Convex hulls of random walks

JAMES FERGAL WILFRED MCREDMOND

Convex hulls of random walks

James McRedmond

A Thesis presented for the degree of
Doctor of Philosophy

The Statistics and Probability Group
Department of Mathematical Sciences
Durham University
United Kingdom

September 2019

Dedicated to

Anne Curley

And

Louis McRedmond

Convex hulls of random walks

James McRedmond

Submitted for the degree of Doctor of Philosophy

September 2019

Abstract: We study the convex hulls of random walks establishing both law of large numbers and weak convergence statements for the perimeter length, diameter and shape of the hull. It should come as no surprise that the case where the random walk has drift, and the zero-drift case behave differently. We make use of several different methods to gain a better insight into each case.

Classical results such as Cauchy's surface area formula, the law of large numbers and the central limit theorem give some preliminary law of large number results.

Considering the convergence of the random walk and then using the continuous mapping theorem leads to intuitive results in the case with drift where, under the appropriate scaling, non-zero, deterministic limits exist. In the zero-drift case the random limiting process, Brownian motion, provides insight into the behaviour of such a walk. We add to the literature in this area by establishing tighter bounds on the expected diameter of planar Brownian motion. The Brownian motion process is also useful for proving that the convex hull of the zero-drift random walk has no limiting shape.

In the case with drift, a martingale difference method was used by Wade and Xu to prove a central limit theorem for the perimeter length. We use this framework to establish similar results for the diameter of the convex hull. Time-space processes give degenerate results here, so we use some geometric properties to further what is known about the variance of the functionals in this case and to prove a weak convergence statement for the diameter. During the study of the geometrical properties, we show that, only finitely often is there a single face in the convex minorant (or concave majorant) of such a walk.

Declaration

The work in this thesis is based on research carried out in the The Statistics and Probability Group, Department of Mathematical Sciences, Durham University, United Kingdom. No part of this thesis has been submitted elsewhere for any other degree or qualification and it is all my own work unless referenced to the contrary in the text.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to Andrew and Ostap for their ideas, patience and supervision over so many years. I couldn't have developed in this subject area without their vast knowledge and brilliant guidance throughout.

I am indebted to my wonderful family for their support and love throughout my education (and whole life!), and grateful that the probability theory conversations haven't tired them out yet!

Of course, being a PhD student is no fun without the other students. So to all in the Trinity Offices and the board game group, thank you for such great times inside and out of the department. I am extremely grateful to the probability students for their discussions and input not only in relation to my work but also for broadening my knowledge of many other models of random processes. Also thanks to Chang for the interesting collaboration which is included in this work.

Thanks also to the Lawson 5 for making my undergraduate degree what it was, leading to the choice to continue into doctoral studies. And to Mariana and Claire, for their unwavering friendship and support.

I am also grateful to the probability group in Durham, and those members of PiNE who created an inclusive community in which to work giving a brilliant platform for early stage researchers to develop. The reading group, quarterly meetings and coffee room discussion have been great!

Thank you to Vlad and Mikhail for their thorough comments and suggestions for improvement.

Sist, men ikke minst, jeg vil gjerne takke Lina. Takk for at du har vært så tålmodig de siste 4 årene og jeg gleder meg til fremtiden.

This work was supported by an EPSRC studentship (EP/M507854/1).

Contents

List of Figures

Chapter 1

Introduction

In a world where financial output requires justification, it seems appropriate to open by justifying why we should study convex hulls of random walks – or limit theorems for convex hulls of random walks which could be a more specific title for this thesis.

There is no denying the relevance of random walks. Whether you are studying the stock market, disease spreading, election campaign policy or animal movements, you could consider the change in price, infection rate, popularity or location as a random incremental change and so, by studying an associated random walk, you could estimate many quantities of interest.

What about convex hulls though? We discuss specific applications below, but we can also consider Occam's razor which is the notion that the simpler solution can often be more correct, or at least that it is often better to appeal to the simplest solution. With respect to random walks, more often than not, this theory is applied – how many times is the exact distribution of the nth step of a random walk described? Instead we simplify our question to something more tractable such as recurrence or transience, or a simpler probability bound. Convex hulls are nothing more than another extension of this simplification. If one knows about the shape and size of a convex hull, upper and lower bounds on the shape, size and position of the walk can be found.

Finally, but similarly, why consider limit theorems? Again, this gives a simpler framework in which to work, but that is not all. In fact, after a little thought,[1] it is entirely

[1] In this case, after discussing a similar question with Nick Bingham, to whom I am grateful for his comments and suggestions.

natural to ask about limits to infinity, after all, the natural numbers are an infinite set themselves. Any process which recurs every hour, minute or second can be associated with this infinite set, and more often than not, if a limit exists, we will see this limiting behaviour if we watch the process for long enough. Just ask your favourite casino's owner if they agree!

In this rest of this chapter we will discuss some of the existing work specific to random walks, convex hulls, and the crossover of these two topics. The first of these topics is a vast area of research in itself and one which we cannot cover in its entirety. Thus we must choose to focus on a subset of the topic, here presenting results concerning fluctuation theory which we feel leads quite naturally into the study of convex hulls. After this historical overview, we describe some specific applications of convex hulls and convex hulls of random walks across the natural sciences. We then provide a brief outline of the remainder of the thesis and introduce some specific random walks which will be used to demonstrate our results in action.

Finally, we end this chapter by going on a quick tour of some relevant mathematical concepts that are used in the material of this thesis. This section will be particularly useful for readers unfamiliar with the topics covered, but for the more advanced reader, this section can be used simply for reference. The reason for the length of this introduction is two-fold. Firstly to be comprehensive, but also because our results are built on these foundations many which are generic or technical results in random walk theory that would hinder the flow of the main body of work. We also use this section as a reference guide for any notation specific to this thesis.

1.1 Historical context and applications

This section is separated into four parts. We begin with introductions to random walk theory, and problems concerning convex shapes and convex hulls of random points. These two subsections are deliberately conversational to give a flavour of the context in which the more specific topic of convex hulls of random walks sits. This is the third subsection, which we discuss in more depth with results explicitly stated with as much of the current literature discussed as possible. We finish with a brief discussion of areas of science where convex hulls are and could in the future be applied.

1.1.1 Random walk theory

The term random walk was first coined by Karl Pearson in 1905 when he wrote a letter to Nature asking for help from the readership [Pea05a]. His posed problem was that of a man walking a distance of l yards from his starting position, then changing direction (or not) randomly, before walking a further l yards, until he had made n movements. In particular, Pearson wondered what the probability is of the man being in an annulus a given distance from his starting position after carrying out this procedure.

The response came from Lord Rayleigh [Ray05], informing Pearson that he had already studied the problem several decades earlier in a different context. His focus was concerning sound wave vibrations of a fixed pitch (magnitude) but varying phase (angle) [Ray80]. Both a summary of Rayleigh's result and one further key contribution to the field, namely the analogy of a drunken man, came in the next edition of Nature where Pearson wrote [Pea05b]:

> The lesson of Lord Rayleigh's solution is that in open country the most probable place to find a drunken man who is at all capable of keeping on his feet is somewhere near his starting point!

Further to Pearson's and Rayleigh's work, notable mentions should go to Louis Bachelier, who was the first to link the topic of mathematical finance to these probabilistic processes in his doctoral thesis of 1900 [Bac00], and the little known scientist Albert Einstein, who provided a theory to Robert Brown's observations on the movement of particles suspended in a liquid, which he termed Brownian motion [Ein05; Bro28]. We will not discuss any applications to finance, but Brownian motion certainly features heavily throughout the work.

With the theory now known in the scientific community, there were many directions of study which were pursued across the globe. One such example is the dichotomy between recurrence and transience (the notions of returning to any given state eventually against visiting states only finitely many times) which for a particularly simple random walk led Shizuo Kakutani to follow in Pearson's footsteps with his jest [Dur10, p.191]:

> A drunk man will eventually find his way home, but a drunk bird may get lost forever.

That is, in two dimensions the process is recurrent but in three dimensions it becomes transient and thus might never return home.

In this work, we consider results on the convex hull of random walks. There are plenty of previous works on this specific topic, but even before the term convex hull entered this area of study, there were many works on the extreme values, and in particular maxima, of random walks. Through Cauchy's surface area formula for convex shapes, see e.g. [Gru07, p.106], which gives an expression for the surface area as the integral of projected lengths over all angles, we see that these results are closely related, even if the original authors were not necessarily aware of this fact.

The study of extreme values of random walks became known as fluctuation theory. The early results were largely related to the proportion of time that a one dimensional walk is on the positive side of the origin. We will denote the positions of the walk at time n as S_n and use T_n to denote the proportion of time on the positive side up to time n. One early paper which sparked an abundance of work, was Lévy's paper 'Sur certains processus stochastiques homogènes' [Lév40b]. Here, Lévy proved an arcsine law for T_n where the underlying walk was the simple symmetric random walk, that is $\mathbb{P}(S_{n+1} - S_n = 1) = \mathbb{P}(S_{n+1} - S_n = -1) = 1/2$ and the $(S_{n+1} - S_n)_{n\geq 1}$ are independent with $S_0 = 0$, which stated

$$\lim_{n\to\infty} \mathbb{P}(T_n < x) = \frac{2}{\pi} \arcsin \sqrt{x}, \text{ for } x \in (0,1).$$

Intuitively, this result says that over a long period the the walk is more likely to spend the majority of time on one side or another, and in fact, the least likely outcome is to spend around half the time on each side of the origin.

From 1947 to 1952 this result was generalised again and again, first to walks where the increments had mean 0 and variance 1 by Erdős and Kac [EK47], then Sparre Andersen twice published the result depending on some conditions of symmetry which allowed the necessity for the increments to be independent of one another to be relaxed [SA49; SA50], and then Maruyama and Udagawa separately relaxed the conditions for these latest results, only requiring a central limit theorem to hold for the walks [Mar51; Uda52]. These results continue to be of interest to the present day. Sparre Andersen's results on symmetric increments have been generalised in the last few years by Kabluchko, Vysotsky and Zaporozhets [KVZ16] to allow for higher dimensions, which of course requires generalisation of what is meant by 'the positive side'; in their work they consider absorption of the origin in the convex hull, particularly appropriate for our current context. This generalisation can be compared to the choice of Bingham and Doney [BD88], who, in 1988 presented results on arcsine laws for Brownian motion in higher dimensions, where they considered 'the positive side' to be taken as having all components positive.

Many of the earlier arcsine law results were established using combinatorial arguments, but at the same time, Chung and Feller [CF49] applied the powerful tool of generating functions to confirm Erdős and Kac's result in the case of the simple symmetric random walk and also to determine a nice result which contrasted with those found by Lévy. In particular, if N_{2n} is the number of steps of a $2n$ step walk where the position directly before the step is taken or the position directly after the step is taken (or indeed both of these positions) is positive, then for any $r \in \{0, \dots, n\}$

$$\mathbb{P}(N_{2n} = 2r \mid S_{2n} = 0) = \frac{1}{n+1}.$$

This appears in contrast to the arcsine law, because it states that a walk returning to the origin at a given time has an equal probability of having spent any viable proportion of time on either side of the origin up until its moment of return. Specifically, the walk is just as likely to have half its steps on the positive side as it is to have spent all or no time on the positive side of the origin.

