoints. The proposition follows, since if Brownian motion has a non-strict local maximum, there are two disjoint rational intervals where Brownian motion has the same maximum.

Exercise 2.4. **(i)** If $A \in \mathcal{F}(S)$, then $A \cap \{T \le t\} = (A \cap \{S \le t\}) \cap \{T \le t\} \in \mathcal{F}^+(t)$.

(ii) By (i), $\mathcal{F}(T) \subset \mathcal{F}(T_n)$ for all n, which proves $\subset$. On the other hand, if $A \in \bigcap_{n=1}^{\infty} \mathcal{F}(T_n)$, then for all $t \ge 0$,

$$A \cap \{T < t\} = \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} A \cap \{T_n < t\} \in \mathcal{F}^+(t).$$

(iii) Look at the discrete stopping times T_n defined in the previous example. We have, for any Borel set $A \subset \mathbb{R}^d$,

$$\{B(T_n) \in A\} \cap \big\{T_n \le k2^{-n}\big\} = \bigcup_{m=0}^{k} \Big(\{B(m2^{-n}) \in A\} \cap \{T_n = m2^{-n}\}\Big) \in \mathcal{F}^+(k2^{-n}).$$

Hence $B(T_n)$ is $\mathcal{F}(T_n)$-measurable, and as $T_n \downarrow T$, we get that $B(T) = \lim B(T_n)$ is $\mathcal{F}(T_n)$-measurable for any n. Hence $B(T)$ is $\mathcal{F}(T)$-measurable by part (ii).

Exercise 2.5. If $T = 0$ almost surely, there is nothing to show, hence we may assume $\mathbb{E}[T] > 0$.

(a) By construction, T_n is the sum of n independent random variables with the law of T, hence, by the law of large numbers, almost surely,

$$\lim_{n\to\infty} \frac{T_n}{n} = \mathbb{E}[T] > 0,$$

which, by assumption, is finite. This implies, in particular, that $T_n \to \infty$ almost surely, and together with the law of large numbers for Brownian motion, Corollary 1.11, we get almost surely, $\lim_{n\to\infty} B(T_n)/T_n = 0$. The two limit statements together show that, almost surely,

$$\lim_{n\to\infty} \frac{B(T_n)}{n} = \lim_{n\to\infty} \frac{B(T_n)}{T_n} \lim_{n\to\infty} \frac{T_n}{n} = 0.$$

(b) Again by construction, $B(T_n)$ is the sum of n independent random variables with the law of $B(T)$, which we conveniently denote $X_1, X_2, \ldots$. As

$$\lim_{n\to\infty} \frac{X_n}{n} = \lim_{n\to\infty} \frac{B(T_n)}{n} - \lim_{n\to\infty} \frac{B(T_{n-1})}{n} = 0,$$

the event $\{|X_n| \ge n\}$ occurs only finitely often, so that the Borel-Cantelli lemma implies

$$\sum_{n=0}^{\infty} \mathbb{P}\big\{|X_n| \ge n\big\} < \infty.$$

Hence we have that

$$\mathbb{E}[B(T)] = \mathbb{E}|X_n| \le \sum_{n=0}^{\infty} \mathbb{P}\big\{|X_n| \ge n\big\} < \infty.$$

(c) By the law of large numbers, almost surely,

$$\lim_{n\to\infty} \frac{B(T_n)}{n} = \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} X_j = \mathbb{E}[B(T)].$$

Exercise 2.7. Let S be a nonempty, closed set S with no isolated points. To see that it is uncountable, we construct a subset with the cardinality of $\{1,2\}^{\mathbb{N}}$. Start by choosing a point $x_1 \in S$. As this point is not isolated there exists a further, different point $x_2 \in S$. Now pick two disjoint closed balls B_1, B_2 around these points. Again, as x_1 is not isolated, we can find two points in $B_1 \cap S$, around which we can put disjoint balls contained in $B_1 \cap S$, similarly for $B_2 \cap S$, and so on. Now there is a bijection between $\{1,2\}^{\mathbb{N}}$ and the decreasing sequences of balls in our construction. The intersection of the balls in each such sequence contains, as S is closed, at least one point of S, and two points belonging to two different sequences are clearly different. This completes the proof.

Exercise 2.11. By Fubini's theorem,

$$\mathbb{E}[T^\alpha] = \int_0^\infty \mathbb{P}\{T > x^{1/\alpha}\}\,dx \le 1 + \int_1^\infty \mathbb{P}\big\{M(x^{1/\alpha}) < 1\big\}\,dx.$$

Note that, by Brownian scaling, $\mathbb{P}\{M(x^{1/\alpha}) < 1\} \le C\, x^{-\frac{1}{2\alpha}}$ for a suitable constant $C > 0$, which implies that $\mathbb{E}[T^\alpha] < \infty$, as required.

Exercise 2.14. By Exercise 2.13 the process $\{X(t) \colon t \ge 0\}$ defined by

$$X(t) = \exp\big\{2bB(t) - 2b^2 t\big\} \qquad \text{for } t \ge 0,$$

defines a martingale. Observe that $T = \inf\{t > 0 \colon B(t) = a + bt\}$ is a stopping time for the natural filtration, which is finite exactly if $B(t) = a + bt$ for some $t > 0$. Then

$$\mathbb{P}\{T < \infty\} = e^{-2ab}\,\mathbb{E}\big[X(T)\,\mathbb{1}\{T < \infty\}\big],$$

and because $\{X^T(t) \colon t \ge 0\}$ is bounded, the right hand side equals e^{-2ab}.

Exercise 2.15. Use the binomial expansion of $(B(t) + (B(t+h) - B(t)))^3$ to deduce that $X(t) = B(t)^3 - 3tB(t)$ defines a martingale. We know that $\mathbb{P}_x\{T_R < T_0\} = x/R$. Write $\tau_* = \tau(\{0,R\})$. Then

$$\begin{aligned} x^3 &= \mathbb{E}_x[X(0)] = \mathbb{E}_x[X(\tau_*)] = \mathbb{P}_x\{T_R < T_0\}\,\mathbb{E}_x\big[X(\tau_*)\,|\,T_R < T_0\big] \\ &= \mathbb{P}_x\{T_R < T_0\}\,\mathbb{E}_x\big[R^3 - 3\tau_* R\,|\,T_R < T_0\big] = (x/R)(R^3 - 3\gamma R) = x\,(R^2 - 3\gamma)\,. \end{aligned}$$

Solving the last equation for γ gives the claim.

Exercise 2.18. Part (a) can be proved similarly to Theorem 2.47, which in fact is the special case $\lambda = 0$ of this exercise. For part (b) choose $u \colon U \to \mathbb{R}$ as a bounded solution of

$$\tfrac{1}{2}\Delta u(x) = \lambda\, u(x), \qquad \text{for } x \in U\,,$$

with $\lim\limits_{x \to x_0} u(x) = f(x_0)$ for all $x_0 \in \partial U$. Then

$$X(t) = e^{-\lambda t} u(B(t)) - \int_0^t e^{-\lambda s}\big(\tfrac{1}{2}\Delta u(B(s)) - \lambda u(B(s))\big)\,ds$$

defines a martingale. For any compact $K \subset U$ we can pick a twice continuously differentiable function $v\colon \mathbb{R}^d \to \mathbb{R}$ with $v = u$ on K and $v = 0$ on U^{c} Apply the optional stopping theorem to stopping times $S = 0$, $T = \inf\{t \geq 0\colon B(T) \notin K\}$. to get, for every $x \in K$,

$$u(x) = \mathbb{E}[X(0)] = \mathbb{E}[X(T)] = \mathbb{E}_x\big[e^{-\lambda T} f(B(T))\big]\,.$$

Now choose a sequence $K_n \uparrow U$ of compacts and pass to the limit on the right hand side of the equation.

Exercise 3.2. To prove the result for $k = 1$ estimate $|u(x) - u(y)|$ in terms of $|\,x - y|$ using the mean value formula for harmonic functions and the fact that, if x and y are close, the volume of the symmetric difference of $\mathcal{B}(x, r)$ and $\mathcal{B}(y, r)$ is bounded by a constant multiple of $r^{d-1}|\,x - y|$. For general k note that the partial derivatives of a harmonic function are themselves harmonic, and iterate the estimate.

Exercise 3.4. Define a random variable Y by $Y := X$, if $X > \lambda\mathbb{E}[X]$, and $Y := 0$, otherwise. Applying the Cauchy-Schwarz inequality to $\mathbb{E}[Y] = \mathbb{E}[Y \mathbb{1}\{Y > 0\}]$ gives

$$\mathbb{E}[Y\,\mathbb{1}\{Y > 0\}] \leq \mathbb{E}[Y^2]^{1/2}\left(\mathbb{P}\{Y > 0\}\right)^{1/2},$$

hence, as $X \geq Y \geq X - \lambda\mathbb{E}[X]$, we get

$$\mathbb{P}\big\{X > \lambda\,\mathbb{E}[X]\big\} = \mathbb{P}\{Y > 0\} \geq \frac{\mathbb{E}[Y]^2}{\mathbb{E}[Y^2]} \geq (1 - \lambda)^2\,\frac{\mathbb{E}[X]^2}{\mathbb{E}[X^2]}\,.$$

Exercise 3.6. For $d \geq 3$, choose a and b such that $a + br^{2-d} = \tilde{u}(r)$, and $a + bR^{2-d} = \tilde{u}(R)$. Notice that the harmonic functions given by $u(x) = \tilde{u}(|\,x|\,)$ and $v(x) = a + b|\,x|^{2-d}$ agree on ∂D. They also agree on D by Corollary 3.7. So $u(x) = a + b|\,x|^{2-d}$. By similar consideration we can show that $u(x) = a + b\log|\,x|$ in the case $d = 2$.

Exercise 3.7. Let $x, y \in \mathbb{R}^d$, $a = |\,x - y|$. Suppose u is a positive harmonic function. Then

$$\begin{aligned} u(x) &= \frac{1}{\mathcal{L}\mathcal{B}(x,R)}\int_{\mathcal{B}(x,R)} u(z)\,dz \\ &\leq \frac{\mathcal{L}\mathcal{B}(y,R+a)}{\mathcal{L}\mathcal{B}(x,R)}\,\frac{1}{\mathcal{L}\mathcal{B}(y,R+a)}\int_{\mathcal{B}(y,R+a)} u(z)\,dz = \frac{(R+a)^d}{R^d}\,u(y). \end{aligned}$$

This converges to $u(y)$ as $R \to \infty$, so $u(x) \leq u(y)$, and by symmetry, $u(x) = u(y)$ for all x, y. Hence u is constant.

Exercise 3.8. Uniqueness is clear, because there is at most one *continuous* extension of u. Let $D_0 \subset D$ be a ball whose closure is contained in D, which contains x. u is bounded and harmonic on $D_1 = D_0 \setminus \{x\}$ and continuous on $\overline{D}_1 \setminus \{x\}$. Show that this already implies that $u(z) = \mathbb{E}_z[u(\tau(D_1))]$ on D_1 and that the right hand side has an obvious harmonic extension to $D_1 \cup \{x\}$, which defines the global extension.

Exercise 3.12. To obtain joint continuity one can show equicontinuity of $G(x,\,\cdot\,)$ and $G(\,\cdot\,,x)$ in $D\setminus\mathcal{B}(x,\varepsilon)$ for any $\varepsilon>0$. This follows from the fact that these functions are harmonic, by Exercise 3.11, and the estimates of Exercise 3.2.

Exercise 3.13. Recall that

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

The expectation can be evaluated (one can see how in the proof of Theorem 3.43). The final answer is

$$G(x,y)=\begin{cases}-\frac{1}{\pi}\log|\,x/R-y/R|+\frac{1}{\pi}\log\left|\frac{x}{|x|}-|\,x|yR^{-2}\right|, & \text{if } x\neq 0,\ x,y\in\mathcal{B}(0,R),\\ -\frac{1}{\pi}\log|y/R| & \text{if } x=0,\ y\in\mathcal{B}(0,R).\end{cases}$$

Exercise 3.14. Suppose $x,y\notin\overline{\mathcal{B}(0,r)}$ and $A\subset\mathcal{B}(0,r)$ compact. Then, by the strong Markov property applied to the first hitting time of $\partial\mathcal{B}(0,r)$,

$$\mu_A(x,\,\cdot\,)=\int_{\partial\mathcal{B}(0,r)}\mu_A(z,\,\cdot\,)\,d\mu_{\partial\mathcal{B}(0,r)}(x,dz)\,.$$

Use Theorem 3.43 to show that, for $B\subset A$ Borel, $\mu_{\partial\mathcal{B}(0,r)}(x,B)\le C\mu_{\partial\mathcal{B}(0,r)}(y,B)$ for a constant C not depending on B. Complete the argument from there.

Exercise 4.1. Let $\alpha=\log 2/\log 3$. For the upper bound it suffices to find an efficient covering of C by intervals of diameter ε. If $\varepsilon\in(0,1)$ is given, let n be the integer such that $1/3^n<2\varepsilon\le 1/3^{n-1}$ and look at the sets

$$\Big[\sum_{i=1}^n\frac{x_i}{3^i},\sum_{i=1}^n\frac{x_i}{3^i}+\varepsilon\Big]\ \text{ for } (x_1,\ldots,x_n)\in\{0,2\}^n.$$

These sets obviously cover C and each of them is contained in an open ball centred in an interval of diameter 2ε. Hence

$$M(C,\varepsilon)\le 2^n=3^{\alpha n}=3^\alpha\big(3^{n-1}\big)^\alpha\le 3^\alpha(1/\varepsilon)^\alpha\,.$$

This implies $\overline{\dim}_M C\le\alpha\,.$

For the lower bound we may assume we have a covering by intervals $(x_k-\varepsilon,x_k+\varepsilon)$, with $x_k\in C$, and let n be the integer such that $1/3^{n+1}\le 2\varepsilon<1/3^n$. Let $x_k=\sum_{i=1}^\infty x_{i,k}3^{-i}$. Then

$$B(x_k-\varepsilon,x_k+\varepsilon)\cap C\subset\Big\{\sum_{i=1}^\infty\frac{y_i}{3^i}\ :\ y_1=x_{1,k},\ldots,y_n=x_{n,k}\Big\},$$

and we need at least 2^n sets of the latter type to cover C. Hence,

$$M(C,\varepsilon)\ge 2^n=3^{\alpha n}=(1/3)^\alpha\big(3^{n+1}\big)^\alpha\ge(1/3)^\alpha(1/\varepsilon)^\alpha\,.$$

This implies $\underline{\dim}_M C\ge\alpha\,.$

Exercise 4.2. Given $\varepsilon \in (0,1)$ find the integer n such that $1/(n+1)^2 \leq \varepsilon < 1/n^2$. Then the points in $\{1/k : k > n\} \cup \{0\}$ can be covered by $n+1$ intervals of diameter ε. n further balls suffice to cover the remaining n points. Hence

$$M(E,\varepsilon) \leq 2n+1 \leq \tfrac{2n+1}{n}\,(1/\varepsilon)^{1/2}\,,$$

implying $\overline{\dim}_M(E) \leq 1/2$. On the other hand, as the distance between neighbouring points is

$$\frac{1}{k} - \frac{1}{k+1} = \frac{1}{k(k+1)} \geq \frac{1}{(k+1)^2}\,,$$

we always need at least $n-1$ sets of diameter ε to cover E, which implies

$$M(E,\varepsilon) \geq n-1 \geq \tfrac{n-1}{n+1}\,(1/\varepsilon)^{1/2}\,,$$

hence $\underline{\dim}_M(E) \geq 1/2$.

Exercise 4.3. Suppose E is a bounded metric space with $\underline{\dim}_M E < \alpha$. Choose $\varepsilon > 0$ such that $\underline{\dim} E_M < \alpha - \varepsilon$. Then, for every k there exists $0 < \delta < \frac{1}{k}$ and a covering $E_1, \ldots, E_n$ of E by sets of diameter at most δ with $n \leq \delta^{-\alpha+\varepsilon}$. The α-value of this covering is at most $n\delta^{\alpha} \leq \delta^{\varepsilon}$, which tends to zero for large k. Hence $\mathcal{H}^{\alpha}_{\infty}(E) = 0$, and $\dim E \leq \alpha$.