As Erdős and Kac generalised Lévy's results, this theorem was generalised by Lipschutz, in [Lip52], to walks with mean 0 and variance 1 on the condition they had a finite 4th moment. A few years later, Baxter presented a paper on 'Wiener process distributions of the "arcsine law" type', which apart from citing Erdős and Kac's paper, required only one cited result, Chung and Feller's uniform distribution result. After this, there was a period where Chung and Feller's paper, although the results were well-known, did not attract much attention, however this is possibly due to Feller's seminal book, first appearing in 1950, which contained a chapter on the fluctuations of coin tossing, including this theorem [Fel68, pp.67–97]. Despite this, since the millennium there has been a renewed interest in Chung and Feller's paper itself, with over 20 references, including one particular paper which used combinatorial methods in the style of Sparre Andersen to investigate a topic of interest to us - maxima of random walks [HW16].

This point in time, the early 50's, was the beginning of these functionals being considered together. Feller's book was published, Chung and Erdős had described a generalisation of the Borel-Cantelli lemma which they applied to the number of zeros and number of positive terms of the simple symmetric random walk [CE52], and Lipschutz developed this paper further to a whole class of "number of" events [Lip53]. Similarly, Darling [Dar51] considered random walks with symmetric increments, establishing a theorem on the ordering of the random walk points. In turn, this was shown to give a distribution for the position of the maximum and a distribution for the number of walk points which are positive.

One further paper which considered the first time a walk attains its maximum, last time it attains its minimum, and the number of positive walk points, was the first of two papers titled 'On the fluctuations of sums of random variables' by Sparre Andersen [SA53]. The second such paper [SA54], published the very next year, is, to the author's knowledge, the first serious consideration of the convex hull, or at least the convex minorant, of a random walk and was the product of this ongoing study of functionals, or fluctuations of random walks.

The study of random walks has of course continued beyond the specific topic of their convex hulls. One recent, particularly active researcher with over 200 papers, of which many relate to random walks, some specifically to convex hulls of random walks is Satya Majumdar. Some papers are natural extensions of the results in the 1950's, with several of his works with co-authors Mounaix and Schehr considering not just the first maximum but the first two maxima of random walks [MMS13; MMS14; MSM16], and also his paper, this time only with Schehr, using similar order statistic ideas as Darling did in his work 60 years previously [SM12]. We leave further study of random walks, not with a convex hulls focus, to the interested reader and could suggest a whole host of books on the subject but suffice by suggesting the following [Rév13; Fel68; Fel71; Dur10; Gut05; Kal02; LL10; Spi76; MPW17].

1.1.2 Convex hulls

Completely separated from the study of random walks, the study of convex shapes and convex hulls has been a source of persistent interest going as far back as the works of Archimedes in the 3rd century BC, see e.g. 'The Works of Archimedes' by Heath [Hea97] or the nice commentary by the late Stephen Hawking in his book 'God Created the Integers' [Haw06, pp. 119–239]. Indeed, it is suggested by Gruber [Gru07, p.41] that Archimedes was the first to explicitly define convexity with the axioms in his work 'On the Sphere and the Cylinder'.

Convex sets themselves appear in many branches of applied science including other areas of mathematics. Of course, convex analysis and optimization relies heavily on convex sets and convex functions, see e.g. [Roc70], and can be applied to many useful problems such as the solution space in the simplex method of operations research, see e.g. [FP93; Sai95]. Other natural sciences also make use of convex sets, including the idea of balance in consumption in economics [NS08, p. 94] and in ecological studies of species competition [ML64].

In terms of convex hulls of random points in mathematics, there has been considerable work throughout the last century, with many simple-to-state puzzles probably being the foundation of the work. Such problems include the combinatorial problem posed in a paper by Erdős and Szekeres [ES35] which asks how many points you need in the plane, with no 3 lying on a straight line, such that you can be certain of finding a subset of n points such that the resulting n-gon is convex. Both the proof for the original case concerning quadrilaterals, and the more general question regarding n points are attributed to Esther Klein, who went on to marry Szekeres giving the problem its nickname of the 'happy ending problem'. A similar problem is Sylvester's four point problem, unsurprisingly posed by Sylvester in the Educational Times in 1864 [Syl64] and nicely discussed by Pfiefer in his paper of 1989 [Pfi89]. The problem is to show that the probability of 4 points taken 'at random in an infinite plane' of forming a non-convex polygon is 1/4. The problem is ill posed with differing solutions attained depending on the interpretation of the random selection method. However, this problem posed as randomly selecting points from a certain finite convex plane (e.g. a circle or a specified polygon) where the randomness is considered as uniformly chosen points in the given plane (i.e. a point has a probability of falling in a given subset equal to the proportion of the area of the set covered by the subset), has been solved [Wat65; Woo67]. As has the generalisation of considering a choice of $n+2$ points in the n-dimensional unit ball, and asking whether their convex hull contains $n+1$ points [Kin69; Gro73]. For nice discussion of such results see e.g. [KM63; Pey97].

These problems were then reconsidered in the 1960's from a different perspective; instead of how many vertices to form a convex n-gon or the probability of a single point being inside the convex hull, what is the expected number of vertices in the hull? Or what is the expected perimeter length and area of the hull of these n random points? This was what Rényi and Sulanke [RS63; RS64], Efron [Efr65], and more recently Massé [Mas99; Mas00] and Reitzner [Rei03] considered in their papers.

One extension to these questions was considered by Rogers [Rog78]: whether two sets of points in the plane have disjoint convex hulls. Jewell and Romano then showed that this problem, in a simple form, was equivalent to considering the probability that a given number of arcs of fixed length, when randomly placed on the circle, would form a cover of the circumference [JR82]. Problems of this type have continued to be solved in recent years. Reitzner considered the same problem as Rogers but where the sets of points were restricted to lying in a convex body themselves [Rei00], and Groeneboom considered a similar problem to Rényi and Sulanke of the number of vertices in the convex hull of n points, but now restricted to lying in a convex polygon [Gro12].

A further set of interesting results related to convex shapes are the Bárány-Vershik-Sinai results on the limit shapes of convex polygons. So-called because each of the three authors independently proved similar results in $1994-95$, see [Bár95; Ver94; Sin94]. As Bogachev and Zarbaliev state in their paper generalizing the theorems, see [BZ11], the results concern the limit shape of a typical convex curve from some set of convex curves. For example, Bárány's first theorem considers the typical shape of a convex polygons in the square $[-1,1]^2$ with vertices on the lattice $n^{-1}\mathbb{Z}^2$ as $n \to \infty$. If we consider a point $x \in [-1,1]^2$, and at each step n, calculate the proportion of the convex polygons for which the point is in the interior, call this $\rho_n(x)$, then there exists a limiting shape L for which $\rho_n(x) \to 1$ if $x \in \operatorname{int} L$, the interior of L and $\rho_n(x) \to 0$ if $x \notin L$. The shape L is the convex set $L = \left\{(x,y) \in \mathbb{R}^2 : \sqrt{1-|x|} + \sqrt{1-|y|} \geq 1\right\}$.

This is just a flavour of some of results that have been studied on the convex hulls of random points, a topic which continues to be studied right up to the present day, e.g. [KRR18]. However, the focus of this work is more specifically on convex hulls of random walks, so for further detailed discussion of hulls of random points, the survey by Majumdar is recommended [MCRF10]. This survey also covers random walks but we aim to give more extensive and up to date coverage of this area of study in what follows.

1.1.3 Convex hulls of random walks

We begin our discussion of convex hulls of random walks with Sparre Andersen's results from [SA54]. As Majumdar et al. mention in their survey [MCRF10], Lévy had commented, somewhat heuristically, on the shape of the curve of Brownian motion [Lév48], which is in effect the convex hull, but Sparre Andersen seems to be the first to provide some rigorous results relating to such concepts. After the majority of the paper where Sparre Andersen considered random variables such as the first time to attain the maximum value, as was discussed above, he presented results on the number H_n of values $i = 1, \ldots, n-1$ such that S_i coincides with the largest convex minorant of the sequence $S_0, \ldots, S_n$. Here, the increments of the walk are denoted $Z_i := S_i - S_{i-1}$, and these increments are one-dimensional. Thus Sparre Andersen defines the convex minorant in terms of sequences of numbers: a sequence $b_0, \ldots, b_n$ is called convex if the sequence $b_1 - b_0, b_2 - b_1, \ldots, b_n - b_{n-1}$ are non-decreasing, and then a sequence $a_0, \ldots, a_n$ has a unique, largest, convex minorant sequence $b_0, \ldots, b_n$. This sequence always has $b_0 = a_0$ and $b_n = a_n$, and then either $b_i = a_i$ for $i = 1, \ldots, n-1$ or $b_i = (k-j)^{-1} \left((k-i)a_j + (i-j)a_k\right)$ where k is the smallest index larger than i such that $b_k = a_k$ and j is the largest index smaller than i such that $b_j = a_j$. For the random walk, we construct the convex minorant from the sequence $a_i = S_i$, $i = 0, \ldots, n$.

Graphically, the walk and convex minorant can be represented by plotting the time-space diagram of the random walk, and drawing the shortest path from S_0 to S_n which, at any give time, always has a spatial value less than or equal to the spatial value of the random walk at the same time value. For example, in Figure 1.1 we have the time-space diagram of a random walk, in black, with $n = 10$ and its convex minorant, in green. The point at $(4, -3)$ represents S_4 which is one of the walk points where $b_i = S_i$ because the walk and convex minorant intersect. The point at $(8, -1)$ is b_8 which is an interpolation of $b_7 = S_7$ and $b_9 = S_9$. For this walk, $H_n = 4$ because the walk and convex minorant intersect at 4 indices, not including the points at 0 or $n = 10$.

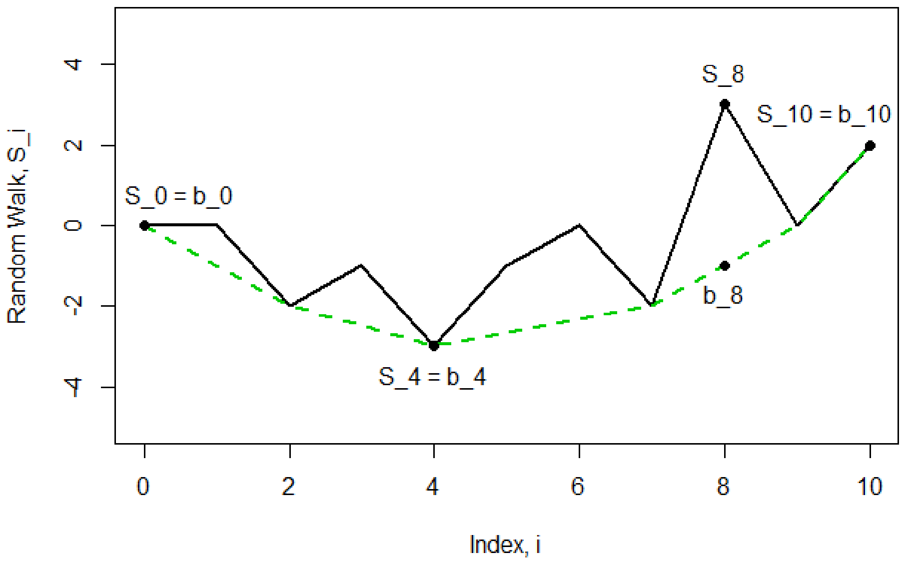


Figure 1.1: Convex Minorant (green) of a random walk (black)

In the paper, Sparre Andersen describes the distribution of H_n on the condition that the increments Z_i are independent and drawn from continuous distributions. He does this by establishing the generating function as

$$H_n(t) := \sum_{m=0}^{n-1} \mathbb{P}(H_n = m)t^m = n^{-1} \prod_{m=1}^{n-1} (1 + m^{-1}t),$$

which coincides with the generating function for a sum of $n - 1$ Bernoulli random variables $Y_1, \ldots, Y_{n-1}$, which take the value 1 with probability $(i + 1)^{-1}$ and 0 with probability $i(i + 1)^{-1}$. This enables us to establish the properties

$$\mathbb{E}(H_n) = \sum_{i=1}^{n-1} (i + 1)^{-1}, \qquad \mathbb{V}\text{ar}(H_n) = \sum_{i=1}^{n-1} i(i + 1)^{-2},$$

which tells us that we should expect approximately $\log n$ random walk points to lie on the convex minorant. Of course, this result also says that we would expect approximately $\log n$ points to lie on the concave majorant, by symmetry, and so the convex hull which is just the convex minorant concatenated to the concave majorant is also expected to have approximately $2 \log n$ vertices, and thus this many faces.

Convex minorants have applications themselves in isotonic estimation in statistics - the estimation of functions which are known to be non-decreasing in some way. An example of such a study is the paper by Leurgans from 1982 [Leu82].

Further study of convex minorants in the period immediately after Sparre Andersen's work was largely from a combinatorial perspective. Spitzer [Spi56] used cyclic permutations to relate the characteristic function of $\max(S_0, \ldots, S_n)$ to the sum of the characteristic functions of $\max(0, S_i)$ for $i = 1, \ldots, n$, a result which Brunk generalized [Bru64]. This, along with further results on cyclic permutations by Stam [Sta83], were used by Goldie in 1989 [Gol89] to analyse the convex minorant of a one dimensional random walk, with the conditions Sparre Andersen had described. Let an increment Z_i 'belong to the jth side of the greatest convex minorant' if exactly j of the random walk points from $S_0, \ldots, S_{i-1}$ have the property $S_k = b_k$ for $k = 0, ..., i-1$. Then, considering the increments and the sides to which they belong, Goldie established that the event, A_i, that the ith smallest increment belongs to a new side, i.e. one which none of the $i-1$ smaller increments belong to, has $P(A_i) = 1/i$ with all the A_i, $i = 1, \ldots, n$ independent.

Interestingly however, Qiao and Steele proved in 2002 that the concave majorant of a random walk consists of a single line infinitely often [QS05]. This contrasts to the results of Goldie and Sparre Andersen which both suggest that we would expect there to be order $\log n$ faces when considering a fixed length of the walk. We pick up this theme in Chapter 6.

Another combinatorial paper which came after Sparre Andersen and Spitzer's combinatorial lemma was that of Baxter on 'A combinatorial lemma for complex numbers' [Bax61]. In terms of random walks, Baxter was considering a two dimensional walk where for any two vectors, each created by adding together a non-empty subset of consecutive increments, must not be parallel, which is satisfied with probability 1 if the increments are drawn from a continuous distribution. Then he noted that there is only one cyclic permutation of the increments $Z_1, \ldots, Z_n$ such that the random walk stays positive throughout. Further to this, any edge of the convex hull is made up of a sum of some subset, A, of the increments. If we say $|A| = m$, then the edge created by adding this specific subset appears in exactly $2(m-1)!(n-m)!$ of the possible permutations (not restricting to cyclic permutations). Using only these two properties, Baxter established that we expect exactly 2 of the increments $Z_1, \ldots, Z_n$ to be edges of the hull. He also verified the result of Sparre Andersen that we would expect $2 \log n$ faces in the hull. Finally, he verified the Spitzer-Widom formula on the expected perimeter length of the convex hull, which we will discuss below. Two years later, in work with Barndorff-Nielsen, Baxter generalized his results to higher dimensions [BNB63].

At the same time as these results, Spitzer, in collaboration with Widom [SW61], considered the expected perimeter length of the convex hull of a planar random walk. This question has been conveyed by another analogy involving our drunken friend, this time as a gardener, in [WX15a]:

> On each of n unsteady steps, a drunken gardener drops a seed. Once the flowers have bloomed, what is the minimum length of fencing required to enclose the garden?

Spitzer and Widom approached this question with the usual combinatorial mindset, but this time in conjunction with Cauchy's surface area formula for convex shapes, see e.g. [Gru07, p.106]. The formula states that the perimeter length, L, of a convex shape can be determined as

$$L = \int_0^\pi D(\theta) d\theta, \tag{1.1.1}$$

where $D(\theta)$ is the length of the projection of the shape onto a line with direction θ. In random walk terms, the convex shape is the convex hull, and thus to calculate the perimeter length we consider

$$D(\theta) = \max_{0 \le i \le n} S_i \cdot \mathbf{e}_\theta - \min_{0 \le i \le n} S_i \cdot \mathbf{e}_\theta,$$

where $\mathbf{e}_\theta$ is the unit vector in direction θ. The combinatorial identity used alongside Cauchy's formula was a lemma of Kac [Kac54], but he attributes the concise proof of the lemma to Dyson. This lemma requires us to consider all the permutations of the n increments, so let $\pi : 1, \ldots, n \mapsto \pi_1, \ldots, \pi_n$ be such a permutation. Then the result states

$$\sum_{\pi}\left(\max_{0\le i\le n} S_{\pi_i} - \min_{0\le i\le n} S_{\pi_i}\right) = \sum_{\pi}\sum_{i=1}^{n}\frac{1}{i}\|S_{\pi_i}\|,$$

where the S_{π_i} must be real numbers and the notation used for the first sum on both sides means over all possible permutations. Of course having all S_{π_i} real is not what was promised, this was Spitzer and Widom's contribution, to combine the two results so that Kac's lemma became

$$\sum_{\pi} L_n^{\pi} = 2\sum_{\pi}\sum_{i=1}^{n}\frac{1}{i}\|S_{\pi_i}\|,$$

where L_n^{π} is the perimeter length of the convex hull of the random walk under the permutation π. The remarkable theorem that follows is the equation that arises when we take expectations,

$$\mathbb{E}\, L_n = 2\sum_{i=1}^{n}\frac{1}{i}\,\mathbb{E}\,\|S_i\|. \tag{1.1.2}$$

Despite the elegance of this result, there was very little in the years that followed other than the papers by Baxter mentioned above. It wasn't until 1993 that Snyder and Steele [SS93] further studied the distribution of L_n and established an upper bound on its variance. Specifically, letting $\mu := \mathbb{E}\, Z$,

$$\mathbb{V}\mathrm{ar}(L_n) \le \frac{\pi^2 n}{2}\left(\mathbb{E}(\|Z\|^2) - \|\mu\|^2\right),$$

if the increments of the random walk $Z_1, \dots, Z_n$ are all independent and distributed like Z. Then, if $\mathbb{E}\,\|Z\|^2 < \infty$, this is sufficient to show that

$$n^{-1}L_n \xrightarrow{\text{a.s.}} 2\|\mu\| \text{ as } n \to \infty. \tag{1.1.3}$$

See also Theorem 2.1.1 and Theorem 3.3.11 below. They also established bounds on the tail probabilities of $L_n - \mathbb{E}\, L_n$ but only in the case where the increments are bounded. Further, they also used Baxter's combinatorial lemma to reaffirm several known results in different ways to the previous expositions, however they also showed it was possible to establish further results such as the expected sum of squares of the face lengths of the convex hull, $L_n^{(2)}$ for which they established,

$$\mathbb{E}\, L_n^{(2)} = 2n\left(\mathbb{E}(\|Z\|^2) - \|\mu\|^2\right).$$

Steele then continued to consider combinatorial results which could help in the study of such functionals of convex hulls and in 2002 presented his paper on the Bohnenblust-Spitzer algorithm [Ste02]. The application of the combinatorial formulae established by the algorithm gives further results on the distribution of functionals such as the number of faces (of the time-space diagram of a one dimensional walk), but not on the variance of the perimeter length. However, after establishing the expected length of the concave majorant was approximately $n\sqrt{1+\mu^2}$ for large n, Steele makes a passing comment that has particular relevance to our later work where we prove his intuition to not only be correct for the one dimensional time-space diagram but for two dimensional walks as well [Ste02, p241],

> This interesting geometric formula tells us that the expected length of the concave majorant grows exactly like the length of the line from $(0, 0)$ to the point $(n, \mathbb{E}(S_n)) = (n, n\mu)$.

The question of improving the upper bound on the variance of L_n to at least an asymptotic result was answered by Wade and Xu in two papers in 2015 [WX15a; WX15b]. In the first, they considered the case where $\|\mu\| > 0$, the case with drift. In this work, they showed (Theorem 1.3 of [WX15a]) that if $\mathbb{E}(\|Z\|^2) < \infty$ and $\mu \neq \mathbf{0}$, then, as $n \to \infty$,

$$n^{-1/2}|L_n - \mathbb{E}\, L_n - 2(S_n - \mathbb{E}\, S_n) \cdot \hat{\mu}| \to 0, \text{ in } L^2. \tag{1.1.4}$$

This result was enough to establish the asymptotic expression for the variance,

$$\lim_{n\to\infty} n^{-1}\, \mathbb{V}\text{ar}\, L_n = 4\, \mathbb{E}\left(((Z-\mu)\cdot\hat{\mu})^2\right) \tag{1.1.5}$$

where we have used $\hat{\mu}$ to mean the unit vector in the direction of μ. In turn, this was enough to describe a central limit theorem for L_n in the case where the right hand side of (1.1.5) was non-zero. The exceptional case refers to the walks where there is no variance in the direction of the mean and so include the time-space diagrams of one dimensional walks. This is the topic of Chapter 6.