Exercise 4.4. Indeed, as $E \subset F$ implies $\dim E \leq \dim F$, it is obvious that

$$\dim \bigcup_{k=1}^{\infty} E_k \geq \sup\big\{ \dim E_k \,:\, k \geq 1 \big\}.$$

To see the converse, we use

$$\begin{aligned}
\mathcal{H}^{\alpha}_{\infty}\Big(\bigcup_{k=1}^{\infty} E_k\Big) &\leq \inf\Big\{ \sum_{k=1}^{\infty}\sum_{j=1}^{\infty} |E_{j,k}|^{\alpha} \,:\, E_{1,k}, E_{2,k}, \ldots \text{ covers } E_k \Big\} \\
&= \sum_{k=1}^{\infty} \inf\Big\{ \sum_{j=1}^{\infty} |E_{j,k}|^{\alpha} \,:\, E_{1,k}, E_{2,k}, \ldots \text{ covers } E_k \Big\} = \sum_{k=1}^{\infty} \mathcal{H}^{\alpha}_{\infty}(E_k)\,.
\end{aligned}$$

Hence,

$$\begin{aligned}
\dim \bigcup_{k=1}^{\infty} E_k &\leq \sup\Big\{\alpha \geq 0 \,:\, \mathcal{H}^{\alpha}_{\infty}\big(\bigcup_{k=1}^{\infty} E_k\big) > 0\Big\} \leq \sup\Big\{\alpha \geq 0 \,:\, \sum_{k=1}^{\infty} \mathcal{H}^{\alpha}_{\infty}\big(E_k\big) > 0\Big\} \\
&\leq \sup_{k=1}^{\infty} \sup\Big\{\alpha \geq 0 \,:\, \mathcal{H}^{\alpha}_{\infty}\big(E_k\big) > 0\Big\}\,.
\end{aligned}$$

This proves the converse inequality.

Exercise 4.6. Suppose that f is surjective and α-Hölder continuous with Hölder constant $C > 0$, and assume that $\mathcal{H}^{\alpha\beta}(E_1) < \infty$. Given $\varepsilon, \delta > 0$ we can cover E_1 with sets $B_1, B_2, \dots$ of diameter at most δ such that

$$\sum_{i=1}^{\infty} |B_i|^{\alpha\beta} \leq \mathcal{H}^{\alpha\beta}(E_1) + \varepsilon\,.$$

Note that the sets $f(B_1), f(B_2), \dots$ cover E_2 and that $|f(B_i)| \leq C\,|B_i|^{\alpha} \leq C\,\delta^{\alpha}$. Hence

$$\sum_{i=1}^{\infty} |f(B_i)|^{\beta} \leq C^{\beta} \sum_{i=1}^{\infty} |B_i|^{\alpha\beta} \leq C^{\beta}\,\mathcal{H}^{\alpha\beta}(E_1) + C^{\beta}\,\varepsilon,$$

from which the claimed result for the Hausdorff measure readily follows.

Exercise 4.8. Start with $d = 1$. For any $0 < a < 1/2$ let $C(a)$ be the Cantor set obtained by iteratively removing from each construction interval a central interval of $1 - 2a$ of its length. Note that at the nth level of the construction we have 2^n intervals each of length a^n. It is not hard to show that $C(a)$ has Hausdorff dimension $\log 2/\log(1/a)$, which solves the problem for the case $d = 1$.

For arbitrary dimension d and given α we find a such that $\dim C(a) = \alpha/d$. Then the Cartesian product $C(a) \times \overset{d}{\dots} \times C(a)$ has dimension α. The upper bound is straightforward, and the lower bound can be verified, for example, from the mass distribution principle, by considering the natural measure that places mass $1/2^{dn}$ to each of the 2^{dn} cubes of sidelength a^n at the nth construction level.

Exercise 4.13. Recall that it suffices to show that $\mathcal{H}^{1/2}(\mathsf{Rec}) = 0$ almost surely. In the proof of Lemma 4.21 the maximum process was use to define a measure on the set of record points: this measure can be used to define 'big intervals' analogous to the 'big cubes' in the proof of Theorem 4.18. A similar covering strategy as in this proof yields the result.

Exercise 5.1. Use the Borel-Cantelli lemma for the events

$$E_n = \Big\{ \sup_{n \leq t < n+1} B(t) - B(n) \geq \sqrt{a \log n} \Big\}$$

and test for which values of a the series $\mathbb{P}(E_n)$ converges. To estimate the probabilities, the reflection principle and Lemma II.3.1 will be useful.

Exercise 5.2. The lower bound is immediate from the one-dimensional statement. For the upper bound pick a finite subset $S \subset \partial\mathcal{B}(0,1)$ of directions such that, for every $x \in \partial\mathcal{B}(0,1)$ there exists $\tilde{x} \in S$ with $|x - \tilde{x}| < \varepsilon$. Almost surely, all Brownian motions in dimension one obtained by projecting $\{B(t) : t \geq 0\}$ on the line determined by the vectors in S satisfy the statement. From this one can infer that the limsup under consideration is bounded from above by $1 + \varepsilon$.

Exercise 5.3. Let $T_a = \inf\{t > 0 \colon B(t) = a\}$. The proof of the upper bound can be based on the fact that, for $A < 1$ and $q > 1$,

$$\sum_{n=1}^{\infty} \mathbb{P}\big\{\psi(T_1 - T_{1-q^n}) < \tfrac{1}{A}\, 2^{-n}\big\} < \infty.$$

Exercise 5.4. Define the stopping time $\tau_{-1} = \min\{k : S_k = -1\}$ and recall the definition of p_n from (2.3). Then

$$p_n = \mathbb{P}\{S_n \geq 0\} - \mathbb{P}\{S_n \geq 0, \tau_{-1} < n\}.$$

Let $\{S_j^* : j \geq 0\}$ denote the random walk reflected at time τ_{-1}, that is

$$\begin{aligned} S_j^* &= S_j && \text{for } j \leq \tau_{-1}, \\ S_j^* &= (-1) - (S_j + 1) && \text{for } j > \tau_{-1}. \end{aligned}$$

Note that if $\tau_{-1} < n$ then $S_n \geq 0$ if and only if $S_n^* \leq -2$, so

$$p_n = \mathbb{P}\{S_n \geq 0\} - \mathbb{P}\{S_n^* \leq -2\}.$$

Using symmetry and the reflection principle, we have

$$p_n = \mathbb{P}\{S_n \geq 0\} - \mathbb{P}\{S_n \geq 2\} = \mathbb{P}\big\{S_n \in \{0, 1\}\big\},$$

which means that

$$\begin{aligned} p_n &= \mathbb{P}\{S_n = 0\} &= \binom{n}{n/2} 2^{-n} && \text{for } n \text{ even}, \\ p_n &= \mathbb{P}\{S_n = 1\} &= \binom{n}{(n-1)/2} 2^{-n} && \text{for } n \text{ odd}. \end{aligned}$$

Recall that Stirling's Formula gives $m! \sim \sqrt{2\pi} m^{m+1/2} e^{-m}$, where the symbol $\sim$ means that the ratio of the two sides approaches 1 as $m \to \infty$. One can deduce from Stirling's Formula that $p_n \sim \sqrt{2/\pi n}$, which proves the result.

Exercise 5.5. Denote by $I_n(k)$ the event that k is a point of increase for $S_0, S_1, \ldots, S_n$ and by $F_n(k) = I_n(k) \setminus \bigcup_{i=0}^{k-1} I_n(i)$ the event that k is the first such point. The events that $\{S_k$ is largest among $S_0, S_1, \ldots S_k\}$ and that $\{S_k$ is smallest among $S_k, S_{k+1}, \ldots S_n\}$ are independent, and therefore $\mathbb{P}(I_n(k)) = p_k p_{n-k}$.

Observe that if S_j is minimal among $S_j, \ldots, S_n$, then any point of increase for $S_0, \ldots, S_j$ is automatically a point of increase for $S_0, \ldots, S_n$. Therefore for $j \leq k$ we can write

$$F_n(j) \cap I_n(k) = \\ F_j(j) \cap \{S_j \leq S_i \leq S_k \text{ for all } i \in [j,k]\} \cap \{S_k \text{ is minimal among } S_k, \ldots, S_n\}.$$

The three events on the right-hand side are independent, as they involve disjoint sets of summands; the second of these events is of the type considered in Lemma 5.9. Thus,

$$\begin{aligned} \mathbb{P}(F_n(j) \cap I_n(k)) &\geq \mathbb{P}(F_j(j))\, p_{k-j}^2\, p_{n-k} \\ &\geq p_{k-j}^2\, \mathbb{P}(F_j(j))\, \mathbb{P}\{S_j \text{ is minimal among } S_j, \ldots, S_n\}, \end{aligned}$$

since $p_{n-k} \geq p_{n-j}$. Here the two events on the right are independent, and their intersection is precisely $F_n(j)$. Consequently $\mathbb{P}(F_n(j) \cap I_n(k)) \geq p_{k-j}^2 \mathbb{P}(F_n(j))$.

Decomposing the event $I_n(k)$ according to the first point of increase gives

$$\begin{aligned}\sum_{k=0}^{n} p_k p_{n-k} = \sum_{k=0}^{n} \mathbb{P}(I_n(k)) &= \sum_{k=0}^{n}\sum_{j=0}^{k} \mathbb{P}(F_n(j)\cap I_n(k)) \\ \geq \sum_{j=0}^{\lfloor n/2\rfloor} \sum_{k=j}^{j+\lfloor n/2\rfloor} p_{k-j}^2 \mathbb{P}(F_n(j)) &\geq \sum_{j=0}^{\lfloor n/2\rfloor} \mathbb{P}(F_n(j)) \sum_{i=0}^{\lfloor n/2\rfloor} p_i^2 \,.\end{aligned} \tag{0.1}$$

This yields an upper bound on the probability that $\{S_j : j = 0, \ldots, n\}$ has a point of increase by time $n/2$; but this random walk has a point of increase at time k if and only if the reversed walk $\{S_n - S_{n-i} : i = 0, \ldots, n\}$ has a point of increase at time $n-k$. Thus, doubling the upper bound given by (0.1) proves the statement.

Exercise 5.7. In the proof of Exercise 5.5 we have seen that,

$$\sum_{k=0}^{n} p_k p_{n-k} = \sum_{k=0}^{n} \mathbb{P}(I_n(k)) = \sum_{k=0}^{n}\sum_{j=0}^{k} \mathbb{P}(F_n(j)\cap I_n(k))\,.$$

By Lemma 5.9, we have, for $j \leq k \leq n$,

$$\begin{aligned}\mathbb{P}(F_n(j)\cap I_n(k)) &\leq \mathbb{P}(F_n(j)\cap \{S_j \leq S_i \leq S_k \text{ for } j \leq i \leq k\}) \\ &\leq \mathbb{P}(F_n(j)) p^2_{\lfloor (k-j)/2\rfloor}.\end{aligned}$$

Thus,

$$\sum_{k=0}^{n} p_k p_{n-k} \leq \sum_{k=0}^{n}\sum_{j=0}^{k} \mathbb{P}(F_n(j)) p^2_{\lfloor (k-j)/2\rfloor} \leq \sum_{j=0}^{n} \mathbb{P}(F_n(j)) \sum_{i=0}^{n} p^2_{\lfloor i/2\rfloor}.$$

This implies the statement.

Exercise 5.9. Suppose that X is an arbitrary random variable with vanishing expectation and finite variance. For each $n \in \mathbb{N}$ divide the intersection of the support of X with the interval $[-n, n]$ into finitely intervals with mesh $< \frac{1}{n}$. If $x_1 < \cdots < x_m$ are the partition points, construct the law of X_n by placing, for any $j \in \{0, \ldots, m\}$, atoms of size $P\{X \in [x_j, x_{j+1})\}$ in position $E[X \mid x_j \leq X < x_{j+1}]$, using the convention $x_0 = -\infty$ and $x_{m+1} = \infty$. By construction, X_n takes only finitely many values.

Observe that $E[X_n] = 0$ and X_n converges to X in distribution. Moreover, one can show that $\tau_n \to \tau$ almost surely. This implies that $B(\tau_n) \to B(\tau)$ almost surely, and therefore also in distribution, which implies that X has the same law as $B(\tau)$. Fatou's lemma implies that

$$\mathbb{E}[\tau] \leq \liminf_{n\uparrow\infty} \mathbb{E}\big[\tau_n\big] = \liminf_{n\uparrow\infty} E\big[X_n^2\big] < \infty.$$

Hence, by Wald's second lemma, $E[X^2] = \mathbb{E}[B(\tau)^2] = \mathbb{E}[\tau]$.

Exercise 6.4. From Exercise 2.15 we get, for any $x \in (0,1)$ that

$$\mathbb{E}_x\big[T_1 \,\big|\, T_1 < T_0\big] = \frac{1-x^2}{3}, \qquad \mathbb{E}_x\big[T_0 \,\big|\, T_0 < T_1\big] = \frac{2x-x^2}{3},$$

where T_0, T_1 are the first hitting times of the points 0, resp. 1.

Define stopping times $\tau_0^{(x)} = 0$ and, for $j \geq 1$,

$$\sigma_j^{(x)} = \inf\{t > \tau_{j-1}^{(x)} \colon B(t) = x\}, \qquad \tau_j^{(x)} = \inf\{t > \sigma_j^{(x)} \colon B(t) \in \{0,1\}\}.$$

Let $N^{(x)} = \min\{j \geq 1 \colon B(\tau_j^{(x)}) = 1\}$. Then $N^{(x)}$ is geometric with parameter x. We have

$$\int_0^{T_1} \mathbb{1}\{0 \leq B(s) \leq 1\}\, ds = \lim_{x \downarrow 0} \sum_{j=1}^{N^{(x)}} (\tau_j^{(x)} - \sigma_j^{(x)}).$$

and this limit is increasing. Hence

$$\begin{aligned}
\mathbb{E} \int_0^{T_1} & \mathbb{1}\{0 \leq B(s) \leq 1\}\, ds \\
&= \lim_{x\downarrow 0} \mathbb{E}\big[N^{(x)} - 1\big] \mathbb{E}\big[\tau_1^{(x)} - \sigma_1^{(x)} \,\big|\, B(\tau_1^{(x)}) = 0\big] + \lim_{x\downarrow 0} \mathbb{E}\big[\tau_1^{(x)} - \sigma_1^{(x)} \,\big|\, B(\tau_1^{(x)}) = 1\big] \\
&= \lim_{x\downarrow 0} \Big(\frac{1}{x} - 1\Big) \frac{2x - x^2}{3} + \lim_{x\downarrow 0} \frac{1-x^2}{3} = 1.
\end{aligned}$$

Exercise 6.5. Observe that $\mathbb{E}\exp\{\lambda Z_j\} = e^\lambda/(2-e^\lambda)$ for all $\lambda < \log 2$, and hence, for a suitable constant$C$ and all small $\lambda > 0$,

$$\mathbb{E}\exp\big\{\lambda(Z_j - 2)\big\} \leq \lambda^2 + C\lambda^3,$$

by a Taylor expansion. Using this for $\lambda = \frac{\varepsilon}{2}$ we get from Chebyshev's inequality,

$$\begin{aligned}
\mathbb{P}\Big\{ \sum_{j=1}^k (Z_j - 2) > m\varepsilon \Big\} &\leq \exp\{-m\tfrac{\varepsilon^2}{2}\} \Big(\mathbb{E}\exp\{\tfrac{\varepsilon}{2}\,(Z_j-2)\}\Big)^k \\
&\leq \exp\big\{-m\tfrac{\varepsilon^2}{2}\big\} \exp\big\{m\big(\tfrac{\varepsilon^2}{4} + C\tfrac{\varepsilon^3}{8}\big)\big\},
\end{aligned}$$

which proves the the more difficult half of the claim. The inequality for the lower tail is obvious.

Exercise 6.6. We have that

$$\mathbb{P}\Big\{\frac{(X+\ell)^2}{2} \leq t\Big\} = \mathbb{P}\big\{-\sqrt{2t} - \ell \leq X \leq \sqrt{2t} - \ell\big\}.$$

So the density of the left hand side is

$$\frac{1}{2\sqrt{\pi t}}\, e^{-(2t+\ell^2)/2} \Big[e^{\ell\sqrt{2t}} + e^{-\ell\sqrt{2t}}\Big],$$

which by Taylor expansion is

$$\frac{1}{\sqrt{\pi t}}\, e^{-(2t+\ell^2)/2} \sum_{k=0}^\infty \frac{(\ell\sqrt{2t})^{2k}}{(2k)!}.$$

Recall that $X^2/2$ is distributed as Gamma($\frac{1}{2}$), and given N the sum $\sum_{i=1}^N Z_i$ is distributed as Gamma(N). By conditioning on N, we get that the density of the right hand side is

$$\sum_{k=0}^{\infty} \frac{\ell^{2k} e^{-\ell^2/2} t^{k-1/2} e^{-t}}{2^k k! \Gamma(k+\frac{1}{2})} .$$

Recall that

$$\Gamma(k+\frac{1}{2}) = \frac{\sqrt{\pi}(2k)!}{2^{2k}k!} ,$$

and so the densities of both sides are equal.