In the second of the two papers [WX15b], the authors considered the convergence of the convex hull of the random walk to that of Brownian motion by using a continuous mapping argument and Donsker's theorem. In particular, they showed that, for walks with $\mu = \mathbf{0}$, as $n \to \infty$,

$$n^{-1/2} L_n \xrightarrow{\text{d}} \mathcal{L}(\Sigma^{1/2} h_1), \qquad \text{and} \qquad n^{-1} A_n \xrightarrow{\text{d}} \mathcal{A}(\Sigma^{1/2} h_1) = a_1 \sqrt{\det \Sigma},$$

where we have used the notation $\mathcal{L}$ and $\mathcal{A}$ to mean the perimeter length and area of a set, respectively, $\Sigma := \mathbb{E}[(Z - \mu)(Z - \mu)^\top]$, the covariance matrix for the increments, and h_1 and a_1 to be the hull of Brownian motion run for unit time and the area of said hull, respectively. From this distribution result, they established convergence of the mean of L_n in the zero drift case as

$$\lim_{n \to \infty} n^{-1/2} \, \mathbb{E} \, L_n = 4 \, \mathbb{E} \, \|Y\|,$$

if $Y \sim \mathcal{N}(0, \Sigma)$ is a Normal random variable. Likewise, they found

$$\lim_{n \to \infty} n^{-1} \, \mathbb{E} \, A_n = \frac{\pi}{2} \sqrt{\det \Sigma}.$$

For the case where there is drift, $\|\mu\| > 0$, the hull does not converge to that of two dimensional Brownian motion, but it does to the convex hull of the space-time diagram of one-dimensional Brownian motion, $\tilde{h}_1$. This led to the result: if $\mathbb{E} \, \|Z\|^p < \infty$ for some $p > 2$, and $\sigma^2_{\mu_\perp} > 0$ where $\sigma^2_{\mu_\perp}$ is the variance in the direction perpendicular to the mean, formally defined at (1.3.5) below, then

$$\lim_{n \to \infty} n^{-3/2} \, \mathbb{E} \, A_n = \frac{1}{3} \|\mu\| \sqrt{2\pi\sigma^2_{\mu_\perp}}. \tag{1.1.6}$$

The variance of these two functionals was also studied with the following convergence to the variance of the respective quantity in terms of Brownian motion as follows:

- Suppose $\mu = \mathbf{0}$ and $\mathbb{E} \, \|Z\|^p < \infty$ for $p > 2$, then

$$\lim_{n \to \infty} n^{-1} \, \mathbb{V}\text{ar} \, L_n = \mathbb{V}\text{ar}(\mathcal{L}(\Sigma^{1/2} h_1)).$$

- Suppose $\mu = \mathbf{0}$ and $\mathbb{E} \, \|Z\|^p < \infty$ for $p > 4$, then

$$\lim_{n \to \infty} n^{-2} \, \mathbb{V}\text{ar} \, A_n = \mathbb{V}\text{ar}(a_1) \det \Sigma.$$

- Suppose $\|\mu\| > 0$ and $\mathbb{E}\,\|Z\|^p < \infty$ for $p > 4$, then

$$\lim_{n\to\infty} n^{-3}\,\mathbb{V}\mathrm{ar}\,A_n = \mathbb{V}\mathrm{ar}(\mathcal{A}(\tilde{h}_1))\|\mu\|^2\sigma^2_{\mu_\perp}.$$

It is no mistake or oversight that these results do not include an equivalent result for L_n in the case $\|\mu\| > 0$ which would of course agree with the first of the two papers. The fact that the limiting object in this case is the space-time diagram of one dimensional Brownian motion, means that the scaling in the time direction is n and in the space direction is $n^{1/2}$ which explains the $n^{-3/2}$ in (1.1.6). Unfortunately, under this different scaling knowing only the length of an edge of the hull is not enough to know how it scales, the angle of the edge is also required to determine the scaled length. Thus, knowing the perimeter length without more details on the angles of the edges of the hull is also not enough.

One result that has been attained for the perimeter length is a large deviation result. Akopyan and Vysotsky [AV16] have shown $\mathbb{P}(L_n \geq 2cn)$ for $c > \|\mu\|$ decays exponentially, and likewise for deviations on the lower side.

Finally on the perimeter length and area, there are also some results when some central symmetry is assumed with continuous increments. Grebenkov, Lanoiselée and Majumdar [GLM17] found expansions of $\mathbb{E}\,L_n$ in this case, showing that, when we have finite variance, the second term of the expansion, after the $n^{1/2}$ term, is in fact constant. If we do not have finite variance, then the order of the terms in the expansion depends on which is the largest moment that is in fact finite in the density function of the increments. The authors also established similar results for the expansion of $\mathbb{E}\,A_n$ but this was only for the case of Gaussian increments.

There are other functionals of the convex hulls of random walks which have also been studied. In [KVZ17b], Kabluchko, Vysotsky and Zaporozhets determined the expected number of faces of the convex hull. Vysotsky and Zaporozhets had previously studied the probability that a multidimensional walk with centrally symmetric increments absorbs the origin into its hull in [VZ18], but their distribution-free results only were proven in two dimensions, however in a later work, also with Kabluchko, they were able to complete the proof using a different methodology [KVZ17a]. The method of Vysotsky and Zaporozhets did have the advantage of proving a multi-dimensional generalisation of the Spitzer-Widom formula.

Some earlier studies had also considered other functionals. In fact, they considered any functional, Ψ, of the convex hull that is monotone with respect to the convex hull set, and satisfied an affine scaling property. The first of these papers was by Khoshnevisan [Kho92a], in which he proved a law of the iterated logarithm result: let $\mathcal{H}_n$ be the convex hull at time n, then for some $\alpha > 0$ which is determined by the affine scaling property,

$$\limsup_{n\to\infty} \frac{\Psi(\mathcal{H}_n)}{(2n \log\log n)^{\alpha/2}} = c_\Psi \text{ a.s.}$$

where c_Ψ is a deterministic constant depending on the choice of functional. Also in this paper, Khoshnevisan proved the related lower bound,

$$\liminf_{n\to\infty} \left(\frac{\log\log n}{n}\right)^{\alpha/2} \Psi(\mathcal{H}_n) = c'_\Psi \text{ a.s.}$$

where α and Ψ are as before, but c'_Ψ is a different deterministic constant. Both this paper, and the second paper by Kuelbs and Ledoux [KL98], were actually focused on convex hulls of Brownian motion which we are about to discuss below. The contribution with respect to random walks of the Kuelbs and Ledoux paper was to clear up some edge cases which required some careful consideration beyond Khoshnevisan's proofs.

Prior to the contents of this thesis, this is the extent of what was known about the convex hulls of random walks. However, there has been considerable study of the convex hulls of Brownian motion, which the results above indicate is a closely related topic.

Important and relevant works in this area are the Ph.D. thesis of El Bachir [EB83] which expected area, the useful paper of Eldan [Eld14] which established explicit formulae for the volumes of n dimensional Brownian motion, and the note by Takács [Tak80] answering a question by Letac [Let78] on the expected perimeter length of standard Brownian motion. Some slight variants on the theme of convex hulls of Brownian motion have been studied by Majumdar and a host of co-authors; in [RFMC09] the convex hull of multiple Brownian motion paths is studied, and in the two papers [CBM15a; CBM15b] a single Brownian motion but with a restriction on the plane to have a reflecting wall is considered. Majumdar has also been involved with a numerous papers which use a numerical approach to estimate the distributions of the volume and surface area of the convex hull of Brownian motion in higher dimensions [SHM17], a single walk in two dimensions that is not necessarily Brownian [CHM15], multiple random walks in two dimensions that are not necessarily Brownian [DCHM16], and of self avoiding random walks [SHM18]. This is not an exhaustive list of results on convex hulls of Brownian motion, so for further results in the area see [Kho92b; KZ16; RF13] and the references therein.

A slightly broader view is taken when studying the convex hulls of Lévy processes, a class which includes Brownian motion. Until recent years very little work had been done in this area but now there are a few references [KLM12; MW16; RFW17; RF14]. Some results on the convex minorant of Lévy processes can also be found in [PUB12] and this work, along with some similar results for Brownian motion and other processes is summarised in the survey paper [APRUB11].

This is not all the processes that have been studied either. In [RMR11] Cauchy's formula is used to study the expected perimeter length and the expected area for a random acceleration process in two dimensions which is not even Markovian.

Most of the random walk results depend on some combinatorial identities and possibly use Cauchy's formula too. However, this approach that Wade and Xu used in their paper on the zero drift case [WX15b], of considering the limit to Brownian motion and then studying that process is a strategy that, unbeknownst to Wade and Xu, was also used in the non-zero drift case in the book by Whitt [Whi02] to study random walks without considering the convex hull. This strategy is also one which we employ in Chapter 3.

1.1.4 Applications

As far as applications of convex hulls of random walks are concerned, the most cited is the application to ecology and the home range of animals, see [Wor95; Wor87]. This idea was pursued by Luković, Geisel and Eule [LGE13], who studied the convex hulls of some continuous random walks which they compared to the search strategies of Mediterranean seabirds and animals ambushing their prey. This has further relevance to bridges, random walks where the end point is fixed to be back at the origin, and to multiple walks, which can be used to model the foraging pattern of an animal or pack of animals, that return to a fixed location to sleep each night, as described in e.g. [GLM17].

Convex hulls themselves have many applications to more specific situations than the convex set appearances mentioned earlier. In statisical analysis, convex hull peeling, sometimes referred to as the onion layer problem, is a method used to determine an ordering of central tendency of points in a data set. By creating the convex hull of the data, removing the points in the hull boundary and calling these the least central points, then repeating, we can establish a grouping of the data. Knowing about the characteristics of the convex hull is particularly important in order to establish the efficiency of this procedure. Some works on this topic are by Eddy [Edd82] and Brozius [Bro89].

Another application which specifically uses the convex hulls is in pattern recognition within images where algorithms to find a convex hull are often reported on in the hope to speed up computer programs. Many references exist in this area including [AT78; MT85; Hus88; Ye95].

In biology and medicine, convex hulls are also used to both approximate the surface of a protein, which is particularly useful in helping to identify the situations in which a given protein could be useful [MAHPS95]. Similarly, convex hull classification algorithms are used to identify proteins [YMBH15], or even predict psychosis onset [Bed+15].

This is only a selection of the uses of convex hulls to demonstrate the possibility that convex hulls of random walks could find further uses in the future beyond the study of the home range of animals.

1.2 Thesis outline

First, as mentioned above, we give a detailed introduction to the theory required, starting with basic probability theory and building to some specific results relating to random walks. Then, we will describe the examples which will be used in simulations throughout the thesis and briefly mention some comments on how we carried out the simulations.

In Chapter 2 we present some law of large numbers results related to the perimeter length of the convex hull, diameter and ratio of the two functionals. Then we establish the first few terms of the expansion of the expectation of the perimeter length and the diameter. We also demonstrate these results through simulations, including plots of the walks we will be using for our examples.

We then move to functional limit theorems. In this, the longest chapter of the thesis, we present the key results, namely the functional law of large numbers and functional central limit theorem which, with the continuous mapping theorem, allow us to determine the convergence of the convex hull and related functionals. Basic examples include the maximum functional and a generalisation of the arc-sine law, but we also establish further convergence results for our two main functionals, the perimeter length of the convex hull and the diameter of the hull. Brownian motion is the limiting object in some of the results, so this chapter also includes a discussion of the diameter of planar Brownian motion, in particular improving what is known about the expectation of this diameter.

In Chapter 4 we look further into the shape of the convex hull, using the notion of the zero drift walk converging to Brownian motion that was discussed in Chapter 3. We establish a zero-one law, and then find that the ratio of the perimeter length and diameter does not converge in the zero drift case.

Then, in Chapter 5, we turn to a different method, using martingale differences, in order to prove the central limit theorem for the diameter in the case with non-zero drift. As with the similar result for the perimeter length of the convex hull which was established in [WX15a], these results do not hold for a certain class of walks, the time-space processes. Thus, in Chapter 6, we fill in the gap for the diameter, establishing the limiting distribution for the diameter in this case. For the perimeter length, we do not get the limiting distribution, but show the variance grows slower than any polynomial in n, the number of steps. The heuristic and motivation behind the proofs in this case leads to some further results regarding the faces of the time-space processes which are also presented in Chapter 6.

1.3 Mathematical prerequisites

Some of the key mathematical themes that we will need to cover before embarking on our original work include measures, metrics and convexity. We also include some standard probability theory for ease of reference. For a comprehensive introduction to metric spaces see [Bar95], a nice text on convexity is [Gru07] and for a more detailed exposition of the probability theory see [Gut05].

We begin with the definition of convexity. If we are considering a set $A \subseteq \mathbb{R}^d$ then we say A is convex if for all $x, y \in A$, $x + \lambda(y - x) \in A$ for all $\lambda \in [0, 1]$.

A key definition for us will be the convex hull of a set, intuitively, this is the set of all points that are between points in the set. More formally hull A is the smallest convex set containing A. Note this definition means the convex hull of a convex set is the set itself.

Two final pieces of notation for specific sets that we will use are $\mathbb{S}^{d-1} := \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\| = 1\}$ for the unit sphere in $\mathbb{R}^d$ and for the unit ball in $\mathbb{R}^d$ we write $\mathbb{B}^d := \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\| \leq 1\}$.

As well as notation of sets, we will require the concept of algebras and σ-algebras, for definitions see for example [Bar95, § 9, § 13]. A natural link between the sets and σ-algebras is the notion of generating a σ-algebra from a set of subsets, $\mathcal{E}$, which we denote $\sigma(\mathcal{E})$. The generated σ-algebra is the smallest σ-algebra which contains all the subsets in $\mathcal{E}$. Note also that an algebra itself is a set (of subsets) so we can generate a σ-algebra from an algebra. If an algebra and σ-algebra are both generated by the same finite set, then it is clear they will in fact be the same, but if two sets are not exactly the same, can we numerate how different it is in a coherent way? For this, we will of course use measures, again see [Bar95, § 9] for a definition. A couple of examples of measures we mention here for notational purposes are:

Lebesgue measure: For intervals in $\mathbb{R}$, such as the open interval (x, y) with $x, y \in \mathbb{R}$ or the closed interval $[x, y]$, the Lebesgue measure is $\mu(x, y) = \mu[x, y] = y - x$.

Probability measure: If a measure is a probability measure, then it must satisfy the additional condition $\mu(\Omega) = 1$ where Ω is the sample space. We use the standard notation of replacing the μ in this case with $\mathbb{P}$.

In particular, we call the triple $(\Omega, \mathcal{F}, \mathbb{P})$ a probability triple where Ω is the sample space and $\mathcal{F}$ is a σ-algebra of subsets of Ω.

For this next technical result, we will require one explicit piece of set notation. For sets A and B we denote the symmetric difference as $A \triangle B$ which is defined as the elements that are in one and only one of the sets A and B.

Lemma 1.3.1. *Let $\mathcal{A}$ be an algebra and $\sigma(\mathcal{A})$ the generated σ-algebra. Then for any set $A \in \sigma(\mathcal{A})$ and $\epsilon > 0$, there exists a set $A' \in \mathcal{A}$ such that $\mathbb{P}(A \triangle A') < \epsilon$.*

We also need some notation for metric spaces, for some elementary definitions see for example [Rud76, p. 30]. We will use d to denote a metric and call the pair $(\mathcal{S}, d)$ a metric space where $\mathcal{S}$ is the underlying set.

A set E in the metric space $(\mathcal{S}, d)$ is called *open* if, for any $x_1 \in E$, we can find some $\epsilon > 0$, such that $d(x_1, x_2) < \varepsilon$ implies $x_2 \in E$. Recall, the complement of a set is denoted $E^c := \mathcal{S} \setminus E$ and we call E *closed* if E^c is open. We denote the closure of E by $\operatorname{cl} E$, and define the boundary of E by $\partial E := \operatorname{cl} E \cap \operatorname{cl} E^c$. The interior of $E \subseteq \mathbb{R}^d$ is $\operatorname{int} E := E \setminus \partial E$. We also use the notation E^ε to represent the set of points at a distance of at most ε from E, so $E^\varepsilon := \{x \in \mathcal{S} : d(x, E) \leq \varepsilon\}$. Often, we will use this notation without explicitly declaring that we will take the metric d to be the Euclidean metric on $\mathbb{R}^d$, defined below.

With the definition of open sets above, we can define a *compact* set as one for which any open cover, that is a cover formed from a collection of open subsets, has a finite subcover. Specifically, if $E_1, E_2, \ldots$ are open sets such that $\mathcal{S} \subseteq \cup_{i=1}^{\infty} E_i$ then there exists some finite subset $\{E_{i_1}, E_{i_2}, \ldots, E_{i_n}\} \subset \{E_1, E_2, \ldots\}$ such that $\mathcal{S} \subseteq \cup_{k=1}^{n} E_{i_k}$.

If a generating set of a σ-algebra is all of the open sets in $\mathbb{R}$, then we call the generated σ-algebra the *Borel* σ-algebra, denoted $\mathcal{B}$ and members of $\mathcal{B}$ are called Borel sets. We write $\mathcal{B}_d$ for the Borel σ-algebra on $\mathbb{R}^d$.

Particular examples of metric spaces that we will use are now described and we introduce some specific notation to represent each metric.

Euclidean metric on $\mathbb{R}$: The absolute value of the difference between two numbers. For $x, y \in \mathbb{R}$ we denote this $d(x, y) = \rho(x, y) := |x - y|$, with $|x|$ denoting the absolute value of a number.

Euclidean metric on $\mathbb{R}^n$: For the higher dimensional space we use vector distance. For $\mathbf{x} = (x_1, x_2, \dots, x_d)^\top \in \mathbb{R}^d$ we denote the Euclidean norm as $\|x\| := \sqrt{x_1^2 + \ldots + x_d^2}$. Then for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ we define the Euclidean distance by $d(\mathbf{x}, \mathbf{y}) = \rho_E(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|$.

Euclidean distance between a point and a set, or two sets: this is the minimum distance between a point, x, and the boundary of a set, A, defined as $d(x, A) = \rho_E(x, A) := \inf_{y \in A} d(x, y)$. Note that this is not a metric in itself because we do not have a definition for the distance between any two elements of the space if the space contains sets and points, because our definition does not admit taking two sets. We could take the Euclidean distance between two sets; for two sets A and B we set $\rho_E(A, B) := \inf_{x \in A} \inf_{y \in B} \rho(x, y) = \inf_{x \in A} \rho_E(x, B) = \inf_{y \in B} \rho_E(y, A)$. However, this is not a metric either because the distance is 0 if the sets have a common element, but having a common element does not mean $A = B$. Nevertheless, these Euclidean distances are useful to have defined. For a metric for such a space, we will use the Hausdorff metric, see [Gru07, p. 84] for further details.

Hausdorff metric on $\mathfrak{S}_0^d$: we use the notation $\mathfrak{S}_0^d$ to denote the set of bounded subsets of $\mathbb{R}^d$ containing $\mathbf{0}$. For $A, B \in \mathfrak{S}_0^d$ we define the Hausdorff metric by either of the following equivalent definitions

$$d(A, B) = \rho_H(A, B) := \max\left\{\sup_{x \in A} \rho_E(x, B), \sup_{y \in B} \rho_E(y, A)\right\}, \tag{1.3.1}$$

$$d(A, B) = \rho_H(A, B) := \inf\{\varepsilon \geq 0 : A \subseteq B^\varepsilon \text{ and } B \subseteq A^\varepsilon\}. \tag{1.3.2}$$

We note here that when discussing vectors in $\mathbb{R}^d$ we will assume all vectors are column vectors. Also, it will often be convenient to normalise a vector in $\mathbb{R}^d \setminus \{\mathbf{0}\}$ so that it has unit length. For this we write $\hat{\mathbf{x}} := \mathbf{x}/\|\mathbf{x}\|$. We then use the convention $\hat{\mathbf{0}} = \mathbf{0}$.

Some further metric spaces we wish to consider concern spaces of functions. In particular, we will restrict ourselves to *measurable functions*. For measurable spaces $(S, \mathcal{F})$ and $(S', \mathcal{F}')$, a function, $f : S \mapsto S'$ is *measurable* if, for any $B \in \mathcal{F}'$, $f^{-1}(B) \in \mathcal{F}$. We note that this definition does not in fact require a measure to be defined.

However, we will be using some specific metric spaces so we will use this definition to describe our sets. First, call the set of bounded, measurable[2] $f : [0, 1] \mapsto \mathbb{R}^d$ the

[2]where the σ-algebras are taken to be the Borel ones

set of *trajectories* and denote them $\mathcal{M}^d := \mathcal{M}^d[0,1]$ where here and throughout we use the notation $f[0,t]$ to denote the *interval image* for $t \in [0,1]$, formally $f[0,t] := \{f(x) : x \in [0,t]\}$. Then for $f \in \mathcal{M}^d$ we write $D_f \subset [0,1]$ for the set of discontinuities of f, that is $D_f := \{c \in [0,1] : \lim_{x \to c} \neq f(c)\}$. Then, we call the set of continuous functions $\mathcal{C}^d := \mathcal{C}^d[0,1]$ defined as $\mathcal{C}^d := \{f \in \mathcal{M}^d : D_f = \emptyset\}$. Finally, we denote the set of right-continuous functions with left hand limits (often called cádlág functions) as $\mathcal{D}^d := \mathcal{D}^d[0,1]$. These are the functions $f \in \mathcal{M}^d$ such that

1. For $0 \leq t < 1, f(t+) = \lim_{s \downarrow t} f(s)$ exists and $f(t+) = f(t)$.

2. For $0 < t \leq 1, f(t-) = \lim_{s \uparrow t} f(s)$ exists.

Note that functions in $\mathcal{D}^d$ are bounded, and have (at most) countably many discontinuities of the first type (jump discontinuities): see [Bil99, pp. 121–122]. For any of these sets, we often add the restriction $f(0) = \mathbf{0}$, and call the induced subsets $\mathcal{M}_0^d := \{f \in \mathcal{M}^d : f(0) = \mathbf{0}\}$, $\mathcal{C}_0^d := \{f \in \mathcal{C}^d : f(0) = \mathbf{0}\}$ or $\mathcal{D}_0^d := \{f \in \mathcal{D}^d : f(0) = \mathbf{0}\}$. For ease of notation, we will also use $\mathcal{M}$, $\mathcal{C}$ and $\mathcal{D}$ when considering $\mathcal{M}^1$, $\mathcal{C}^1$ and $\mathcal{D}^1$.