Exercise 7.1. Use that $\int_0^T H(s)\,dB(s) = \int_0^\infty H^T(s)\,dB(s)$.

Exercise 7.3. First establish a Taylor formula of the form

$$\begin{aligned}\big|f(x,y) - f(x_0,y_0) - \nabla_y f(x_0,y_0)\cdot(y-y_0)&\\ - \nabla_x f(x_0,y_0)\cdot(x-x_0) - \tfrac{1}{2}(x-x_0)^T \mathrm{Hes}_x f(x_0,y_0)(x-x_0)\big|&\\ \le \omega_1(\delta,M)\,|y-y_0| + \omega_2(\delta,M)|x-x_0|^2,&\end{aligned}$$

where $\mathrm{Hes}_x f = (\partial_{ij} f)$ is the $d\times d$-Hessian matrix of second derivatives in the directions of x, and

$$\omega_1(\delta,M) = \sup_{\substack{x_1,x_2\in[-M,M]^d, y_1,y_2\in[-M,M]^m\\ |x_1-x_2|\wedge|y_1-y_2|<\delta}} \big|\nabla_y f(x_1,y_1) - \nabla_y f(x_2,y_2)\big|,$$

and the modulus of continuity of $\mathrm{Hes}_x f$ by

$$\omega_2(\delta,M) = \sup_{\substack{x_1,x_2\in[-M,M]^d, y_1,y_2\in[-M,M]^m\\ |x_1-x_2|\wedge|y_1-y_2|<\delta}} \|\mathrm{Hes}_x f(x_1,y_1) - \mathrm{Hes}_x f(x_2,y_2)\|,$$

where $\|\cdot\|$ is the operator norm of a matrix. Then argue as in the proof of Theorem 7.13.

Exercise 7.4. First use Brownian scaling and the Markov property, as in the original proof of Theorem 2.33 to reduce the problem to showing that the distribution of $B(T(1))$ (using the notation of Theorem 2.33) is the Cauchy distribution.

The map defined by $f(z) = \frac{z}{2-z}$, for $z\in\mathbb{C}$, takes the half-plane $\{(x,y)\colon x<1\}$ onto the unit disk and $f(0)=0$. The image measure of harmonic measure on $V(1)$ from 0 is the harmonic measure on the unit sphere from the origin, which is uniform. Hence the harmonic measure $\mu_{V(1)}(0,\cdot)$ is the image measure of the uniform distribution ϖ on the unit sphere under f^{-1}, which can be calculated using the derivative of f.

Exercise 7.5. Use that $\theta(t) = W_2(H(t))$ and $\lim_{t\uparrow\infty} H(t) = \infty$.

Exercise 7.7. Suppose h is supported by $[0,b]$ and look at the partitions given by $t_k^{(n)} = bk2^{-n}$, for $k = 0, \dots, 2^n$. By Theorem 7.32 and Theorem 6.18 we can choose a continuous modification of the process $\{\int_0^t \mathrm{sign}(B(s)-a)\, dB(s) \,:\, a \in \mathbb{R}\}$. Hence the Lebesgue integral on the left hand side is also a Riemann integral and can be approximated by the sum

$$\sum_{k=0}^{2^n-1} b2^{-n} h(t_k^{(n)}) \Big(\int_0^t \mathrm{sign}(B(s) - t_k^{(n)})\, dB(s) \Big) = \int_0^t F_n(B(s))\, dB(s),$$

where

$$F_n(x) = \sum_{k=0}^{2^n-1} b2^{-n} h(t_k^{(n)})\, \mathrm{sign}\big(x - t_k^{(n)}\big), \qquad \text{for } n \in \mathbb{N}.$$

This is a uniformly bounded sequence, which is uniformly convergent to the Lebesgue integral

$$F(x) = \int_{-\infty}^{\infty} h(a)\, \mathrm{sign}(x-a)\, da.$$

Therefore the sequence of stochastic integrals converges in L^2 to the stochastic integral $\int_0^\infty F(B(s))\, dB(s)$, which is the right hand side of our formula.

Exercise 8.1. Suppose that u is subharmonic and $\mathcal{B}(x,r) \subset U$. Let τ be the first exit time from $\mathcal{B}(x,r)$, which is a stopping time. As $\Delta u(z) \geq 0$ for all $z \in U$ we see from the multidimensional version of Itô's formula that

$$u(B(t \wedge \tau)) \leq u(B(0)) + \sum_{i=1}^d \int_0^{t\wedge\tau} \frac{\partial u}{\partial x_i}(B(s))\, dB_i(s).$$

Note that $\partial u/\partial x_i$ is bounded on the closure of $\mathcal{B}(x,r)$, and thus everything is well-defined. We can now take expectations, and use Exercise 7.1 to see that

$$\mathbb{E}_x\big[u(B(t\wedge\tau))\big] \leq \mathbb{E}_x\big[u(B(0))\big] = u(x).$$

Now let $t \uparrow \infty$, so that the left hand side converges to $\mathbb{E}_x[u(B(\tau))]$ and note that this gives the mean value property for spheres. The result follows by integrating over r.

Exercise 8.2. Let u be a solution of the Poisson problem on U. Define open sets $U_n \uparrow U$ by

$$U_n = \big\{ x \in U \,:\, |x-y| > \tfrac{1}{n} \text{ for all } y \in \partial U \big\}.$$

Let τ_n be the first exit time of U_n, which is a stopping time. As $\frac{1}{2}\Delta u(x) = -g(x)$ for all $x \in U$ we see from the multidimensional version of Itô's formula that

$$u(B(t\wedge\tau_n)) = u(B(0)) + \sum_{i=1}^d \int_0^{t\wedge\tau_n} \frac{\partial u}{\partial x_i}(B(s))\, dB_i(s) - \int_0^{t\wedge\tau_n} g(B(s))\, ds.$$

Note that $\partial u/\partial x_i$ is bounded on the closure of U_n, and thus everything is well-defined. We can now take expectations, and use Exercise 7.1 to see that

$$\mathbb{E}_x\big[u(B(t\wedge\tau_n))\big] = u(x) - \mathbb{E}_x \int_0^{t\wedge\tau_n} g(B(s))\, ds.$$

Note that both integrands are bounded. Hence, as $t \uparrow \infty$ and $n \to \infty$, bounded convergence yields that

$$u(x) = \mathbb{E}_x \int_0^\tau g(B(s))\, ds,$$

where we have use the boundary condition to eliminate the left hand side.

Exercise 9.2. Use decompositions as in the proof of Theorem 9.22 to transfer the results of Theorem 9.8 from intersections of independent Brownian motions to self-intersections of one Brownian motion.

Exercise 9.4. A counterexample as required in part (d) can be constructed as follows: Let A_1 and A_2 be two disjoint closed sets on the line such that the Cartesian squares A_i^2 have Hausdorff dimension less than 1/2 yet the Cartesian product $A_1 \times A_2$ has dimension strictly greater than 1/2. Let A be the union of A_1 and A_2. Then Brownian motion $\{B(t)\colon t \geq 0\}$ on A is 1-1 with positive probability (if $B(A_1)$ is disjoint from $B(A_2)$) yet with positive probability $B(A_1)$ intersects $B(A_2)$.

For instance let A_1 consist of points in $[0,1]$ where the binary n^{th} digit vanishes whenever $(2k)! \leq n < (2k+1)!$ for some k. Let A_2 consist of points in $[2,3]$ where the binary n^{th} digit vanishes whenever $(2k-1)! \leq n < (2k)!$ for some k. then $\dim(A_i^2) = 0$ for $i = 1,2$ yet $\dim(A_1 \times A_2) \geq \dim(A_1 + A_2) = 1$, in fact $\dim(A_1 \times A_2) = 1$.

Exercise 9.5. Let $\{B_1(t)\colon 0 \leq t \leq 1\}$ be the first component of the planar motion. By Kaufman's theorem, almost surely,

$$\dim S(a) = 2 \dim\{t \in [0,1]\colon B_1(t) = a\}$$

and, as in Corollary 9.30, the dimension on the right equals 1/2 for every $a \in (\min\{x\colon (x,y) \in B[0,t]\}, \max\{x\colon (x,y) \in B[0,t]\})$.

Exercise 10.2. For every decomposition $E = \bigcup_{i=1}^\infty E_i$ of E into bounded sets, we have, using countable stability of Hausdorff dimension,

$$\sup_{i=1}^{\infty} \overline{\dim}_M E_i \geq \sup_{i=1}^{\infty} \dim E_i = \dim \bigcup_{i=1}^{\infty} E_i = \dim E\,,$$

and passing to the infimum yields the statement.

Exercise 10.7 The argument is sketched in [**La99**].

Exercise 10.9. For (a) note that Theorem 10.28 can be read as a criterion to determine the packing dimension of a set E by hitting it with a limsup random fractal. Hence $\dim_P(A \cap E)$ can be found by evaluating $\mathbb{P}\{A \cap A' \cap E = \emptyset\}$ for A' an independent copy of A. Now use that $A \cap A'$ is also a discrete limsup fractal.

Exercise 10.10 To apply Theorem 7.24 for the proof of Lemma 10.40 (a) we shift the cone by defining a new tip $\tilde{z}$ as follows:

- If $\alpha < \pi$ the intersection of the line through x parallel to the central axis of the cone with the boundary of the dual cone,
- if $\alpha > \pi$ the intersection of the line through x parallel to the central axis of the cone with the boundary of the cone.

Note that $z + W[\alpha, \xi] \subset \tilde{z} + W[\alpha, \xi]$ and there exists a constant $C > 1$ depending only on α such that $|z - \tilde{z}| < C\delta$. There is nothing to show if $C\delta > \varepsilon/2$ and otherwise

$$\mathbb{P}_x\big\{B(0, T_\varepsilon(z)) \subset z + W[\alpha, \xi]\big\} \le \mathbb{P}_x\big\{B(0, T_{\varepsilon/2}(\tilde{z})) \subset \tilde{z} + W[\alpha, \xi]\big\}\,.$$

By shifting, rotating and scaling the Brownian motion and by Theorem 7.24 we obtain an upper bound for the right hand side of

$$\mathbb{P}_1\big\{B(0, T_{\frac{\varepsilon}{\delta}(C+\frac{1}{2})^{-1}}(0)) \subset W[\alpha, 0]\big\} = \tfrac{2}{\pi}\,\arctan\big(C_0\,\big(\tfrac{\delta}{\varepsilon}\big)^{\frac{\pi}{\alpha}}\big) \le C_1\,\big(\tfrac{\delta}{\varepsilon}\big)^{\frac{\pi}{\alpha}},$$

where $C_0, C_1 > 0$ are suitable constants.

Appendix II: Background and prerequisites

1. Convergence of distributions on metric spaces

In this section we collect the basic facts about convergence in distribution. While this is a familiar concept for real valued random variables, for example in the central limit theorem, we need a more abstract viewpoint, which allows to study convergence in distribution for random variables with values in metric spaces, like for example function spaces.

If random variables $\{X_n : n \geq 0\}$ converge in distribution, strictly speaking it is their *distributions* and not the *random variables* themselves which converge. This distinguishes convergence in distribution from the types of convergence for random variables, like

- almost sure convergence,
- convergence in probability,
- L^1-convergence (and L^p-convergence).

These types of convergence refer to a sequences of random variables $\{X_n : n \geq 0\}$ converging to a random variable X on the *same* probability space. The values of the approximating sequences lead to conclusions about the values of the limit random variable. This is entirely different for *convergence in distribution*, which we now study. Intuitively if $\{X_n : n \geq 0\}$ converges in distribution to X, this just means that the shape of the distributions of X_n for large n is like the shape of the distribution of X. Sample values from X_n allow no inference towards sample values from X and, indeed, there is no need to define X_n and X on the same probability space. In fact, convergence in distribution is only related to the convergence of the distributions of the random variables and not to the random variables themselves.

We start by giving a definition of convergence in distributions for random variables in metric spaces, explore some of its properties and then show that the concept of convergence in distribution for real-valued random variables is consistent with our definition.

Definition 1.1. *Suppose (E, ρ) is a metric space and $\mathcal{A}$ the Borel-σ-algebra on E. Suppose that X_n and X are E-valued random variables. Then we say that X_n* converges in distribution *to X, if, for every bounded continuous $g : E \to \mathbb{R}$,*

$$\lim_{n\to\infty} \mathbb{E}[g(X_n)] = \mathbb{E}[g(X)].$$

We write $X_n \Rightarrow X$ for convergence in distribution. ◇

Remark 1.2. If $X_n \Rightarrow X$ and $g\colon E \to \mathbb{R}$ is continuous, then $g(X_n) \Rightarrow g(X)$. But note that, if $E = \mathbb{R}$ and $X_n \Rightarrow X$, this does not imply that $\mathbb{E}[X_n]$ converges to $\mathbb{E}[X]$, as $g(x) = x$ is not a bounded function on $\mathbb{R}$. ◇

Here is an alternative approach, which shows that convergence in distribution is in fact a convergence *of* the distributions. The statement of the following proposition is trivial.

PROPOSITION 1.3. *Let $Prob(E)$ be the set of probability measures on $(E, \mathcal{A})$. A sequence $\{P_n : n \geq 0\} \subset Prob(E)$ converges weakly to a limit $P \in Prob(E)$ if, for every continuous, bounded function $g : E \to \mathbb{R}$,*

$$\lim_{n\to\infty} \int g \, dP_n = \int g \, dP \, .$$

Then the limit of a convergent sequence is uniquely determined. Suppose that X_n and X are E-valued random variables. Then X_n converges in distribution to X, if and only if the distributions of X_n converge weakly to the distribution of X.

Proof. Only the uniqueness of the limit needs proof. If P and Q are two limits of the same sequence, then $\int f \, dP = \int f \, dQ$ for all bounded continuous $f : E \to \mathbb{R}$. For every open set $G \subset E$ we may choose an increasing sequence $f_n(x) = n\rho(x, G^c) \wedge 1$ of continuous functions converging to $\mathbb{1}_G$ and infer from monotone convergence that $P(G) = Q(G)$. Now $P = Q$ follows from the Uniqueness theorem for probability measures. ∎

EXAMPLE 1.4.

- Suppose $E = \{1, \ldots, m\}$ is finite and $\rho(x, y) = 1 - \mathbb{1}_{\{x=y\}}$. Then $X_n \Rightarrow X$ if and only if $\lim_{n\to\infty} \mathbb{P}\{X_n = k\} = \mathbb{P}\{X = k\}$ for all $k \in E$.
- Let $E = [0, 1]$ and $X_n = 1/n$ almost surely. Then $X_n \Rightarrow X$, where $X = 0$ almost surely. However, note that $\lim_{n\to\infty} \mathbb{P}\{X_n = 0\} = 0 \neq \mathbb{P}\{X = 0\} = 1$.

◇

THEOREM 1.5. *Suppose a sequence $\{X_n : n \geq 0\}$ of random variables converges almost surely to a random variable X (of course, all on the same probability space). Then X_n converges in distribution to X.*

Proof. Suppose g is bounded and continuous. The $g(X_n)$ converges almost surely to $g(X)$. As the sequence is bounded it is also uniformly integrable, hence convergence holds also in the L^1-sense and this implies convergence of the expectations, i.e. $\mathbb{E}[g(X_n)] \to \mathbb{E}[g(X)]$. ∎

THEOREM 1.6 (Portmanteau theorem). *The following statements are equivalent*

(i) $X_n \Rightarrow X$.
(ii) *For all closed sets $K \subset E$, $\limsup_{n\to\infty} \mathbb{P}\{X_n \in K\} \leq \mathbb{P}\{X \in K\}$.*
(iii) *For all open sets $G \subset E$, $\liminf_{n\to\infty} \mathbb{P}\{X_n \in G\} \geq \mathbb{P}\{X \in G\}$.*
(iv) *For all Borel sets $A \subset E$ with $\mathbb{P}\{X \in \partial A\} = 0$, $\lim_{n\to\infty} \mathbb{P}\{X_n \in A\} = \mathbb{P}\{X \in A\}$.*
(v) *For all bounded measurable functions $g : E \to \mathbb{R}$ with $\mathbb{P}\{g \text{ is discontinuous at } X\} = 0$ we have $\mathbb{E}[g(X_n)] \to \mathbb{E}[g(X)]$.*

Proof. **(i)⇒(ii)** Let $g_n(x) = 1 - (n\rho(x, K) \wedge 1)$, which is continuous and bounded, is 1 on K and converges pointwise to $\mathbb{1}_K$. Then, for every n,

$$\limsup_{k\to\infty} \mathbb{P}\{X_k \in K\} \leq \limsup_{k\to\infty} \mathbb{E}[g_n(X_k)] = \mathbb{E}[g_n(X)]\,.$$

Let $n \to \infty$. The integrand on the right hand side is bounded by 1 and converges pointwise and hence in the L^1-sense to $\mathbb{1}_K(X)$.

(ii)⇒(iii) Follows from $\mathbb{1}_G = 1 - \mathbb{1}_K$ for the closed set $K = G^c$.

(iii)⇒(iv) Let G be the interior and K the closure of A. Then, by assumption, $\mathbb{P}\{X \in G\} = \mathbb{P}\{X \in K\} = \mathbb{P}\{X \in A\}$ and we may use (iii) and (ii) (which follows immediately from (iii)) to get

$$\limsup_{n\to\infty} \mathbb{P}\{X_n \in A\} \leq \limsup_{n\to\infty} \mathbb{P}\{X_n \in K\} \leq \mathbb{P}\{X \in K\} = \mathbb{P}\{X \in A\},$$
$$\liminf_{n\to\infty} \mathbb{P}\{X_n \in A\} \geq \liminf_{n\to\infty} \mathbb{P}\{X_n \in G\} \geq \mathbb{P}\{X \in G\} = \mathbb{P}\{X \in A\}.$$

(iv)⇒(v) From (iv) we infer that the convergence holds for g of the form $g(x) = \sum_{n=1}^N a_n \mathbb{1}_{A_n}$ where A_n satisfies $\mathbb{P}\{X \in A_n\} = 0$. Let us call such functions elementary. Given g as in (v) we observe that for every $a < b$ with possibly a countable set of exceptions

$$\mathbb{P}\big\{X \in \partial\{x : g(x) \in (a, b]\}\big\} = 0\,.$$

Indeed, if $X \in \partial\{x : g(x) \in (a, b]\}$ then either g is discontinuous in X or $g(X) = a$ or $g(X) = b$. The first event has probability zero and so have the last two except possibly for a countable set of values of a, b. By decomposing the real axis in small suitable intervals we thus obtain an increasing sequence g_n and a decreasing sequence h_n of elementary functions both converging pointwise to g. Now, for all k,

$$\limsup_{n\to\infty} \mathbb{E}[g(X_n)] \leq \limsup_{n\to\infty} \mathbb{E}[h_k(X_n)] = \mathbb{E}[h_k(X)]\,,$$

and

$$\liminf_{n\to\infty} \mathbb{E}[g(X_n)] \geq \liminf_{n\to\infty} \mathbb{E}[g_k(X_n)] = \mathbb{E}[g_k(X)]\,.$$

and the right sides converge, as $k \to \infty$, by bounded convergence, to $\mathbb{E}[g(X)]$.

(v)⇒(i) This is trivial. ∎

To remember the directions of the inequalities in the Portmanteau theorem it is useful to recall the last example $X_n = 1/n \to 0$ and choose $G = (0, 1)$ and $K = \{0\}$ to obtain cases where the opposite inequalities fail.

We now show that the convergence of distribution as defined here agrees with the familiar concept in the case of real random variables.

Theorem 1.7 (Helly-Bray theorem). *Let X_n and X be real valued random variables and define the associated distribution functions $F_n(x) = \mathbb{P}\{X_n \leq x\}$ and $F(x) = \mathbb{P}\{X \leq x\}$. Then the following assertions are equivalent.*

(a) *X_n converges in distribution to X,*
(b) *$\lim_{n\to\infty} F_n(x) = F(x)$ for all x such that F is continuous in x.*

Proof. (a)⇒(b) Use property (iv) for the set $A = (-\infty, x]$.
(b)⇒(a) We choose a dense sequence $\{x_n\}$ with $\mathbb{P}\{X = x_n\} = 0$ and note that every open set $G \subset \mathbb{R}$ can be written as the countable union of disjoint intervals $I_k = (a_k, b_k]$ with a_k, b_k chosen from the sequence. We have

$$\lim_{n\to\infty} \mathbb{P}\{X_n \in I_k\} = \lim_{n\to\infty} F_n(b_k) - F_n(a_k) = F(b_k) - F(a_k) = \mathbb{P}\{X \in I_k\}\,.$$

Hence, for all N,

$$\liminf_{n\to\infty} \mathbb{P}\{X_n \in G\} \geq \sum_{k=1}^{N} \liminf_{n\to\infty} \mathbb{P}\{X_n \in I_k\} = \sum_{k=1}^{N} \mathbb{P}\{X \in I_k\}\,,$$

and as $N \to \infty$ the last term converges to $\mathbb{P}\{X \in G\}$. ∎

Finally, we note the useful fact that for nonnegative random variables X_n, rather then testing convergence of $\mathbb{E}[f(X_n)]$ for *all* continuous bounded functions f, it suffices to consider functions of a rather simple form.

PROPOSITION 1.8. *Suppose* $(X_1^{(n)}, \ldots, X_m^{(n)})$ *are random vectors with nonnegative entries, then*

$$(X_1^{(n)}, \ldots, X_m^{(n)}) \Longrightarrow (X_1, \ldots, X_m)\,,$$

if an only if, for any $\lambda_1, \ldots, \lambda_m \geq 0$,

$$\lim_{n\uparrow\infty} \mathbb{E}\Big[\exp\Big\{-\sum_{j=1}^{m} \lambda_j X_j^{(n)}\Big\}\Big] = \mathbb{E}\Big[\exp\Big\{-\sum_{j=1}^{m} \lambda_j X_j\Big\}\Big]\,.$$

The function $\phi(\lambda_1, \ldots, \lambda_m) = \mathbb{E}[\exp\{-\sum_{j=1}^{m} \lambda_j X_j\}]$ is called the Laplace transform of $(X_1, \ldots, X_m)$ and thus the proposition states in other words that the convergence of nonnegative random vectors is equivalent to convergence of their Laplace transforms. The proof, usually done by approximation, can be found in...

2. Gaussian random variables

In this section we have collected the facts about Gaussian random vectors, which are used in this book. We start with a useful estimate for standard normal random variables, which is quite precise for large x.

LEMMA 3.1. *Suppose X is standard normally distributed. Then, for all $x > 0$,*

$$\frac{x}{x^2+1}\frac{1}{\sqrt{2\pi}}\,e^{-x^2/2} \leq \mathbb{P}\{X > x\} \leq \frac{1}{x}\frac{1}{\sqrt{2\pi}}\,e^{-x^2/2}.$$

Proof. The right inequality is obtained by the estimate

$$\mathbb{P}\{X > x\} \leq \frac{1}{\sqrt{2\pi}} \int_x^\infty \frac{u}{x}\,e^{-u^2/2}\,du = \frac{1}{x}\frac{1}{\sqrt{2\pi}} e^{-x^2/2}\,.$$

For the left inequality we define

$$f(x) = xe^{-x^2/2} - (x^2+1)\int_x^\infty e^{-u^2/2}\,du\,.$$

Observe that $f(0) < 0$ and $\lim_{x\to\infty} f(x) = 0$. Moreover,

$$f'(x) = (1 - x^2 + x^2 + 1)e^{-x^2/2} - 2x \int_x^\infty e^{-u^2/2}\,du = -2x\Big(\int_x^\infty e^{-u^2/2}\,du - \frac{e^{-x^2/2}}{x}\Big),$$

which is positive for $x > 0$, by the first part. Hence $f(x) \le 0$, proving the lemma. ■

We now look more closely at random vectors with normally distributed components. Our motivation is that they arise, for example, as vectors consisting of the increments of a Brownian motion. Let us clarify some terminology.

Definition 3.2. *A random variable $X = (X_1, \ldots, X_d)^{\mathrm{T}}$ with values in $\mathbb{R}^d$ has the* d-dimensional standard Gaussian distribution *if its d coordinates are standard normally distributed and independent.* ◇

More general Gaussian distributions can be derived as linear images of standard Gaussians. Recall, e.g. from Definition 1.2 in Chapter 1, that a random variable Y with values in $\mathbb{R}^d$ is called *Gaussian* if there exists an m-dimensional standard Gaussian X, an $d \times m$ matrix A, and a d dimensional vector b such that $Y^{\mathrm{T}} = AX + b$. The *covariance matrix* of the (column) vector Y is then given by

$$\mathrm{Cov}(Y) = \mathbb{E}\big[(Y - \mathbb{E}Y)(Y - \mathbb{E}Y)^{\mathrm{T}}\big] = AA^{\mathrm{T}},$$

where the expectations are defined componentwise.

Our next lemma shows that applying an orthogonal $d\times d$ matrix does not change the distribution of a standard Gaussian random vector, and in particular that the standard Gaussian distribution is rotationally invariant. We write I_d for the $d \times d$ identity matrix.

Lemma 3.3. *If A is an orthogonal $d \times d$ matrix, i.e. $AA^{\mathrm{T}} = I_d$, and X is a d-dimensional standard Gaussian vector, then AX is also a d-dimensional standard Gaussian vector.*

Proof. As the coordinates of X are independent, standard normally distributed, X has a density

$$f(x_1, \ldots, x_d) = \prod_{i=1}^d \frac{1}{\sqrt{2\pi}}\, e^{-x_i^2/2} = \frac{1}{(2\pi)^{d/2}}\, e^{-|x|^2/2},$$

where $|\cdot|$ is the Euclidean norm. The density of AX is (by the transformation rule) $f(A^{-1}x)\,|\det(A^{-1})|$. The determinant is 1 and, since orthogonal matrices preserve the Euclidean norm, the density of X is invariant under A. ■

Corollary 3.4. *Let X_1 and X_2 be independent and normally distributed with expectation* 0 *and variance $\sigma^2 > 0$. Then $X_1 + X_2$ and $X_1 - X_2$ are independent and normally distributed with expectation* 0 *and variance* $2\sigma^2$.

Proof. The vector $(X_1/\sigma, X_2/\sigma)^{\mathrm{T}}$ is standard Gaussian by assumption. Look at

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

This is an orthogonal matrix and applying it to our vector yields $((X_1 + X_2)/(\sqrt{2}\sigma), (X_1 - X_2)/(\sqrt{2}\sigma))$, which thus must have independent standard normal coordinates. ∎

The next proposition shows that the distribution of a Gaussian random vector is determined by its expectation and covariance matrix.

PROPOSITION 3.5. *If X and Y are d-dimensional Gaussian vectors with $\mathbb{E}X = \mathbb{E}Y$ and $\mathrm{Cov}(X) = \mathrm{Cov}(Y)$, then X and Y have the same distribution.*

Proof. It is sufficient to consider the case $\mathbb{E}X = \mathbb{E}Y = 0$. By definition, there are standard Gaussian random vectors X_1 and X_2 and matrices A and B with $X = AX_1$ and $Y = BX_2$. By adding columns of zeros to A or B, if necessary, we can assume that X_1 and X_2 are both k-vectors, for some k, and A, B are both $d \times k$ matrices. Let $\mathcal{A}$ and $\mathcal{B}$ be the vector subspaces of $\mathbb{R}^k$ generated by the row vectors of A and B, respectively. To simplify notation assume that the first $l \leq d$ row vectors of A form a basis of $\mathcal{A}$. Define the linear map $L\colon \mathcal{A} \to \mathcal{B}$ by

$$L(A_i) = B_i \text{ for } i = 1, \ldots, l.$$

Here A_i is the i^{th} row vector of A, and B_i is the i^{th} row vector of B. Our aim is to show that L is an orthogonal isomorphism and then use the previous proposition. Let us first show that L is an isomorphism. Our covariance assumption gives that $AA^{\mathrm{T}} = BB^{\mathrm{T}}$. Assume there is a vector $v_1A_1 + \ldots v_lA_l$ whose image is 0. Then the d-vector

$$v = (v_1, \ldots, v_l, 0, \ldots, 0)$$

satisfies $vB = 0$. Hence

$$\|vA\|^2 = vAA^{\mathrm{T}}v^{\mathrm{T}} = vBB^{\mathrm{T}}v^{\mathrm{T}} = 0\,.$$

We conclude that $vA = 0$. Hence L is injective and $\dim \mathcal{A} \leq \dim \mathcal{B}$. Interchanging the roles of A and B gives that L is an isomorphism. As the entry (i, j) of $AA^{\mathrm{T}} = BB^{\mathrm{T}}$ is the scalar product of A_i and A_j as well as B_i and B_j, the mapping L is orthogonal. We can extend it on the orthocomplement of $\mathcal{A}$ to an orthogonal map $L\colon \mathbb{R}^k \to \mathbb{R}^k$ (or an orthogonal $k \times k$-matrix). Then $X = AX_1$ and $Y = BX_2 = AL^{\mathrm{T}}X_2$. As $L^{\mathrm{T}}X_2$ is standard Gaussian, by Lemma 3.3, X and Y have the same distribution. ∎

In particular, comparing a d-dimensional Gaussian vector with $\mathrm{Cov}(X) = I_d$ with a Gaussian vector with d independent entries and the same expectation, we obtain the following fact.

COROLLARY 3.6. *A Gaussian random vector X has independent entries if and only if its covariance matrix is diagonal. In other words, the entries in a Gaussian vector are uncorrelated if and only if they are independent.*

We now show that the Gaussian nature of a random vector is preserved under taking limits.

PROPOSITION 3.7. *Suppose $\{X_n : n \in \mathbb{N}\}$ is a sequence of Gaussian random vectors and $\lim_n X_n = X$, almost surely. If $b := \lim_{n\to\infty} \mathbb{E}X_n$ and $C := \lim_{n\to\infty} \mathrm{Cov}\, X_n$ exist, then X is Gaussian with mean b and covariance matrix C.*

Proof. A variant of the argument in Proposition 3.5 shows that X_n converges in law to a Gaussian random vector with mean b and covariance matrix C. As almost sure convergence implies convergence of the associated distributions, this must be the law of X. ∎

Lemma 3.8. *Suppose X, Y are independent and normally distributed with mean zero and variance σ^2, then X^2+Y^2 is exponentially distributed with mean $2\sigma^2$.*

Proof. For any bounded, measurable $f\colon \mathbb{R}\to\mathbb{R}$ we have, using polar coordinates,

$$\begin{aligned}\mathbb{E}f(X^2+Y^2) &= \frac{1}{2\pi\sigma^2}\int f(x^2+y^2)\exp\big\{-\tfrac{x^2+y^2}{2\sigma^2}\big\}\,dx\,dy = \frac{1}{\sigma^2}\int_0^\infty f(r^2)\exp\big\{-\tfrac{r^2}{2\sigma^2}\big\}\,r\,dr\\ &= \frac{1}{2\sigma^2}\int_0^\infty f(a)\exp\big\{-\tfrac{a}{2\sigma^2}\big\}\,da = \mathbb{E}f(Z)\,,\end{aligned}$$

where Z is exponential with mean $2\sigma^2$. ∎

3. Martingales in discrete time

In this section we recall the essentials from the theory of martingales in discrete time. A more thorough introduction to this delightful subject is Williams [**Wi91**].

Definition 4.1. *A* filtration $(\mathcal{F}_n : n\geq 0)$ *is an increasing sequence*

$$\mathcal{F}_0\subset\mathcal{F}_1\subset\cdots\subset\mathcal{F}_n\subset\cdots$$

of σ-algebras. Let $\{X_n : n\geq 0\}$ be a stochastic process in discrete time and $(\mathcal{F}_n : n\geq 0)$ be a filtration. The process is a martingale *relative to the filtration if, for all $n\geq 0$,*

- *X_n is measurable with respect to $\mathcal{F}_n$,*
- *$\mathbb{E}|X_n|<\infty$, and*
- *$\mathbb{E}[X_{n+1}\,|\,\mathcal{F}_n]=X_n$, almost surely.*

If we have '$\geq$' in the last condition, then $\{X_n\colon n\geq 0\}$ is called a submartingale, *if '$\leq$' holds it is called a* supermartingale. ⋄

Remark 4.2. Note that for a submartingale $\mathbb{E}[X_{n+1}]\geq\mathbb{E}[X_n]$, for a supermartingale $\mathbb{E}[X_{n+1}]\leq\mathbb{E}[X_n]$, and hence for a martingale we have $\mathbb{E}[X_{n+1}]=\mathbb{E}[X_n]$. ⋄

Loosely speaking, a stopping time is a random time such that the knowledge about a random process at time n suffices to determine whether the stopping time has happened at time n or not. Here is a formal definition.