Possible metric spaces to consider are the following.

Supremum metric on $\mathcal{M}^d$, $\mathcal{C}^d$ or $\mathcal{D}^d$: For $f \in \mathcal{M}^d$ define the supremum norm of f as $\|f\|_\infty := \sup_{0 \leq t \leq 1} \|f(t)\|$. Then the supremum metric between two functions $f, g \in \mathcal{M}^d$ is defined as

$$\rho_\infty(f,g) := \|f - g\|_\infty = \sup_{0 \leq t \leq 1} \|f(t) - g(t)\|. \tag{1.3.3}$$

Since $\mathcal{C}^d$ and $\mathcal{D}^d$ are subsets of $\mathcal{M}^d$, this definition extends naturally to the spaces $(\mathcal{C}^d, \rho_\infty)$ and $(\mathcal{D}^d, \rho_\infty)$.

Skorokhod metric on $\mathcal{M}^d$, $\mathcal{C}^d$ and $\mathcal{D}^d$: Let Λ denote the class of strictly increasing, continuous mappings of $[0,1]$ onto itself. Note, if $\lambda \in \Lambda$, then $\lambda(0) = 0$, $\lambda(1) = 1$ and $\lambda^{-1} \in \Lambda$. Then for functions $f, g \in \mathcal{M}^d$, define the Skorokhod metric as

$$\rho_S(f,g) := \inf_{\lambda \in \Lambda} \{\|\lambda - I\|_\infty \vee \|f - g \circ \lambda\|_\infty\} \tag{1.3.4}$$

where I is the identity map on $[0,1]$. Again, this extends naturally to $(\mathcal{C}^d, \rho_S)$ and $(\mathcal{D}^d, \rho_S)$. See [Pol84, p. 123] for further details.

Kolmogorov-Billingsley metric on $\mathcal{M}^d$, $\mathcal{C}^d$ and $\mathcal{D}^d$: For $\lambda \in \Lambda$, as described above, let

$$\|\lambda\|^\circ := \sup_{s<t} \left| \log \frac{\lambda(t) - \lambda(s)}{t - s} \right|.$$

Then for $f, g \in \mathcal{M}^d$, we define the Kolmogorov-Billingsley metric, see [Bil99; Kol56], as $\rho_S^\circ(f,g) := \inf_{\lambda \in \Lambda}\{\|\lambda\|^\circ \vee \|f - g \circ \lambda\|_\infty\}$.

The latter two metrics can be considered as accounting for small perturbations in time as well as space when considering distance between functions. For further discussion and a motivating example, see Section A.1.1.

Note the following simple fact which we will use later.

Lemma 1.3.2. *For any $f, g \in \mathcal{M}^d$ we have $\rho_S(f,g) \le \rho_\infty(f,g)$.*

Proof. The infimum in (1.3.4) is bounded above by the value at $\lambda = I$. □

Of course, we won't just be considering deterministic functions. We will consider random variables, measurable functions from the probability triple $(\Omega, \mathcal{F}, \mathbb{P})$ to a measurable space $(S, \mathcal{S})$ (sometimes also a metric space), where we use the standard notation that for $A \subseteq \mathcal{S}$, $\mathbb{P}(Z \in A) := \mathbb{P}\{\omega \in \Omega : Z(\omega) \in A\}$. We also use the standard notation $\mathbb{E}(Z)$ to represent the expectation of the real-valued random variable Z, but will omit the brackets if no ambiguity ensues. If $Z \in \mathbb{R}^d$, the expectation will be taken to be component-wise. Also, we refer to the property that, for two random variables Z_1 and Z_2 in $\mathbb{R}^d$, we have $\mathbb{E}[Z_1 + Z_2] = \mathbb{E}\, Z_1 + \mathbb{E}\, Z_2$, as the linearity of expectation. Further related theory can be found at, for example, [Dur10, §1].

Using this definition, on $\mathbb{R}$ we specify the case where $g(Z) = (Z - \mathbb{E}\, Z)^2$ as the variance, that is $\mathbb{V}\text{ar}[Z] := \mathbb{E}[(Z - \mathbb{E}\, Z)^2]$. When Z is a random variable in $\mathbb{R}^d$ we denote the covariance matrix of Z by Σ which is defined as $\Sigma := (Z - \mathbb{E}\, Z)(Z - \mathbb{E}\, Z)^\top$ where $\mathbf{z}^\top$ is used to denote the transpose of the vector $\mathbf{z}$. Thus, Σ is a d by d covariance matrix, which we will say is positive definite if all of its eigenvalues are positive - heuristically a positive definite covariance matrix implies that the walk does not live on lower dimensional subspace of the whole space. Being positive definite is good, because then Σ has a unique nonnegative-definite symmetric square-root $\Sigma^{1/2}$ satisfying $(\Sigma^{1/2})^2 = \Sigma$.

On the other hand, we use lower case σ^2 to represent $\mathbb{E}(\|Z - \mathbb{E}\, Z\|^2)$ which in fact gives $\sigma^2 = \operatorname{tr}\Sigma$. We often use the notation $\mu := \mathbb{E}\, Z$ and when $\|\mu\| > 0$, we denote $\hat{\mu} := \mathbb{E}\, Z \cdot \|\, \mathbb{E}\, Z\|^{-1}$, the unit vector in the direction of the expectation. Using this notation we split the variance into the direction of the mean and the direction perpendicular to the mean. Thus, denote

$$\sigma_{\mu}^2 := \mathbb{E}[((Z - \mu) \cdot \hat{\mu})^2], \text{ and } \sigma_{\mu_\perp}^2 := \sigma^2 - \sigma_{\mu}^2. \tag{1.3.5}$$

An important example of a random variable is the Normal distribution both in one dimension and multiple dimensions.

Normal random variable: We write $Z \sim \mathcal{N}(\mu, \sigma^2)$ with $\mu \in \mathbb{R}$ and $\sigma^2 > 0$, if Z is a random variable on $\mathbb{R}$ with probability density function $f(x) = (2\pi\sigma^2)^{-1/2} e^{-(x-\mu)^2/2\sigma^2}$. Here $\mathbb{E}\, Z = \mu$ and $\mathbb{V}\mathrm{ar}\, Z = \sigma^2$. We use ζ to denote the *standard Normal*, where $\mu = 0$ and $\sigma^2 = 1$.

Multivariate Normal random variable: The multivariate normal distribution in d dimensions is written $Z \sim \mathcal{N}_d(\mu, \Sigma)$, where Σ is the covariance matrix and μ the expectation, now a d-dimensional vector. When the determinant of the covariance matrix $\det(\Sigma) > 0$, we define the Normal distribution by the density function $f(\mathbf{x}) = (2\pi)^{-d/2} \det(\Sigma)^{-1/2} e^{-(\mathbf{x}-\mu)^\top \Sigma^{-1} (\mathbf{x}-\mu)/2}$, where $\mathbf{x}$ is also a d-dimensional vector. The standard d-dimensional Normal random variable has covariance matrix I_d, the d-dimensional identity matrix, and mean vector $\mathbf{0}$. In the degenerate case where Σ is a $d \times d$ square matrix of zeros, we define the multivariate Normal distribution by the point mass at μ, that is $\mathbb{P}(Z = \mu) = 1$.

Establishing the expectation and variance from the distribution function is standard, but we are able to use such quantities to obtain information about the tails of the distribution through the rightly celebrated Markov and Chebyshev inequalities, see [Gut05, p.120,p.121].

Theorem 1.3.3. *Let $r > 0$ and take $a > 0$, then $\mathbb{P}(|Z| > a) \le \mathbb{E}[|Z|^r]/a^r$.*

Theorem 1.3.4. *Let $a > 0$, then $\mathbb{P}(|Z - \mathbb{E}\, Z| > a) \le \mathbb{V}\mathrm{ar}[Z]/a^2$.*

One other celebrated inequality which will be particularly useful when trying to get a handle on the expectation of the product of two random variables is the Cauchy-Schwarz inequality [Gut05, p.130].

Theorem 1.3.5. *If* $\mathbb{E}\,|Z_1|^2 < \infty$ *and* $\mathbb{E}\,|Z_2|^2 < \infty$*, then* $|\mathbb{E}\,Z_1 Z_2| \leq \mathbb{E}\,|Z_1 Z_2| \leq \sqrt{\mathbb{E}\,Z_1^2\,\mathbb{E}\,Z_2^2}$.

In order to study convex hulls we will need to consider not only statistics of a single random variable Z but a sequence of random variables $Z_1, Z_2, \ldots$. If considering the long term behaviour of such a sequence, we will also need a notion of limits and then a way to compare a random sequence in the limit. Let

$$\limsup_{n\to\infty} x_n := \lim_{n\to\infty} \sup_{m\geq n} x_m,$$

which exists in $(-\infty, \infty]$ by monotonicity. Note that, if $\limsup_{n\to\infty} x_n = c$ then for any $\varepsilon > 0$, $x_n < c + \varepsilon$ all but finitely often and $x_n > c - \varepsilon$ infinitely often so this is the smallest upper bound for the sequence. Likewise,

$$\liminf_{n\to\infty} x_n := \lim_{n\to\infty} \inf_{m\geq n} x_m.$$

Similarly, if $\liminf_{n\to\infty} x_n = c$ then for any $\varepsilon > 0$, $x_n > c - \varepsilon$ all but finitely often and $x_n < c + \varepsilon$ infinitely often so this is the largest lower bound for the sequence. Then the limit is simply defined as

$$\lim_{n\to\infty} x_n = \begin{cases} c & \text{if } \liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n = c; \\ \text{does not exist} & \text{if } \liminf_{n\to\infty} x_n \neq \limsup_{n\to\infty} x_n. \end{cases}$$

Of course, this is a convenient definition specific to sequences of real numbers where the lim sup considers an upper bound, and lim inf a lower bound, and both of these are functions that will be useful in themselves. If we wish to consider a sequence of numbers in say $\mathbb{R}^2$ then we cannot simply consider two bounds, so we do not define lim sup and lim inf in this case and define the limit as $\lim_{n\to\infty} \mathbf{x}_n = \mathbf{x}$ with $\mathbf{x} \in \mathbb{R}^2$, if for any $\epsilon > 0$ there exists N such that $\|\mathbf{x}_n - \mathbf{x}\| < \epsilon$ for all $n \geq N$.