Definition 4.3. *A random variable T with values in $\{0,1,2,\ldots\}\cup\{\infty\}$ is called a* stopping time *if $\{T\leq n\}=\{\omega : T(\omega)\leq n\}\in\mathcal{F}_n$ for all $n\geq 0$.* ⋄

If $\{X_n : n \geq 0\}$ is a supermartingale and T a stopping time, then it is easy to check that the process

$$\{X_n^T : n \geq 0\} \qquad \text{defined by } X_n^T = X_{T \wedge n}$$

is a supermartingale. If $\{X_n : n \geq 0\}$ is a martingale, then both $\{X_n : n \geq 0\}$ and $\{-X_n : n \geq 0\}$ are supermartingales and, hence, we have,

$$\mathbb{E}\big[X_{T \wedge n}\big] = \mathbb{E}\big[X_0\big], \qquad \text{for all } n \geq 0\,.$$

Doob's optional stopping theorem gives criteria when, letting $n \uparrow \infty$, we obtain $\mathbb{E}[X_T] = \mathbb{E}[X_0]$.

THEOREM 4.4 (Doob's optional stopping theorem). *Let T be a stopping time and X a martingale. Then X_T is integrable and $\mathbb{E}\big[X_T\big] = \mathbb{E}\big[X_0\big]$, if one of the following conditions hold:*

(1) *T is bounded, i.e. there is N such that $T < N$ almost surely;*

(2) *$\{X_n^T : n \geq 0\}$ is dominated by an integrable random variable Z, i.e. $|X_{n \wedge T}| \leq Z$ for all $n \geq 0$ almost surely;*

(3) *$\mathbb{E}[T] < \infty$ and there is $K > 0$ such that $\sup_n |X_n - X_{n-1}| \leq K$.*

Proof. Recall that $\mathbb{E}[X_{T \wedge n} - X_0] = 0$. The result follows in case (1) by choosing $n = N$. In case (2) let $n \to \infty$ and use dominated convergence. In case (3) observe that $|X_{T \wedge n} - X_0| = |\sum_{k=1}^{T \wedge n}(X_k - X_{k-1})| \leq KT$. By assumption KT is an integrable function and dominated convergence can be used again. ■

Doob's famous forward convergence theorem gives a sufficient condition for the almost sure convergence of supermartingales to a limiting random variable. See [**Wi91**, 11.5] for the proof.

THEOREM 4.5 (Doob's supermartingale convergence theorem). *Let $\{X_n : n \geq 0\}$ be a supermartingale, which is bounded in L^1, i.e. there is $K > 0$ such that $\mathbb{E}|X_n| \leq K$ for all n. Then there exists an integrable random variable X on the same probability space such that*

$$\lim_{n \to \infty} X_n = X \textit{ almost surely.}$$

REMARK 4.6. Note that if $\{X_n : n \geq 0\}$ is nonnegative, we have $\mathbb{E}[|X_n|] = \mathbb{E}[X_n] \leq \mathbb{E}[X_0] := K$ and thus X_n is automatically bounded in L^1 and $\lim_{n \to \infty} X_n = X$ exists. ◇

A key question is when the almost sure convergence in the supermartingale convergence theorem can be replaced by L^1-convergence (which in contrast to almost sure convergence implies convergence of expectations). A necessary and sufficient criterion for this is *uniform integrability.* A stochastic process $\{X_n : n \geq 0\}$ is called *uniformly integrable* if, for every $\varepsilon > 0$, there exists $K > 0$ such that

$$\mathbb{E}\big[|X_n| \mathbb{1}\{|X_n| \geq K\}\big] < \varepsilon \qquad \text{for all } n \geq 0\,.$$

Sufficient criteria for uniform integrability are

- $\{X_n : n \geq 0\}$ is dominated by an integrable random variable,
- $\{X_n : n \geq 0\}$ is L^p-bounded for some $p > 1$,
- $\{X_n : n \geq 0\}$ is L^1-convergent.

The following lemma is proved in [**Wi91**, 13.7].

LEMMA 4.7. *Any stochastic process $\{X_n : n \geq 0\}$, which is uniformly integrable and almost surely convergent, converges also in the L^1-sense.*

The next result is one of the highlights of martingale theory.

THEOREM 4.8 (Martingale closure theorem). *Suppose that the martingale $\{X_n : n \geq 0\}$ is uniformly integrable. Then there is an integrable random variable X such that*

$$\lim_{n\to\infty} X_n = X \textit{ almost surely and in } L^1.$$

Moreover, $X_n = \mathbb{E}[X \,|\, \mathcal{F}_n]$ for every $n \geq 0$.

Proof. Uniform integrability implies that $\{X_n : n \geq 0\}$ is L^1-bounded and thus, by the martingale convergence theorem, almost surely convergent to an integrable random variable X. Convergence in the L^1-sense follow from Lemma II.4.7. To check the last assertion, we note that X_n is $\mathcal{F}_n$-measurable and let $F \in \mathcal{F}_n$. For all $m \geq n$ we have, by the martingale property, $\int_F X_m \, d\mathbb{P} = \int_F X_n \, d\mathbb{P}$. We let $m \to \infty$. Then $| \int_F X_m \, d\mathbb{P} - \int_F X \, d\mathbb{P}| \leq \int |X_m - X| \, d\mathbb{P} \to 0$, hence we obtain $\int_F X \, d\mathbb{P} = \int_F X_n \, d\mathbb{P}$, as required. ∎

There is a natural converse to the martingale closure theorem, see [**Wi91**, 14.2] for the proof.

THEOREM 4.9 (Lévy's upward theorem). *Suppose that X is an integrable random variable and $X_n = \mathbb{E}[X \,|\, \mathcal{F}_n]$. Then $\{X_n : n \geq 0\}$ is a uniformly integrable martingale and*

$$\lim_{n\to\infty} X_n = \mathbb{E}\big[X \mid \mathcal{F}_\infty\big] \textit{ almost surely and in } L^1,$$

where $\mathcal{F}_\infty = \big(\bigcup_{n=1}^\infty \mathcal{F}_n\big)$ is the smallest σ-algebra containing the entire filtration.

There is also a convergence theorem for 'reverse' martingales, which is called Lévy's downward theorem and is a natural partner to the upward theorem, see [**Wi91**, 14.4] for the proof.

THEOREM 4.10 (Lévy's downward theorem). *Suppose that $(\mathcal{G}_n : n \in \mathbb{N})$ is a collection of σ-algebras such that*

$$\mathcal{G}_\infty := \bigcap_{k=1}^\infty \mathcal{G}_k \subset \cdots \subset \mathcal{G}_{n+1} \subset \mathcal{G}_n \subset \cdots \subset \mathcal{G}_1.$$

An integrable process $\{X_n : n \in \mathbb{N}\}$ is a **reverse martingale** *if almost surely,*

$$X_n = \mathbb{E}\big[X_{n-1} \,\big|\, \mathcal{G}_n\big] \qquad \textit{for all } n \geq 2\,.$$

Then

$$\lim_{n\uparrow\infty} X_n = \mathbb{E}[X_1 \,|\, \mathcal{G}_\infty] \qquad \textit{almost surely.}$$

An important consequence of Theorems II.4.4 and II.4.8 is that the martingale property holds for well-behaved stopping times. For a stopping time T define $\mathcal{F}_T$ to be the σ-algebra of events A with $A \cap \{T \le n\} \in \mathcal{F}_n$. Observe that X_T is $\mathcal{F}_T$-measurable.

THEOREM 4.11 (Optional sampling theorem). *If the martingale $\{X_n : t \ge 0\}$ is uniformly integrable, then for all stopping times $0 \le S \le T$ we have $\mathbb{E}[X_T \,|\, \mathcal{F}_S] = X_S$ almost surely.*

Proof. By the martingale closure theorem, X_n^T converges to X_T in L^1 and $\mathbb{E}[X_T \,|\, \mathcal{F}_n] = X_{T\wedge n} = X_n^T$. Dividing X_T in its positive and its nonpositive part if necessary, we may assume that $X_T \ge 0$ and therefore $X_n^T \ge 0$ almost surely. Taking conditional expectation with respect to $\mathcal{F}_{S\wedge n}$ gives $\mathbb{E}\big[X_T \,\big|\, \mathcal{F}_{S\wedge n}\big] = X_{S\wedge n}$. Now let $A \in \mathcal{F}_S$. We have to show that $\mathbb{E}[X_T \mathbb{1}_A] = \mathbb{E}[X_S \mathbb{1}_A]$. Note first that $A \cap \{S \le n\} \in \mathcal{F}_{S\wedge k}$. Hence, we get $\mathbb{E}[X_T \mathbb{1}\{A \cap \{S \le n\}\}] = \mathbb{E}[X_{S\wedge n} \mathbb{1}\{A \cap \{S \le n\}\}] = \mathbb{E}[X_S \mathbb{1}\{A \cap \{S \le n\}\}]$. Letting $n \uparrow \infty$ and using monotone convergence gives the required result. ■

Of considerable practical importance are martingales $\{X_n : t \ge 0\}$, which are *square integrable.* Note that in this case we can calculate, for $m \ge n$,

$$\begin{aligned} \mathbb{E}\big[X_m^2 \mid \mathcal{F}_n\big] &= \mathbb{E}\big[(X_m - X_n)^2 \mid \mathcal{F}_n\big] + 2\mathbb{E}[X_m \mid \mathcal{F}_n]\, X_n - X_n^2 \\ &= \mathbb{E}\big[(X_m - X_n)^2 \mid \mathcal{F}_n\big] + X_n^2 \ge X_n^2\,, \end{aligned} \tag{3.1}$$

so that $\{X_n^2 : t \ge 0\}$ is a submartingale.

THEOREM 4.12 (Convergence theorem for L^2-bounded martingales). *Suppose that the martingale $\{X_n : t \ge 0\}$ is L^2-bounded. Then there is a random variable X such that*

$$\lim_{n\to\infty} X_n = X \text{ almost surely and in } L^2\,.$$

Proof. From (3.1) and L^2-boundedness of $\{X_n : t \ge 0\}$ it is easy to see that, for $m \ge n$,

$$\mathbb{E}\big[(X_m - X_n)^2\big] = \sum_{k=n+1}^{m} \mathbb{E}\big[(X_k - X_{k-1})^2\big] \le \sum_{k=1}^{\infty} \mathbb{E}\big[(X_k - X_{k-1})^2\big] < \infty\,.$$

Recall that L^2-boundedness implies L^1-boundedness, and hence, by the martingale convergence theorem, X_n converges almost surely to an integrable random variable X. Letting $m \uparrow \infty$ and using Fatou's lemma in the last display, gives L^2-convergence. ■

We now discuss two *martingale inequalities* that have important counterparts in the continuous setting. The first one is Doob's weak maximal inequality.

THEOREM 4.13 (Doob's weak maximal inequality). *Let $\{X_j : j \ge 0\}$ be a submartingale and denote $M_n := \max_{1\le j\le n} X_j$. Then, for all $\lambda > 0$,*

$$\lambda \mathbb{P}\big\{M_n \ge \lambda\big\} \le \mathbb{E}\big[X_n \mathbb{1}\{M_n \ge \lambda\}\big]\,.$$

Proof. Define the stopping time

$$\tau := \begin{cases} \min\{k : X_k \geq \lambda\} & \text{if } M_n \geq \lambda \\ n & \text{if } M_n < \lambda \end{cases}$$

Note that $\{M_n \geq \lambda\} = \{X_\tau \geq \lambda\}$. This implies

$$\lambda \mathbb{P}\{M_n \geq \lambda\} = \lambda \mathbb{P}\{X_\tau \geq \lambda\} = \mathbb{E}\lambda \mathbb{1}\{X_\tau \geq \lambda\} \leq \mathbb{E}X_\tau \mathbb{1}\{X_\tau \geq \lambda\} = \mathbb{E}X_\tau \mathbb{1}\{M_n \geq \lambda\},$$

and the result follows once we demonstrate $\mathbb{E}X_\tau \mathbb{1}\{M_n \geq \lambda\} \leq \mathbb{E}X_n \mathbb{1}\{M_n \geq \lambda\}$. But, as τ is bounded by n and X^τ is a submartingale, we have $\mathbb{E}[X_\tau] \leq \mathbb{E}[X_n]$, which implies that

$$\mathbb{E}\big[X_\tau \mathbb{1}\{M_n < \lambda\}\big] + \mathbb{E}\big[X_\tau \mathbb{1}\{M_n \geq \lambda\}\big] \leq \mathbb{E}\big[X_n \mathbb{1}\{M_n < \lambda\}\big] + \mathbb{E}\big[X_n \mathbb{1}\{M_n \geq \lambda\}\big].$$

Because, by definition of τ, we have $X_\tau \mathbb{1}\{M_n < \lambda\} = X_n \mathbb{1}\{M_n < \lambda\}$, this reduces to

$$\mathbb{E}\big[X_\tau \mathbb{1}\{M_n \geq \lambda\}\big] \leq \mathbb{E}\big[X_n \mathbb{1}\{M_n \geq \lambda\}\big],$$

and this concludes the proof. ∎

The most useful martingale inequality for us is Doob's L^p-maximal inequality.

THEOREM 4.14 (Doob's L^p maximal inequality). *Suppose $\{X_n \colon n \geq 0\}$ is a submartingale. Let $M_n = \max_{1 \leq k \leq n} X_k$ and $p > 1$. Then*

$$\mathbb{E}\big[M_n^p\big] \leq \big(\tfrac{p}{p-1}\big)^p \, \mathbb{E}\big[|X_n|^p\big].$$

We make use of the following lemma, which allows us to compare the L^p-norms of two nonnegative random variables.

LEMMA 4.15. *Suppose nonnegative random variables X and Y satisfy, for all $\lambda > 0$,*

$$\lambda \mathbb{P}\{Y \geq \lambda\} \leq \mathbb{E}[X \mathbb{1}\{Y \geq \lambda\}].$$

Then, for all $p > 1$,

$$\mathbb{E}\big[Y^p\big] \leq \big(\tfrac{p}{p-1}\big)^p \, \mathbb{E}\big[X^p\big].$$

Proof. Using the fact that $X \geq 0$ and $x^p = \int_0^x p\lambda^{p-1} d\lambda$, we can express $\mathbb{E}[X^p]$ as a double integral and apply Fubini's theorem,

$$\mathbb{E}\big[X^p\big] = \mathbb{E}\int_0^\infty \mathbb{1}\{X \geq \lambda\}\, p\lambda^{p-1}\, d\lambda = \int_0^\infty p\lambda^{p-1}\, \mathbb{P}\{X \geq \lambda\}\, d\lambda.$$

Similarly, using the hypothesis,

$$\mathbb{E}\big[Y^p\big] = \int_0^\infty p\lambda^{p-1}\, \mathbb{P}\big\{Y \geq \lambda\big\}\, d\lambda \leq \int_0^\infty p\lambda^{p-2} \mathbb{E}\big[X \mathbb{1}\{Y \geq \lambda\}\big]\, d\lambda.$$

We can rewrite the right hand side, using Fubini's theorem again, and then integrating $p\lambda^{p-2}$ and using Hölder's inequality with $q = p/(p-1)$,

$$\int_0^\infty p\lambda^{p-2}\, \mathbb{E}\big[X \mathbb{1}\{Y \geq \lambda\}\big]\, d\lambda = \mathbb{E}\Big[X \int_0^Y p\lambda^{p-2}\, d\lambda\Big] \; = q\, \mathbb{E}\big[XY^{p-1}\big] \leq q\, \|X\|_p \|Y^{p-1}\|_q.$$

Altogether, this gives $\mathbb{E}[Y^p] \leq q(\mathbb{E}[X^p])^{1/p} (\mathbb{E}[Y^p])^{1/q}$ So, assuming $\mathbb{E}[Y^p] < \infty$, the above inequality gives,

$$\big(\mathbb{E}[Y^p]\big)^{1/p} \leq q\big(\mathbb{E}[X^p]\big)^{1/p},$$

from which the result follows by raising both sides to the p^{th} power. In general, if $\mathbb{E}[Y^p] = \infty$, then for any $n \in \mathbb{N}$, the random variable $Y_n = Y \wedge n$ satisfies the hypothesis of the lemma, and the result follows by letting $n \uparrow \infty$ and applying the monotone convergence theorem. ∎

Proof of Theorem 4.14. If $\{X_n : n \geq 0\}$ is a submartingale, so is $\{|X_n| : n \geq 0\}$. Hence we may assume in the proof that $X_n \geq 0$. By Doob's weak maximal inequality,

$$\lambda \mathbb{P}\big\{M_n \geq \lambda\big\} \leq \mathbb{E}\big[X_n \mathbb{1}\{M_n \geq \lambda\}\big],$$

and applying Lemma 4.15 with $X = X_n$ and $Y = M_n$ gives the result. ∎

4. The max-flow min-cut theorem

Here we give a short proof of a famous result of graph theory, the *max-flow min-cut theorem* of Ford and Fulkerson [**FF56**] in the special case of infinite trees.

Theorem 5.1 (Max-flow min-cut theorem).

$$\max\Big\{\text{strength}\,(\theta) \,:\, \theta \textit{ a flow with capacities } C\Big\} = \inf\Big\{\sum_{e\in\Pi} C(e) \,:\, \Pi \textit{ a cutset}\Big\}.$$

Proof. The proof is a festival of compactness arguments.

First observe that on the left hand side the infimum is indeed a maximum, because if $\{\theta_n\}$ is a sequence of flows with capacities C, then at every edge we have a bounded sequence $\{\theta_n(e)\}$ and by the diagonal argument we may pass to a subsequence such that $\lim \theta_n(e)$ exists simultaneously for all $e \in E$. This limit is obviously again a flow with capacities C.

Secondly observe that every cutset Π contains a finite subset $\Pi' \subset \Pi$, which is still a cutset. Indeed, if this was not the case, we had for every positive integer j a ray $e_1^j, e_2^j, e_3^j, \ldots$ with $e_i^j \notin \Pi$ for all $i \leq j$. By the diagonal argument we find a sequence j_k and edges e_l of order l such that $e_l^{j_k} = e_l$ for all $k \geq l$. Then $e_1, e_2, \ldots$ is a ray and $e_l \notin \Pi$ for all l, which is a contradiction.

Now let θ be a flow with capacities C and Π an arbitrary cutset. We let A be the set of vertices v such that there is a sequence of edges $e_1, \ldots, e_n \notin \Pi$ with $e_1 = (\rho, v_1)$, $e_n = (v_{n-1}, v)$ and $e_j = (v_{j-1}, v_j)$. By our previous observation this set is finite. Let

$$\phi(v,e) := \begin{cases} 1 & \text{if } e = (v,w) \text{ for some } w \in V\,, \\ -1 & \text{if } e = (w,v) \text{ for some } w \in V\,. \end{cases}$$

Then, using the definition of a flow and finiteness of all sums,

$$\begin{aligned}\text{strength}\,(\theta) &= \sum_{e\in E}\phi(\rho,e)\theta(e) = \sum_{v\in A}\sum_{e\in E}\phi(v,e)\theta(e)\\ &= \sum_{e\in E}\theta(e)\sum_{v\in A}\phi(v,e) \le \sum_{e\in\Pi}\theta(e)\\ &\le \sum_{e\in\Pi}C(e)\,.\end{aligned}$$

This proves the first inequality.

For the reverse inequality we restrict attention to finite trees. Let T_n be the tree consisting of all vertices V_n and edges E_n of order $\le n$ and look at cutsets Π consisting of vertices in E_n. A flow of strength $c > 0$ through the tree T_n with capacities C is a mapping $\theta : E_n \to [0, c]$ such that

- for the root we have $\sum_{w\,:\,\overline{w}=\rho}\theta\big(\rho,w\big) = c$,
- for every vertex $v \ne \rho$ with $|v| < n$ we have $\theta\big(\overline{v}, v\big) = \sum_{w\,:\,\overline{w}=v}\theta\big(v,w\big)$,
- $\theta(e) \le C(e)$.

We shall show that

$$\begin{aligned}&\max\Big\{\text{strength}\,(\theta)\ :\ \theta \text{ a flow in } T_n \text{ with capacities } C\Big\}\\ &\ge\ \min\Big\{\sum_{e\in\Pi}C(e)\ :\ \Pi \text{ a cutset in } T_n\Big\}.\end{aligned}\tag{4.1}$$

Once we have this, we get a sequence (θ_n) of flows in T_n with capacities C and strength at least $c = \min\{\sum_{e\in\Pi}C(e)\ :\ \Pi \text{ a cutset in } T\}$. By using the diagonal argument once more we can get a subsequence such that the limits of $\theta_n(e)$ exist for every edge, and the result is a flow θ with capacities C and strength at least c, as required.

To prove (4.1) let θ be a flow of maximal strength c with capacities C in T_n and call a sequence $\rho = v_0, v_1, \ldots, v_n$ with $(v_i, v_{i+1}) \in E_n$ an *augmenting sequence* if $\theta(v_i, v_{i+1}) < C(v_i, v_{i+1})$. If there are augmenting sequences, we can construct a flow $\widetilde{\theta}$ of strength $> c$ by just increasing the flow through every edge of the augmenting sequence by a sufficiently small $\varepsilon > 0$. As θ was maximal this is a contradiction. Hence there is a minimal cutset Π consisting entirely of edges in E_n with $\theta(e) \ge C(e)$. Let A, as above, be the collection of all edges which are connected to the root by edges not in Π. As before, we have

$$\text{strength}\,(\theta) = \sum_{e\in E}\theta(e)\sum_{v\in A}\phi(v,e) = \sum_{e\in\Pi}\theta(e) \ge \sum_{e\in\Pi}C(e)\,,$$

where in the penultimate step we use minimality. This proves (4.1) and finishes the proof. ∎

Index

adapted, 44
almost surely, 22
arcsine law
  for the last sign-change, 139
  for the last zero, 138
  for the position of the maximum, 138
  for the time spent above zero, 140, 207
Azéma-Yor embedding theorem, 132

Bachelier, Louis, 41
ballot problem, 68
Bessel process, 164
Blumenthal's 0-1 law, 45
bounded variation, 35
Brown, Robert, 41
Brownian bridge, 38, 65
Brownian motion
  d-dimensional, 43
  fractional, 41
  monotonicity, 31
  standard, 14
  area, 51
  conformal invariance, 195
  definition, 21
  graph, 103
  image, 103
  Lévy's construction, 22
  linear, 21
  path, 103
  planar, 43
  range, 103
  scaling invariance, 25
  skew-product representation, 199
  standard, 21, 43
  time inversion, 26
  transient, 82, 220
  with drift, 37
  zero set, 52
Brownian scaling, 26

Cameron-Martin space, 213
Cantor set
  Hausdorff dimension, 117
  Minkowski dimension, 117
capacity
  α-capacity, 114, 226
  general definition, 226
  of an edge, 112
  Riesz, 114
Cauchy process, 56
Ciesielski-Taylor identity, 212
compactness argument, 342
conditional expectation, 53
conformal invariance, 18, 195
continuity
  Lévy's modulus, 29
  local Hölder, 31
convergence in distribution, 331
convolution, 50
cooling, 206
countable stability property, 99
countable subadditivity, 101
covering, 97
cut line, 144
cutset, 112

derivative
  upper and lower, 32
differentiability
  in a point, 33
  nowhere, 34
diffusion
  matrix, 14
dimension, 97
Dirichlet problem
  existence, 74
  harmonic measure, 88
  representation of solution, 15
  solution, 73
  uniqueness, 218
dissipation rate, 206

domain, 69
Donsker's invariance principle, 17, 134
Doob's L^p maximal inequality, 341
Doob's weak maximal inequality, 340
Doob, Joe, 68
downcrossing, 147
drift, 37
  vector, 14
Dubins' embedding theorem, 130
Dvoretzky-Erdős test, 77
Dynkin, Eugene, 68

Einstein, Albert, 41
embedding
  Azéma-Yor, 132
  Dubins, 130
  Skorokhod, 130
energy
  α-energy, 109, 226
energy method
  terminology, 109
envelope
  lower, 77
equilibrium measure, 221
Erdős-Kac theorem, 214
exceptional set
  of times, 33
exchangeable, 32
exit from an interval, 210
exit time, 61, 66

fast time, 276
Feynman-Kac formula, 207, 209
filtered probability space, 44
filtration, 44
first arcsine law, 138, 139
first moment method, 243
FKG inequality, 145
flow, 112
Ford and Fulkerson, 112
fractional Brownian motion, 41
Frostman's lemma, 111, 119
Frostman, Otto, 119
functional central limit theorem, 134

gauge function, 120, 171
Gaussian
  random vector, 23
  standard, 335
  process, 41
germ σ-algebra, 45
Green kernel, 83
Green's function
  definition, 83
  properties, 85

Hölder continuous, 101
harmonic function
  definition, 69
  discrete, 15
  history, 238
  maximum principle, 71
  mean value property, 69
  removable, 238
harmonic measure
  conformal invariance, 197
  definition, 87
  from infinity, 90
  on the sphere, 88
  support, 90
Harnack principle, 88
Harris inequality, 126
Hausdorff content, 99
Hausdorff dimension, 16
  and measure, 101
  countable stability, 117
  definition, 99, 100
  of graph, 110
  of range, 110
  of zero set, 108
Hausdorff measure
  definition, 100
  generalised, 171
Hausdorff, Felix, 119
Hawkes' theorem, 243
heat equation, 206
heating, 206
Helly-Bray theorem, 333
Hewitt-Savage 0-1 law, 32
hitting probability, 76, 228

increasing event, 143
independent, 43
independent increments, 13, 21
initial distribution, 14
integral test, 77
intersection equivalence
  definition, 247
intersections
  dimension, 245
  existence, 239, 241, 245
  nonexistence, 239, 241
intrinsic, 97
inversion at the circle, 89
irregular point, 217

Itô's formula
multi-dimensional, 192
nonsmooth case, 202
one-dimensional, 189
with additional dependence, 190

Kakutani's theorem, 226
Kakutani, S., 96
Kaufman's theorem, 116, 264
kernel, 224
Kochen-Stone lemma, 78, 94
Kolmogorov's 0-1 law, 45

Lévy's downward theorem, 339
Lévy's theorem
for local time, 204
for the maximum process, 54
Lévy, Paul, 41
Laplace's method, 200
law of
maximum, 50
law of large numbers, 27
law of the iterated logarithm
failure, 275
for Brownian motion, 121
for general random walk, 130
for simple random walk, 124
Hartman and Wintner, 130, 144
Khinchine, 121, 124, 144
Kolmogorov, 144
Lebesgue's thorn, 235
limsup fractal, 278
Liouville's theorem, 74
local maximum
dense, 46
strict, 46
local time
definition, 202
Lévy's theorem, 204
local time
at zero, 147
lower envelope, 77
Lyons' theorem, 250

macroscopic, 13
Makarov's theorem, 238, 310
Mandelbrot's conjecture, 238
marginal distributions, 22
Markov
strong property, 48
Markov process, 53
Markov property, 43
Markov transition kernel, 53
Markov, Andrej, 68
Martin kernel, 227
martingale, 58
binary splitting, 130
discrete, 337
reverse, 339
uniformly integrable, 338
martingale inequality, 340
mass distribution, 106
mass distribution principle, 106
max-flow min-cut theorem, 112, 342
McKean's theorem, 103, 115
microscopic, 13
Minkowski dimension
lower, 97
upper, 98
modification, 187
modulus normal distribution, 54
modulus of continuity, 27, 29, 38
monotonicity, 101
in an interval, 31
multiple points
existence, 256
nonexistence, 256

neighbourhood recurrent, 76
nonpolar set, 88
nowhere differentiable, 34

optional stopping theorem, 58
Orey and Taylor's theorem, 276
Ornstein-Uhlenbeck diffusion, 27, 65
outer boundary, 90
outer measure, 101

packing, 283
packing number, 283
Paley, Wiener, Zygmund's theorem, 34
Paley-Zygmund inequality, 77, 94
percolation, 250
percolation limit set
definition, 242
with generation dependent probabilities, 248
perfect set, 52
permuation
finite, 32
Poincaré cone condition, 217
point of increase
global, 127
local, 126
point recurrent, 76
Poisson kernel, 88
Poisson problem

definition, 219
uniqueness, 236
Poisson's formula, 88
polar
for Brownian motion, 88
for percolation, 243
points, 52
Portmanteau theorem, 332
potential
α-potential, 109
electrostatic, 87
gravitational, 109
Newtonian, 87
potential kernel, 83, 226
product formula, 286
progressively measurable, 183

quadratic variation, 36, 188

radial potential, 226
random closed set, 247
random field, 159
random walk, 13
Ray's theorem, 170
record time, 106
recurrent, 76
neighbourhood, 76
point, 76
reflection principle, 49, 74, 209
regular point, 217
regularisation, 283
removable, 238
resolvent operator, 211
reverse Hölder, 103, 119
reverse martingale, 339
Riemann mapping theorem, 195
right-continuity, 47

sample path properties, 21
scaling invariance, 17
second arcsine law, 140, 207
second moment method, 244
Skorokhod embedding theorem, 130
smallest upper envelope, 121
spectrum, 276
Spitzer's law, 201
stable subordinator, 56
stationary increments, 13
stochastic integral
as martingale, 187
construction, 184
continuous, 187
Fubini's theorem, 214
up to T, 187
stochastic integration, 18, 183
stochastic process, 21
stopping time
definition, 46
properties, 64
strict, 46
Strassen's law, 144
subharmonic function, 69
submartingale, 58

tail σ-algebra, 45
tail event, 45
Tanaka's formula, 202
Taylor, S. James, 119
transient, 76
trees
terminology, 112
Tricot, Claude, 309
typical times, 33

universal object, 13
universality, 13
upcrossing excursion, 260
upper envelope, 121

value of a covering, 99
value of a packing, 283
visible part of a set, 308

Wald's lemma, 59
first, 59
second, 60
Weierstrass function, 119
Wiener's theorem, 22
winding number, 201

zero set
Hausdorff dimension, 108
Hausdorff measure, 104
no isolated points, 52
uncountable, 53
zero-one law
Galton-Watson, 270
of Kolmogorov, 45
zero-one law
of Blumenthal, 45
of Hewitt-Savage, 32

Bibliography

[dA83] A. de Acosta. A new proof of the Hartman-Wintner law of the iterated logarithm. *Ann. Probab.* **11**, 270–276 (1983).

[Ad85] O. Adelman. Brownian motion never increases: a new proof of a theorem of Dvoretzky, Erdős and Kakutani. *Israel J. of Math.* **50**, 189–192 (1985).

[ABP98] O. Adelman, K. Burdzy and R. Pemantle. Sets avoided by Brownian motion. *Ann. Probab.* **26**, 429–464 (1998).

[AD85] O. Adelman and A. Dvoretzky. Plane Brownian motion has strictly n-multiple points. *Israel J. Math.* **52**, 361–364 (1985).

[Ad90] R. Adler. *An introduction to continuity, extrema and related topics for general Gaussian processes.* Institute Math. Statistics, Hayward CA (1990).

[An87] D. André. Solution directe du problème résolu par M. Bertrand. *C. R. Acad. Sci. Paris* **105**, 436–437 (1887).

[AN04] K. B. Athreya and P.E. Ney. *Branching processes.* Dover Publications, Mineola NY (2004).

[AY79] J. Azéma and M. Yor. Une solution simple au problème de Skorokhod. In: *Seminar on probability XIII*, 90–115, Springer Lecture Notes in Mathematics 721 (1979).

[Ba00] L. Bachelier. Théorie de la speculation. *Ann. Sci. Ecole Norm. Sup.* **17** 21–86 (1900).

[Ba01] L. Bachelier. Théorie mathématique du jeu. *Ann. Sci. Ecole Norm. Sup.* **18** 143–210 (1901).

[BP84] M. T. Barlow and E. Perkins. Levels at which every Brownian excursion is exceptional. *Lecture Notes in Math.* **1059**, Springer, Berlin-New York, 1–28 (1984).

[Ba95] R. Bass. *Probabilistic Techniques in Analysis.* Springer-Verlag, New York (1995).

[BB97] R. Bass and K. Burdzy. Cutting Brownian paths. *Memoir Amer. Math. Soc.* **137** (1997).

[Ba62] G. Baxter. Combinatorial methods in fluctuation theory. *Z. Wahrsch. verw. Gebiete.* **1** 263–270 (1962).

[BPP95] I. Benjamini, R. Pemantle and Y. Peres. Martin capacity for Markov chains. *Ann. Prob.* **23**, 1332–1346 (1995).

[BP94] I. Benjamini and Y. Peres. Tree-indexed random walks on groups and first passage percolation. *Probab. Th. Rel. Fields.* **98**, 91–112 (1994).