Hence, we can now describe several well known types of convergence for random variables.

Convergence almost surely: The sequence $Z_1, Z_2, \ldots$ converges almost surely to Z, write $Z_n \xrightarrow{\text{a.s.}} Z$ if $\mathbb{P}\{\omega : \lim_{n\to\infty} Z_n(\omega) = Z(\omega)\} = 1$.

Remark 1.3.6. Sometimes we will write $\lim_{n\to\infty} Z_n = Z$ a.s. to represent almost sure convergence. Likewise, statements such as $\limsup_{n\to\infty} Z_n = Z$ a.s. mean $\mathbb{P}(\limsup_{n\to\infty} Z_n = Z) = 1$.

Convergence in probability: The sequence $Z_1, Z_2, \ldots$ converges in probability to Z, write $Z_n \xrightarrow{\text{P}} Z$ if for any $\varepsilon > 0$, $\lim_{n\to\infty} \mathbb{P}(|Z_n - Z| > \epsilon) = 0$.

Convergence in L^r: The sequence $Z_1, Z_2, \ldots$ converges in L^r to Z, write $Z_n \to Z$ in L^r, if $\lim_{n\to\infty} \mathbb{E}[|Z_n - Z|^r] = 0$.

Convergence in distribution: The sequence $Z_1, Z_2, \ldots$ converges in distribution to Z, write $Z_n \xrightarrow{\text{d}} Z$ if $\lim_{n\to\infty} \mathbb{P}(Z_n \le c) = \mathbb{P}(Z \le c)$, for all c at which $\mathbb{P}(Z \le c)$ is continuous.

Of course, this definition requires the random variables to have the domain $\mathbb{R}$ for the less-than operator to make sense. One generalisation of convergence in distribution is *weak convergence* which we start by defining for probability measures.

Weak Convergence: The probability measures $P_1, P_2, \ldots$ defined on a metric measure space $(S, \mathcal{S}, \rho)$ converge weakly to P, that is, $P_n \Rightarrow P$, if

$$\int_S f \mathrm{d}P_n \to \int_S f \mathrm{d}P$$

for all bounded, continuous $f : S \to \mathbb{R}$.

As with the other types of convergence, it is often more convenient to speak of weak convergence of random variables. Consider a random element X on $(\Omega, \mathcal{F}, \mathbb{P})$, taking values in a metric measure space $(S, \mathcal{S}, \rho)$. Consider also a sequence of random variables X_n, defined on possibly different probability spaces $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$, but all taking values in the same metric measure space $(S, \mathcal{S}, \rho)$. We associate with $X, X_1, X_2, \ldots$ probability measures $P, P_1, P_2, \ldots$ on $(S, \mathcal{S}, \rho)$ in the natural way: for any $B \in \mathcal{S}$,

$$P(B) = \mathbb{P}(X \in B), \text{ and } P_n(B) = \mathbb{P}_n(X_n \in B). \tag{1.3.6}$$

Definition 1.3.7. In this context, we say that $X_n \Rightarrow X$ if $P_n \Rightarrow P$.

In other words, $X_n \Rightarrow X$ if $\lim_{n\to\infty} \mathbb{E}_n f(X_n) = \mathbb{E} f(X)$ for all bounded, uniformly continuous $f : S \to \mathbb{R}$, where $\mathbb{E}$ and $\mathbb{E}_n$ are expectations under $\mathbb{P}$ and $\mathbb{P}_n$, respectively.

Remark 1.3.8. In the case where $(S, \mathcal{S}, \rho)$ is $(\mathbb{R}^d, \mathcal{B}_d, \rho_E)$, where $\mathcal{B}_d$ is the Borel σ-algebra of $\mathbb{R}^d$, weak convergence reduces to convergence in distribution: see [Kal02, p. 42].

It is well-known that these convergences are linked in the sense that almost sure convergence implies convergence in probability which in turn implies convergence in distribution. Further, convergence in L^r also implies convergence in probability. Hence, we, where appropriate, will state results as convergence almost surely and in L^r, with the other convergence results implicit.

However, there are further conditions upon which further implications of convergence can be satisfied. One such is Lebesgue's dominated convergence theorem, see [Gut05, p.57], which allows us to pass from almost sure convergence to convergence in L^1.

Theorem 1.3.9. *Suppose that $|Z_n| < Y$ for all n, for some Y with $\mathbb{E}\,|Y| < \infty$. If $Z_n \xrightarrow{\text{a.s.}} Z$ as $n \to \infty$, then $Z_n \to Z$ in L^1 as $n \to \infty$.*

This statement can be generalised slightly, so that we dominate the Z_n by a sequence of random variables, not just one fixed Y. This is Pratt's lemma [Gut05, p.221].

Theorem 1.3.10. *Suppose that $|Z_n| < Y_n$ for all n, for some Y_n such that $Y_n \xrightarrow{\text{a.s.}} Y$ as $n \to \infty$ with $\lim_{n\to\infty} \mathbb{E}\,Y_n = \mathbb{E}\,Y$ as $n \to \infty$ where $\mathbb{E}\,Y \in (-\infty, \infty)$. If $Z_n \xrightarrow{\text{a.s.}} Z$ as $n \to \infty$, then $Z_n \to Z$ in L^1 as $n \to \infty$.*

Another link can be created by assuming uniform integrability of the sequence of random variables. A sequence $Z_1, Z_2, \dots$ of random variables is uniformly integrable if, for any $\epsilon > 0$, there exists $C_\epsilon \in [0, \infty)$ such that $\mathbb{E}[Z_n \mathbf{1}\{|Z_n| > C_\epsilon\}] < \epsilon$ for all n, where here and elsewhere we use $\mathbf{1}\{A\}$ to be the indicator function of the event A. The uniformity is in the sense that C_ϵ can be chosen independent of n. The following can be found at [Gut05, p.224].

Theorem 1.3.11. *Suppose that $Z_n \xrightarrow{\text{d}} Z$ as $n \to \infty$ and for some $r > 0$, $\{Z_1^r, Z_2^r, \dots\}$ are uniformly integrable, then*

$$\mathbb{E}\,|Z_n|^r \to \mathbb{E}\,|Z|^r \quad \text{as } n \to \infty.$$

A final convergence result that we will make use of is Slutsky's theorem, see [Gut05, p.249]. This does not connect different types of convergence but, the part of the theorem we state and use, allows us to consider sums of two random variables and carry the sum across the limit as follows.

Theorem 1.3.12. *Let* $Z_1, Z_2, \ldots$ *and* $Y_1, Y_2, \ldots$ *be sequences of random variables such that* $Z_n \xrightarrow{\text{d}} Z$ *as* $n \to \infty$ *and* $Y_n \xrightarrow{\text{P}} c$ *as* $n \to \infty$ *for some constant* c. *Then* $Z_n + Y_n \xrightarrow{\text{d}} Z + c$ *as* $n \to \infty$.

Here we also note Slutsky's result (see e.g. [Bil99, Theorem 3.1]) stated in the context of weak convergence.

Theorem 1.3.13. *Let* $Z, Z_1, Z_2, \ldots$ *and* $Y_1, Y_2, \ldots$ *be sequences of random variables on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *taking values in metric measure space* $(S, \mathcal{S}, \rho)$. *If* $X_n \Rightarrow X$ *and* $\rho(X_n, Y_n) \xrightarrow{\text{P}} 0$, *then* $Y_n \Rightarrow X$.

Now we have defined the convex hull, and looked at sequences of random variables, but we are yet to discuss the random walks that are key to the study. A random walk is simply the set of partial sums of a sequence of random variables, $Z_1, Z_2, \ldots$, often assumed to be independent and identically distributed, which in this context are called increments. We give the following formal, labelled definitions which we will refer to throughout.

($\mathbf{W}_\mu$) Let $d \in \mathbb{N}$, and suppose that $Z, Z_1, Z_2, \ldots$ are i.i.d. random vectors in $\mathbb{R}^d$ with $\mathbb{E}\,\|Z\| < \infty$ and $\mathbb{E}\, Z = \mu$. The random walk $(S_n, n \in \mathbb{Z}_+)$ is the sequence of partial sums $S_n := \sum_{i=1}^n Z_i$ with $S_0 := \mathbf{0}$.

For large parts of this work, our results concern the case when $d = 2$. Here, we will use the more compact notation ($\mathbf{W}_\mu^2$) instead of explicitly stating $d = 2$ at each occurrence. Independence itself could warrant a whole chapter of discussion, see [Dur10, ch.2]. However, we simply state the condition for identically distributed Z_i such that $Z : \Omega \mapsto \mathbb{R}^d$ as

$$\mathbb{P}(Z_1 \in E_1, Z_2 \in E_2, \ldots, Z_n \in E_n) = \mathbb{P}(Z_1 \in E_1)\mathbb{P}(Z_2 \in E_2) \cdots \mathbb{P}(Z_n \in E_n),$$

for any $E_1, \dots, E_n$ Borel subsets of $\mathbb{R}^d$.

There are several classical results of random walks which we use throughout. First the strong law of large numbers due to Kolmogorov [Kol30] which states the average step of the walk converges almost surely to the expected increment.

Theorem 1.3.14. *Consider the random walk as defined at* $(\mathbf{W}_\mu)$*, then* $n^{-1}S_n \xrightarrow{\text{a.s.}} \mu$.

In order to establish many more interesting results, we often require further conditions such as finite variance. We will often impose this or stronger restrictions through use of one of the following two conditions:

(V) Suppose that $\mathbb{E}[\|Z\|^2] < \infty$ and write $\Sigma := \mathbb{E}[(Z-\mu)(Z-\mu)^\top]$. Here Σ is a nonnegative-definite, symmetric d by d matrix; we write $\sigma^2 := \operatorname{tr}\Sigma = \mathbb{E}[\|Z-\mu\|^2]$,

($\mathbf{M}_p$) Suppose that $\mathbb{E}[\|Z\|^p] < \infty$.

With the condition $(\mathbf{V})$, the central limit theorem gives an expression of the size of the error as the average step converges.

Theorem 1.3.15. *Consider the random walk as defined at* $(\mathbf{W}_\mu)$ *with* $(\mathbf{V})$*, then* $n^{-1/2}(S_n - n\mu) \xrightarrow{\text{d}} \zeta$*, where* $\zeta \sim \mathcal{N}(\mathbf{0}, \Sigma)$ *is a d-dimensional Normal random variable.*

This gives us an idea of how the errors are distributed, but sometimes we may care about particularly large/small events away from the mean. In 1-dimension we have the Hartman-Wintner law of the iterated logarithm [HW41] which gives us the order of the lim sup.