[Be83] S. M. Berman. Nonincrease almost everywhere of certain measurable functions with applications to stochastic processes. *Proc. Amer. Math. Soc.* **88**, 141–144 (1983).

[Be91] J. Bertoin. Increase of a Lévy process with no positive jumps. *Stochastics* **37**, 247–251 (1991).

[Bi67] P. Bickel. Some contributions to the theory of order statistics. *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability, Vol I: Statistics*, 575–591 (1967).

[Bi68] P. Billingsley. *Convergence of Probability Measures.* Wiley, New York (1968).

[Bi82] P. Billingsley. Van der Waerden's continuous nowhere differentiable function. *Amer Math. Monthly* **89** 691 (1982).

[Bi86] N. H. Bingham. Variants on the law of the iterated logarithm. *Bull. London Math. Soc.* **18**, 433–467 (1986).

[BP96] C. J. Bishop and Y. Peres. Packing dimension and Cartesian products. *Trans. Amer. Math. Soc.* **348**, 4433–4445 (1996).

[BJPP97] C. J. Bishop, P. W. Jones, R. Pemantle and Y. Peres. The outer boundary of Brownian motion has dimension greater than one. *J. Funct. Analysis* **143**, 309-336 (1997).

[Bl57] R.M. Blumenthal. An extended Markov property. *Trans. Amer. Math. Soc.* **82**, 52–72 (1957).

[BG68] R.M. Blumenthal and R. Getoor. Markov processes and potential theory. Academic Press (1968).

[Bo89] A.N. Borodin. Brownian local time. *Russian Math. Surveys* **44**, 1–51 (1989).

[BS02] A.N. Borodin and P. Salminen. Handbook of Brownian motion. Facts and formulae. 2$^{\text{nd}}$ Edition. Birkhäuser Verlag, Basel (2002).

[Br76] G.R. Brosamler. A probabilistic version of the Neumann problem. *Math. Scand.* **38**, 137–147 (1976).

[Br28] R. Brown. A brief descriptions of microscopical observations made in the months of June, July and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *Ann. Phys.* **14**, 294–313 (1828).

[Bu89] K. Burdzy. Cut points on Brownian paths. *Ann. Probab.* **17**, 1012–1036 (1989).

[Bu90] K. Burdzy. On nonincrease of Brownian motion. *Ann. Probab.* **18**, 978–980 (1990).

[Bu95] K. Burdzy. Labyrinth dimension of Brownian trace. In: *Probability and Mathematical Statistics* Volume dedicated to the memory of Jerzy Neyman **15**, 165–193 (1995).

[BL90] K. Burdzy and G. F. Lawler. Non-intersection exponents for Brownian paths. Part II: Estimates and applications to a random fractal. *Ann. Probab.* **18**, 981–1009 (1990).

[BSM89] K. Burdzy and J. San Martín. Curvature of the convex hull of planar Brownian motion near its minimum point. *Stoch. Proc. Appl.* **33**, 89–103 (1989).

[Ca14] C. Carathéodory. Über das lineare Maß von Punktmengen, eine Verallgemeinerung des Längenbegriffs. *Nachrichten Ges. Wiss. Göttingen*, 406–426 (1914).

[Ca67] L. Carleson. *Selected Problems on Exceptional Sets.* Van Nostrand, Princeton, NJ (1967).

[Ch72] J.P.R. Christensen. On sets of Haar measure zero in abelian Polish groups. *Israel J. Math.* **13**, 255–260 (1972).

[Ch73] K.L. Chung. Probabilistic approach in potential theory to the equilibrium problem. *Ann. Inst. Fourier, Grenoble* **23,3**, 313–322 (1973).

[Ch82] K.L. Chung. *Lectures from Markov processes to Brownian motion.* Springer Verlag (1982).

[Ch02] K.L. Chung. *Green, Brown and Probability & Brownian motion on the line.* World Scientific (2002).

[CFL28] R. Courant, K. Friedrichs and H. Lewy. Über die partiellen Differentialgleichungen der mathematischen Physik. *Math. Annalen* **100**, 32–74 (1928).

[CHM89] M. Cranston, P. Hsu and P. March. Smoothness of the convex hull of planar Brownian motion. *Ann. Probab.* **17**, 144–150 (1989).

[CR81] M. Csörgö and P. Revesz. *Strong Approximations in Probability and Statistics.* Academic Press, New York (1981).

[CT62] Z. Ciesielski and S. J. Taylor. First passage times and sojourn times for Brownian motion in space and the exact Hausdorff measure of the sample path. *Trans. Amer. Math. Soc.* **103** (1962).

[CW90] K.L. Chung and R. J. Williams. *Introduction to stochastic integration.* Second edition. Birkhäuser, Boston, MA (1990).

[Da65] K.E. Dambis. On the decomposition of continuous martingales. *Theor. Prob. Appl.* **10**, 401-410 (1965).

[Da75] B. Davis. Picard's theorem and Brownian motion. *Trans. Amer. Math. Soc.* **213**, 353-362 (1975).

[Da83] B. Davis. On Brownian slow points. *Z. Wahrscheinlichkeitstheorie.* **64**, 359-367 (1983).

[DE06] M. Davis and A. Etheridge. Louis Bachelier's theory of speculation: The origins of modern finance. Princeton University Press, Princeton, NJ (2006).

[DM98] P. Deheuvels and D.M. Mason. Random fractal functional laws of the iterated logarithm. *Studia Sci. Math.Hungar.* **34**, 89-106 (1998).

[DM04] P. Del Moral. Feynman-Kac formulae. Genealogical and interacting particle systems with applications. Springer Verlag (2004).

[DPRZ00] A. Dembo, Y. Peres, J. Rosen and O. Zeitouni. Thick points for spatial Brownian motion: multifractal analysis of occupation measure. *Ann. Probab.* **28(1)**, 1–35 (2000).

[DPRZ01] A. Dembo, Y. Peres, J. Rosen and O. Zeitouni. Thin points for Brownian motion. *Ann. Inst. H. Poincaré Probab. Statist.*, **36**(6), 749–774 (2001).

[Do96] R. A. Doney. Increase of Lévy processes. *Ann. Probab.* **24**, 961–970 (1996).

[Do51] M. D. Donsker. An invariance principle for certain probability limit theorems. *Mem. Amer. Math. Soc.* **6** (1951–52).

[Do49] J. L. Doob. Heuristic approach to the Kolmogorov-Smirnov theorems. *Ann. Math. Stat.* **20**, 393–403 (1949).

[Do53] J.L. Doob. *Stochastic Processes.* Wiley (1953).

[Du68] L. E. Dubins. On a theorem of Skorokhod. *Ann. Math. Statist.* **39**, 2094–2097 (1968).

[DS65] L. E. Dubins and G. Schwarz. On continuous martingales. *Proc. Nat. Acad. Sci. USA* **53**, 913–916 (1965).

[Du02] R. M. Dudley. *Real Analysis and Probability.* Cambridge University Press (2002).

[DL02] T. Duquesne and J.-F. Le Gall. Random trees, Lvy processes and spatial branching processes. *Astérisque* **281**, 1–147 (2002).

[Du95] R. Durrett. *Probability: Theory and examples.* Duxbury Press, Belmont (1995).

[Du96] R. Durrett. *Stochastic calculus: a practical introduction.* CRC Press (1996).

[Dv63] A. Dvoretzky. On the oscillation of the Brownian motion process. *Israel Journal Math.* **1**, 212-214 (1963).

[DE51] A. Dvoretzky and P. Erdős. Some problems on random walk in space. In: *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950.* 353–367 (1951).

[DEK50] A. Dvoretzky, P. Erdős and S. Kakutani Double points of paths of Brownian motion in n-space. *Acta Sci. Math. Szeged* **12** 75–81 (1950).

[DEK54] A. Dvoretzky, P. Erdős and S. Kakutani Multiple points of paths of Brownian motion in the plane. *Bull. Res. Council Israel* **3** 364–371 (1954).

[DEK58] A. Dvoretzky, P. Erdős and S. Kakutani Points of multiplicity $\mathfrak{c}$ of planar Brownian motion. *Bull. Res. Council Israel* **7** 157–180 (1958).

[DEK61] A. Dvoretzky, P. Erdős and S. Kakutani. Nonincrease everywhere of the Brownian motion process. *Proc. 4th Berkeley Symp. Math. Stat. Probab.* **2**, 103-116 (1961).

[DEKT57] A. Dvoretzky, P. Erdős, S. Kakutani and S. J. Taylor Triple points of Brownian paths in 3-space. *Proc. Cambridge Philos. Soc.* **53** 856–862 (1957).

[Dy57] E. B. Dynkin. Inhomogeneous strong Markov processes. *Dokl. Akad. Nauk. SSSR.* **113**, 261–263 (1957).

[Ei05] A. Einstein. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Ann. Physik* **17** 549–560 (1905).

[Ei94] N. Eisenbaum. Dynkin's isomorphism theorem and the Ray-Knight theorems. *Probab. Theory Related Fields* **99** 321–335 (1994).

[EK46] P. Erdős and M. Kac. On certain limit theorems in the theory of probability. *Bull. Amer. Math. Soc.* **52** 292–302 (1946).

[EK47] P. Erdős and M. Kac. On the number of positive sums of independent random variables. *Bull. Amer. Math. Soc.* **53** 1011–1020 (1947).

[Ev85] S. Evans. On the Hausdorff dimension of Brownian cone points. *Math. Proc. Cambridge Philoph. Soc.* **98** 343–353 (1985).

[Ev87a] S. Evans. Potential theory for a family of several Markov processes. *Ann. Inst. H. Poincaré Probab. Statist.* **23** 499-530 (1987).

[Ev87b] S. Evans. Multiple points in the sample paths of a Lévy process. *Probab. Theory Related Fields* **76** 359–367 (1987).

[Fa97a] K. Falconer. *Fractal geometry: mathematical foundations and applications.* Wiley, Chichester (1997).

[Fa97b] K. Falconer. *Techniques in fractal geometry.* John Wiley & Sons, Chichester (1997).

[FH96] K.J. Falconer and J. D. Howroyd. *Projection theorems for box and packing dimension. Math. Proc. Camb. Phil. Soc.* **119**, 287–295 (1996).

[Fe68] W. Feller. *An introduction to probability theory and its applications.* Volume I. 3^{rd} Edition. John Wiley & Sons, New York (1968).

[Fe66] W. Feller. *An introduction to probability theory and its applications.* Volume II. John Wiley & Sons, New York (1966).

[FS89] P.J. Fitzsimmons and T. Salisbury. Capacity and energy for multiparameter Markov processes. *Ann. Inst. Henri Poincaré, Probab.* **25** 325–350 (1989).

[FF56] L. R. Ford Jr. and D. R. Fulkerson. Maximal flow through a network. *Canad. J. Math.* **8**, 399–404 (1956).

[Fr67] B. Fristedt An extension of a theorem of S. J. Taylor concerning the multiple points of the symmetric stable process. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **9** 62–64 (1967).

[FG97] B. Fristedt and L. Gray. *A modern approach to probability theory.* Birkhäuser, Boston (1997).

[FKG71] C. M. Fortuin, P. N. Kasteleyn and J. Ginibre. Correlational inequalities for partially ordered sets. *Comm. Math. Phys.* **22**, 89–103 (1971).

[Fr35] O. Frostman. Potential d'équilibre et capacité des ensembles avec quelques applications a la théorie des fonctions. *Meddel. Lunds Univ. Math. Sem.* **3**, 1–118, (1935).

[Ga40] C. F. Gauss. Allgemeine Lehrsätze in Beziehung auf die im verkehrten Verhältnisse des Quadrats der Entfernung wirkenden Anziehungs- und Abstoßungskräfte. *Gauss Werke* **5** 197–242 (1840).

[GH80] D. Geman and J. Horowitz. Occupation densities. *Ann. Probab.* **8**, 1–67 (1980).

[GHR84] D. Geman, J. Horowitz and J. Rosen. A local time analysis of intersections of Brownian motion in the plane. *Ann. Probab.* **12**, 86–107 (1984).

[Gr28] G. Green. An essay on the application of mathematical analysis to the theories of electricity and magnetism. (1828).

[GP83] C. Greenwood and E. Perkins. A conditioned limit theorem for random walk and Brownian local time on square root boundaries. *Ann. Probab.* **11**, 227-261 (1983).

[Ha19] F. Hausdorff. Dimension und üßeres Maß. *Math. Ann.* **79**, 157–179 (1919).

[Ha60] T. E. Harris. A lower bound for the critical probability in a certain percolation process. *Proc. Camb. Phil. Soc.* **56**, 13–20 (1960).

[HW41] P. Hartman and A. Wintner. On the law of the iterated logarithm. *J. Math.* **63**, 169–176 (1941).

[Ha81] J. Hawkes. Trees generated by a simple branching process. *J. London Math. Soc.*, **24**, 373–384 (1981).

[Ho95] J.D. Howroyd. On dimension and on the existence of sets of finite positive Hausdorff measure. *Proc. Lond. Math. Soc.*, III. **70**, 581-604 (1995).

[HT97] X. Hu and S. J. Taylor. The multifractal structure of stable occupation measure. *Stoch. Proc. Appl.* **66**, 283-299 (1997).

[HSY92] B. R. Hunt, T. Sauer and J. A. Yorke. Prevalence: a translation-invariant "almost every" on infinite-dimensional spaces. *Bull. Amer. Math. Soc.* **27** (2), 217–238 (1992).

[Hu56] G. A. Hunt. Some theorems concerning Brownian motion. *Trans. Amer. Math. Soc.* **81** , 294–319 (1956).

[It44] K. Itô. Stochastic integral. *Proc. Imp. Acad. Tokyo* **20**, 519–524 (1944).

[It51] K. Itô. On a formula concerning stochastic differentials. *Nagoya Math. Journal* **3**, 55–55 (1951).

[IM74] K. Itô and H.P. McKean. *Diffusion processes and their sample paths.* Springer, Berlin, 1974.

[Ka51] M. Kac. On some connections between probability theory and differential and integral equations. *Proc. 2^{nd} Berkeley Symp. Math. Statist. Probab.*, 189–215 (1951).

[Ka76] J.-P. Kahane. Sur les zéros et les instants de ralentissement du mouvement brownien. *C. R. Acad. Sci. Paris* **282**, 431–433 (1976).

[Ka85] J.-P. Kahane. *Some random series of functions.* Cambridge University Press (1985).

[Ka86] J.-P. Kahane. Sur la dimension des intersections. In: Aspects of Mathematics and its Applications (Ed. J.A. Barroso), 419–430, North Holland (1986).

[Ka44a] S. Kakutani. On Brownian motions in n-space. *Proc. Imp. Acad. Tokyo* **20**, 648–652 (1944).

[Ka44b] S. Kakutani. Two dimensional Brownian motion and harmonic functions. *Proc. Imp. Acad. Tokyo* **20**, 706–714 (1944).

[Ka45] S. Kakutani. Markoff process and the Dirichlet problem. *Proc. Japan Acad.* **21**, 227–233 (1945).

[Ka02] O. Kallenberg. *Foundations of modern probability.* Springer Verlag (2002).

[KS88] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus.* Springer Verlag (1988).

[Ka69] R. Kaufman. Une propriété métrique du mouvement brownien. *C. R. Acad. Sci. Paris* **268**, 727-728 (1969).

[Ka75] R. Kaufman. Large increments of Brownian motion. *Nagoya Math. Journal* **56**, 139-145 (1975).

[Kh23] A.Y. Khinchin. Über dyadische Brüche. *Math. Zeitschrift* **18**, 109-116 (1923).

[Kh24] A.Y. Khinchin. Über einen Satz der Wahrscheinlichkeitsrechnung. *Fund. Math.* **6**, 9-20 (1924).

[Kh33] A.Y. Khinchin. *Asymptotische Gesetze der Wahrscheinlichkeitsrechnung.* Springer Verlag, Berlin (1933).

[Kh02] D. Khoshnevisan. *Multiparameter processes.* Springer Verlag (2002).

[Kh99] D. Khoshnevisan. Brownian sheet images and Bessel-Riesz capacity. *Trans. Amer. Math. Soc.* **351**, 2607-2622 (1999)

[KPX00] D. Khoshnevisan, Y. Peres and Y. Xiao. Limsup random fractals. *El. J. Probab.* **5**, Paper 4, pp.1–24 (2000).

[KS00] D. Khoshnevisan and Z. Shi. Fast sets and points for fractional Brownian motion. Séminaire de Probabilités, XXXIV, 393–416, Lecture Notes in Math., 1729, Springer, Berlin, (2000).

[KX05] D. Khoshnevisan and Y. Xiao. Lévy processes: Capacity and Hausdorff dimension. *Ann. Probab.* **33**, 841–878 (2005).

[KM05] A. Klenke and P. Mörters. The multifractal spectrum of Brownian intersection local times. *Ann. Probab.* **33**, 1255-1301 (2005).

[Kn63] F. B. Knight. Random walks and a sojurn density process of Brownian motion. *Trans. Amer. Math. Soc.* **109**, 56–86 (1963).

[Kn81] F. B. Knight. *Essentials of Brownian Motion and Diffusion.* Amer. Math. Soc., Providence, Rhode Island (1981).

[Ko29] A.N. Kolmogorov. Über das Gesetz des iterierten Logarithmus. *Math. Ann.*, **101**, 126-135, (1929).

[KMT75] J. Komlós, P. Major and G. Tusnády. An approximation of partial sums of independent random variables and the sample distribution function. *Z. Wahrsch. verw. Gebiete*, **32**, 111-131, (1975).

[KM02] W. König and P. Mörters. Brownian intersection local times: upper tails and thick points. *Ann. Probab.*, **30(4)**, 1605-1656, (2002).

[La98] M. Laczkovich. *Conjecture and proof.* TypoTeX, Budapest (1998).

[La09] E. Landau. *Verteilung der Primzahlen.* (1909).

[La96a] G.F. Lawler. Hausdorff dimension of cut points for Brownian motion. *El. Journal Probab.* **1** (1996).

[La96b] G.F. Lawler. The dimension of the frontier of planar Brownian motion. *El. Comm. in Probab.* **1**, 29–47 (1996).

[La99] G.F. Lawler. Geometric and fractal properties of Brownian motion and random walk paths in two and three dimensions. In: *Random walks (Budapest 1998)* 219–258, Bolyaj Society Math. Stud., Volume 9, Budapest, 1999.

[LSW01] G. F. Lawler, O. Schramm and W. Werner. The dimension of the Brownian frontier is 4/3. *Math. Res. Lett.* **8**, 401-411 (2001).

[Le24] H. Lebesque. Conditions de régularité, conditions d' irrégularité, conditions de impossibilité dans le problème de Drirchlet. *Comp. Rendu Acad. Sci.* **178**, 349–354 (1924).

[LG85] J.-F. Le Gall. Sur la measure de Hausdorff de la courbe brownienne. In: *Séminaire de probabilités (Strasbourg)* **19** 297–313 (1985).

[LG86] J.-F. Le Gall. The exact Hausdorff measure of Brownian multiple points. In: *Seminar on Stochastic Processes* (1986), Birkhäuser, Basel, pp 107–137.

[LG87] J.-F. Le Gall. Le comportement du mouvement brownian entre les deux instants où il passe par un point double. *J. Funct. Anal.* **71**, 246–262 (1987).

[LG91] J.-F. Le Gall. Some properties of planar Brownian motion. Springer Lecture Notes in Mathematics.

[LG99] J.-F. Le Gall. *Spatial branching processes, random snakes and partial differential equations.* Lecture Notes in Mathematics ETH Zürich. Birkhäuser, Basel (1999).

[LL98] J.-F. Le Gall and Y. Le Jan. *Branching processes in Lvy processes: the exploration process. Ann. Probab.* **26** 213–252 (1998).

[LRS89] J.-F. Le Gall, J.S. Rosen, and N.-R. Shieh. Multiple points of Lévy processes. *Ann. Probab.* **17** 503–515 (1989).

[Le66] E. Lehmann. Some concepts of dependence. *Ann. Math. Statist.* **37**, 1137–1153 (1966).

[Le37] P. Lévy. *Théorie de l'addition des variables aléatoires.* Gauthier-Villars, Paris (1937).

[Le39] P. Lévy. Sur certain processus stochastiques homogénes. *Comp. Math.* **7** 283–339 (1939).

[Le40] P. Lévy. Le mouvement brownien plan. *Amer. J. Math.* **62** 487–550 (1940).

[Le48] P. Lévy. *Processus stochastiques et mouvement Brownien.* Suivi d'une note de M. Loève. Gauthier-Villars, Paris (1948).

[Le51] P. Lévy. La mesure de Hausdorff de la courbe du mouvement brownien à n dimensions. *C. R. Acad. Sci. Paris* **233** 600–602 (1951).

[Le54] P. Lévy. La mesure de Hausdorff de la courbe du mouvement brownien. *Giorn. Ist. Ital. Attuari* **16** 1–37 (1954).

[Li95] M.A. Lifshits. *Gaussian random functions.* Kluwer, Dordrecht (1995).

[Ly90] R. Lyons. Random walks and percolation on trees. *Ann. Probab.* **18**, 931-958 (1990).

[Ly92] R. Lyons. Random walks, capacity and percolation on trees. *Ann. Probab.* **20** 2043–2088 (1992).

[Ly96] R. Lyons. Probabilistic aspects of infinite trees and some applications. In: Trees, pp. 81–94, Birkhäuser, Basel (1996).

[LP05] R. Lyons with Y. Peres. *Probability on trees and networks.* To appear. Preliminary version currently available at *http://php.indiana.edu/ rdlyons/prbtree/prbtree.html*

[MR06] M.B. Marcus and J. Rosen. *Markov processes, Gaussian processes, and local times.* Cambridge University Press, Cambridge (2006).

[Ma85] N.G. Makarov. On the distortion of boundary sets under conformal mappings. *Proc. London Math. Soc.(3)* **51** 369–384 (1985).

[Ma06] A.A. Markov. Extension of the law of large numbers to dependent events. [Russian] *Bull. Soc. Phys. Math. Kazan* **15** 135–156 (1906).

[Ma95] P. Mattila. *Geometry of sets and measures in Euclidean spaces. Fractals and rectifiability.* Cambridge University Press (1999).

[MM97] P. Mattila and R.D. Mauldin. Measure and dimension functions: measurability and densities. *Math. Proc. Camb. Philos. Soc.* **121**, 81–100 (1997).

[McK55] H.P. McKean Jr. Hausdorff-Besicovitch dimension of Brownian motion paths. *Duke Math. J.* **22**, 229–234 (1955).

[Me83] I. Meilijson. On the Azéma-Yor stopping time. In: *Seminar on probability XVII*, 225–226, Springer LNM 986 (1983).

[Me66] P. A. Meyer. *Probability and Potentials.*. Blaisdell Publishing, Waltham (1966).

[MG84] D. P. Minassian and J. W. Gaisser. A simple Weierstrass function. *Amer. Math. Monthly* **91**, 254-256 (1984).

[Mö02] P. Mörters. A pathwise version of Spitzer's law. In: *Limit theorems in Probability and Statistics II*, János Bolyai Mathematical Society 427–436 (2002).

[Ne75] J. Neveu. *Discrete-parameter martingales.* Revised Edidtion. Translated from the French by T. P. Speed. North Holland, Amsterdam (1975).

[NP89] J. Neveu and J.W. Pitman. *The branching process in a Brownian excursion.* Séminaire de Probabilités, XXIII, 248–257, Lecture Notes in Math., 1372, Springer, Berlin (1989).

[Ob04] I. Obłój The Skorokhod embedding problem and its offspring. *Probab. Surv.* **1**, 321–390 (2004).

[ON04] T.C. O'Neil. The Hausdorff dimension of the visible sets of connected compact sets *Trans. Amer. Math. Soc.* to appear, (2004).

[OT74] S. Orey and S. J. Taylor. How often on a Brownian path does the law of the iterated logarithm fail? *Proc. London Math. Soc.(3)* **28**, 174–192 (1974).

[PWZ33] R. E. A. C. Paley, N. Wiener and A. Zygmund. Notes on random functions. *Math. Zeitschrift* **37**, 647-668 (1933).

[Pe95] R. Pemantle. Tree-indexed processes. *Stat. Sci.* **10**, 200–213 (1995).

[Pe97] R. Pemantle. The probability that Brownian motion almost contains a line. *Ann. Inst. H. Poincaré, Probab. Stat.* **33**, 147–165 (1997).

[PP07] R. Pemantle and Y. Peres. What is the probability of intersecting the set of Brownian double points? *Ann. Probab.* **35**, 2044–2062 (2007).

[Pe96a] Y. Peres. Remarks on intersection-equivalence and capacity equivalence. *Ann. Inst. H. Poincaré (Physique Théorique)* **64**, 339–347 (1996).

[Pe96b] Y. Peres. Intersection-equivalence of Brownian paths and certain branching processes. *Commun. Math. Phys.* **177**, 417-434 (1996).

[Pe96c] Y. Peres. Points of increase for random walks. *Israel J. of Math.* **95**, 341–347 (1996).

[Pe99] Y. Peres. *Probability on trees: An introductory climb.* In: Bertoin, J. (ed.) et al., Lectures on probability theory and statistics. Ecole d'été de Probabilités de Saint-Flour XXVII-1997. Berlin: Springer. Lect. Notes Math. 1717, 193-280 (1999).

[Pe81] E. Perkins. The exact Hausdorff measure of the level sets of Brownian motion. *Z. Wahrscheinlichkeitstheorie.* **58**, 373-388 (1981).

[Pe83] E. Perkins. On the Hausdorff dimension of Brownian slow points. *Z. Wahrscheinlichkeitstheorie.* **64**, 369-399 (1983).

[PT87] E. Perkins and S.J. Taylor. Uniform measure results for the image of subsets under Brownian motion. *Probab. Theory rel. Fields.* **76**, 257-289 (1987).

[PT88] E. Perkins and S.J. Taylor. Measuring close approaches ona Brownian path. *Ann. Probab.* **16**, 1458-1480 (1988).

[PW23] H. B. Phillips and N. Wiener. Nets and Dirichlet problem. *J. Math. Phys.* **2**, 105-124 (1923).

[Pi75] J.W. Pitman. One-dimensional Brownian motion and the three-dimensional Bessel process. *Advances in Appl. Probability.* **7** 511–526 (1975).

[PY86] J. Pitman and M. Yor. Asymptotic laws of planar Brownian motion. *Ann. Probab.* **14** 733–779 (1986).

[Po90] H. Poincaré. Sur les équations aux derivées partielles de la physique mathématique. *Amer. J. Math.* **12** 211–779 (1986).

[Po21] G. Pólya. Über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die Irrfahrt im Strassennetz. *Math. Ann.* **84** 149–160 (1921).

[PS78] S.C. Port and C.J. Stone Brownian motion and classical potential theory. Academic Press, New York (1978).

[Pr56] Y.V. Prohorov. Convergence of random processes and limit theorems in probability. *Th. Probab. Appl.* **1** 157–214 (1956).

[Pr90] W.E. Pruitt. The rate of escape of random walk. *Ann. Probab.* **18** 1417–1461 (1990).

[PU89] F. Przytycki and M. Urbański. On the Hausdorff dimension of some fractal sets. *Studia Math.* **93**, 155–186 (1989).

[Ra63a] D. Ray. Sojourn times and the exact Hausdorff measure of the sample path for planar Brownian motion. *Trans. Amer. Math. Soc.* **106**, 436–444 (1963).

[Ra63b] D. Ray. Sojourn times of diffusion processes. *Illinois J. Math.* **7**, 615–630 (1963).

[RY94] D. Revuz and M. Yor. *Continuous martingales and Brownian motion.* Springer Verlag (1994).

[RP49] H. Robbins and E.J.G. Pitman. Application of the method of mixtures to quadratic forms. *Ann. Math. Statist.* **20**, 552–560 (1949).

[Ro99] C.A. Rogers. *Hausdorff measures.* Cambridge University Press (1999).

[RT61] C.A. Rogers and S.J. Taylor. Functions continuous and singular with respect to a Hausdorff measure. *Mathematika* **8**, 1–31 (1961).

[RW00] L.C.G. Rogers and D. Williams. *Diffusions, Markov processes and martingales.* Cambridge University Press (2000).

[Ro69] D. H. Root. The existence of certain stopping times on Brownian motion. *Ann. Math. Statist.* **40**, 715–718 (1969).

[Ro83] J. Rosen. A local time approach to the self-intersections of Brownian paths in space. *Comm. Math. Phys.* **88**, 327–338 (1983).

[Ru86] W. Rudin. Real and complex analysis. 3$^{\text{rd}}$ edition. McGraw-Hill.

[Se95] L. Serlet. Some dimension results for super-Brownian motion. *Probab. Theory Related Fields* **101**, 371–391 (1995).

[Sh67] L.A. Shepp. A first passage problem for the Wiener process. *Ann. Math. Statist.* **38**, 1912–1914.

[Sh98] Z. Shi. Windings of Brownian motion and random walks in the plane. *Ann. Probab.* **26** 112–131 (1998).

[ST99] N.-R. Shieh and S.J. Taylor. Logarithmic multifractal spectrum of stable occupation measure.

[Sk65] A. Skorokhod. *Studies in the theory of random processes.* Addison-Wesley, Reading, MA (1965).

[Sp58] F. Spitzer. Some theorems concerning 2-dimensional Brownian motion. *Trans. Amer. Math. Soc.* **87** 187–197 (1958).

[Sp64] F. Spitzer. Principles of random walk. Van Nostrand, Princeton (1964).

[St75] F. Stern. Conditional expectation of the duration in the classical ruin problem. *Mathematics magazine.* **48** 200-203 (1975).

[St64] V. STRASSEN. An invariance principle for the law of the iterated logarithm. *Z. Wahrsch. Verw. Gebiete* **3**, 211–226 (1964).

[Ta63] H. TANAKA Note on continuous additive functionals of the one-dimensional Brownian path. *Z. Wahrsch. Verw. Gebiete* **1**, 251–157(1963).

[Ta53] S. J. TAYLOR. The Hausdorff α-dimensional measure of Brownian paths in n-space. *Proc. Cambridge Philos. Soc.* **49**, 31–39 (1953).

[Ta55] S. J. TAYLOR. The α-dimensional measure of the graph and set of zeros of a Brownian path. *Proc. Cambridge Philos. Soc.* **51**, 265–274 (1955).

[Ta64] S. J. TAYLOR. The exact Hausdorff measure of the sample path for planar Brownian motion. *Proc. Cambridge Philos. Soc.* **60**, 253–258 (1964).

[Ta66] S. J. TAYLOR. Multiple points for the sample paths of the symmetric stable process. *Zeitschr. Wahrsch. verw. Gebiete* **5**, 247–264 (1966).

[Ta72] S. J. TAYLOR. Exact asymptotic estimates of Brownian path variation. *Duke Math. J.* **39**, 219–241 (1972).

[Ta86] S. J. TAYLOR. The measure theory of random fractals. *Proc. Cambridge Philos. Soc.* **100** 383–406 (1986).

[TT85] S. J. TAYLOR and C. TRICOT. Packing measure and its evaluation for a Brownian path. *Trans. Amer. Math. Soc.* **288** 679–699 (1985).

[TW66] S. J. TAYLOR and J. G. WENDEL. The exact Hausdorff measure of the zero set of a stable process. *Zeitschr. Wahrsch. verw. Gebiete* **6**, 170–180 (1966).

[To88] N. TONGRING. Which sets contain multiple points of Brownian motion? *Math. Proc. Camb. Phil. Soc.* **103**, 181–187 (1988).

[Tr82] C. TRICOT. Two definitions of fractional dimension. *Math. Proc. Cambridge Philos. Soc.* **91**, 57–74 (1982).

[Tr58] H. TROTTER. A property of Brownian motion paths. *Ill. Journal Math.* **2**, 425–433 (1958).

[We04] W. WERNER. Random planar curves and Schramm-Loewner evolutions. In: *Lectures on probability theory and statistics*, Lecture Notes in Math., 1840, 107–195, Springer (2004).

[Wi23] N. WIENER. Differential space. *Journal Math. Phys.* **2**, 131–174 (1923).

[Wi91] D. WILLIAMS. *Probability with martingales.* Cambridge University Press (1991).

[Xi04] Y. XIAO. Random fractals and Markov processes. In: *Fractal Geometry and Applications: A Jubilee of Benoit Mandelbrot.* Eds: M.L. Lapidus, M. van Frankenhuijsen. pp 261–338. American Mathematical Society (2004).

[Yo92] M. YOR. Some aspects of Brownian motion. Part I. Birkhäuser Verlag (1992).

[Za11] S. ZAREMBA. Sur le principe de Dirichlet. *Acta Math.* **34** 293–316 (1911).