Theorem 1.3.16. *Consider the random walk as defined at* $(\mathbf{W}_\mu)$ *with* $d = 1$*, and with* $(\mathbf{V})$*. Then*

$$\limsup_{n\to\infty} \frac{S_n - n\mu}{\sqrt{2\sigma^2 \log\log n}} = 1 \ \text{a.s.}$$

Considering the walk with increments $-Z$ of course shows that the result where the lim sup and 1 are replaced by lim inf and -1 respectively also holds.

It is not always enough to know information about the limit of a random walk, often we require knowledge about the process up to some fixed time n. To this end there are three relevant inequalities we will use, all defined under the assumption that $Z, Z_1, Z_2, \ldots \in \mathbb{R}$. The Azuma-Hoeffding inequality [Gut05, p.120] on tail probabilities of S_n, the Marcinkiewicz-Zygmund inequality [Gut05, p.150] on the expectation of a power of S_n, and Etemadi's inequality [Gut05, p.144] on the maximum of the walk up to time n.

Theorem 1.3.17. *Let $c_1, c_2, \ldots$ be finite positive constants. If $\mathbb{E}\, Z_i = 0$ and $|Z_i| < c_i$ for all for $i = 1, 2, \ldots,$ and $S_0 = 0$. Then for any $a > 0$ and $n \geq 1$,*

$$\mathbb{P}(S_n \geq a) \leq \exp\left(\frac{-a^2}{2\sum_{i=1}^{n} c_i^2}\right).$$

Theorem 1.3.18. *If $\mathbb{E}\, Z = 0$, and $\mathbb{E}\, |Z|^p < \infty$ for some $p \geq 1$, then*

$$C_1\, \mathbb{E}\left((\sum_{i=1}^{n} |Z_i|^2)^{p/2}\right) \leq \mathbb{E}(|S_n|^p) \leq C_2\, \mathbb{E}\left((\sum_{i=1}^{n} |Z_i|^2)^{p/2}\right),$$

for two constants C_1 and C_2 which only depend on p.

Theorem 1.3.19. *Let $a > 0$, then $\mathbb{P}(\max_{1 \leq k \leq n} |S_k| > 3a) \leq 3 \max_{1 \leq k \leq n} \mathbb{P}(|S_k| > a)$.*

We will require some further notation regarding random walks, some of which has already been described in the introduction.

For the convex hull of an n-step random walk we use $\mathcal{H}_n := \text{hull}(S_0, S_1, \ldots, S_n)$. Write L_n for the perimeter length of $\mathcal{H}_n$, and let

$$D_n := \text{diam}\{S_0, S_1, \ldots, S_n\} = \max_{0 \leq i,j \leq n} \|S_i - S_j\| = \text{diam}\, \mathcal{H}_n. \tag{1.3.7}$$

On occasion it will be useful to use $\mathcal{L}(E)$ or $\mathcal{D}(E)$ as the respective functionals for the perimeter and diameter of the set E when E is not the random walk.

Before moving on from our prerequisites we must mention one specific process related to random walks, Brownian motion. This is not defined in the same way as the previously discussed random walks which were sums of random variables, however is defined by three properties, see [Dur10, ch.8]. We call $b_d = (b_d(t), t \in \mathbb{R}_+)$, a standard d-dimensional Brownian motion process if:

i. $b_d(0) = \mathbf{0}$;

ii. for $t_0 < t_1 < \ldots < t_n$, $b_d(t_0), b_d(t_1) - b_d(t_0), \ldots, b_d(t_n) - b_d(t_{n-1})$ are independent;

iii. the jump $b_d(s)-b_d(t)$, with $s>t$ is a random variable distributed as a multivariate Normal, $\mathcal{N}_d(\mathbf{0},(s-t)I_d)$ where I_d is the d-dimensional identity matrix.

To generalise this process, we say $\Sigma^{1/2}b_d$ is *correlated* Brownian motion with covariance matrix Σ. In the case $d=1$ we write simply b for b_1.

The particular relevance of this result to random walk theory comes from a theorem of Donsker (see e.g. [Dur10, p.386] for the $d=1$ case), which states that the path which connects the points of a random walk by straight lines, where $\mathbb{E}\, Z=\mathbf{0}$, behaves like Brownian motion when rescaled by $n^{-1/2}$. We only state this loosely here, with the rigorous formulation saved for Theorem 3.1.5.

Finally, we note a couple of extra bits of notation. We use $\mathbf{1}_A(x)$ to denote the indicator function of a set, that is

$$\mathbf{1}_A(x)=\begin{cases}1 \text{ if } x\in A;\\ 0 \text{ otherwise.}\end{cases}$$

Then, for $x\in\mathbb{R}$ we set $x^+:=\max\{x,0\}$, $x^-:=\max\{-x,0\}$, so that $x=x^+-x^-$ and

$$\operatorname{sgn}(x)=\begin{cases}1 \text{ if } x>0;\\ 0 \text{ if } x=0;\\ -1 \text{ if } x<0.\end{cases}$$

For $x,y\in\mathbb{R}$ we write $x\wedge y:=\min\{x,y\}$ and $x\vee y:=\max\{x,y\}$. Given functions f and g, the function $f\circ g$ is defined by $(f\circ g)(x)=f(g(x))$. The projection of $A\subseteq\mathbb{R}^d$ on to the space perpendicular to $u\in\mathbb{S}^{d-1}$ will be denoted by $A|u^\perp$. For $d=2$, we use the notation $\mathbf{e}_\theta=(\cos\theta,\sin\theta)$ for the unit vector in direction θ.

1.4 Examples and simulation comments

Throughout, we will demonstrate our results by considering some examples of random walks and running various simulations. We will use the notation described above at $(\mathbf{W}_\mu)$ and describe our choices of Z here. The first walk we will consider is the *simple symmetric random walk* in d-dimensions. For this walk, we have

$$Z\in\{\pm\mathbf{e}_1,\pm\mathbf{e}_2,\dots,\pm\mathbf{e}_d\},$$

each with probability $(2d)^{-1}$, where the $\mathbf{e}_i$ are the standard basis vectors of $\mathbb{R}^d$.

The second walk, we will call the d-dimensional *standard Normal random walk*. This walk lives in $\mathbb{R}^d$, and simply takes $Z \sim \mathcal{N}_d(\mathbf{0}, I_d)$, the standard d-dimensional Normal random variable, also formally described above. In both of these first two cases $\mu = \mathbf{0}$.

We will also consider some walks with drift. All of our examples set the first coordinate, which will be the horizontal axis in plots, as the direction of the drift. We note that a simple transformation of the coordinate space could map these walks to a whole class of walks with drift in any direction and it is only our affinity to doing things from left-to-right that motivates this choice. The first walk in this case will be the *random walk with drift and all coordinates Normally distributed*, for which we will consider increments as

$$Z = (\zeta_{\tilde{\mu},\tilde{\sigma}}, \zeta_1, \dots, \zeta_{d-1}),$$

where $\zeta_{\tilde{\mu},\tilde{\sigma}} \sim \mathcal{N}(\tilde{\mu}, \tilde{\sigma}^2)$ for constants $\tilde{\mu}$ and $\tilde{\sigma}$, and $\zeta_i \sim \mathcal{N}(0,1)$ are independent Normal random variables. In our demonstrations, we will use a combination of different choices of $\tilde{\mu}$ and $\tilde{\sigma}$ where it is necessary to show how these parameters impact on the walk and its respective limit theorems. In this case $\mu = (\tilde{\mu}, 0, \dots, 0)$.

Finally we consider a random walk where there is drift, but no randomness in the direction of the drift. The *random walk with drift and no variance in the first coordinate*, will take

$$Z = (\tilde{\mu}, \zeta_1, \dots, \zeta_{d-1}),$$

where, as before, ζ_i are independent standard Normal random variables, and we will consider various values of $\tilde{\mu}$. Note, that this walk coincides with the random walk with Normal drift and standard $(d-1)$ dimensional Normal deviations where we take $\tilde{\sigma} = 0$, if we use the defined distribution, where $\zeta_i = \mu$ a.s., for the degenerate case $\mathcal{N}(\mu, 0)$.

1.4.1 Simulation comments

Writing a simulation of a random walk is a fairly trivial task if you have any programming experience, but writing a program which can simulate millions of steps with thousands of repetitions whilst not making use of a supercomputer becomes far from trivial, especially if you need to calculate the convex hull of these long walks. Whilst many of the plots in this thesis are only short simulations (for presentational purposes), the more demanding programs required some tricks to make the simulations feasible.

Firstly, we aimed to use a cluster of computer cores to run parts of the code in parallel. By splitting the code up in this way we reduce the time to simulate all the steps but 'gluing together' thousands of partial walks is not feasible because the memory required to store all of the individual steps is too great for a normal computer to process, never mind calculating the convex hull of such a walk. However, we can use what we know about convex hulls and convexity to help us.

If we calculate the convex hull of each individual subsection of the walk, and remember the start and endpoints of the walk, we only need to store the vertices of the walk and these two points. In the convex hull of an n step walk there are around $\log n$ vertices (more on this later), so the memory required is much more manageable. Moreover, after gluing together all the convex hulls of the subsections, we can calculate the convex hulls of all the vertices which produces the convex hull of the whole walk. All that remains is to balance out the length of the subsections with the number of subsections, in turn balancing out the time increased by calculating each subsection's convex hull and the time it takes to combine all the individual parts afterwards, with the memory cost of making each individual subsection too long.

As a final comment, we note that we also calculate the diameter by calculating the maximum distance between any two of the vertices in the convex hull. By convexity this will find the two points of the underlying walk which attain the diameter. Of course finding the maximum distance between any two points from a set of size $\log n$ is much faster than between points from a set of size n.

Chapter 2

Laws of large numbers and extensions using classical results

Our first exploration of the convex hull starts by considering the laws of large numbers for the perimeter length and diameter functionals in dimension 2. As discussed in the introduction, there are several results already in the literature, in particular relating to the perimeter length. Most of the results in this section can be heuristically justified by the idea that the walk with drift converges to a line segment under the law of large numbers scaling, and the walk without drift degenerates to a point under the same scaling. For now, we mention this only to explain the intuition behind the results, but this idea is more formally explored in Chapter 3.

The Spitzer and Widom formula (1.1.2) was used by Snyder and Steele to establish the law of large numbers for the perimeter length as described at (1.1.3). Their result requires the condition $\mathbb{E}(\|Z\|^2) < \infty$ and was stated for the case $\mu \neq \mathbf{0}$, but their proof works equally well when $\mu = \mathbf{0}$. Our first contribution in this section is to provide a different proof for this result which removes the need for the second moment to be finite. With a few basic observations and an application of Pratt's lemma, we can extend this to a law of large numbers for the diameter. Despite this extension being relatively simple, and could have been established from Snyder and Steele's law of large numbers for the perimeter length, albeit with stronger assumptions, it does not seem to have appeared explicitly in the literature. These two laws actually give some justification to our heuristic about the shape of the convex hull by considering the ratio of the two quantities.

In the case of drift we present some further results. First, we establish the second order term of the asymptotic expansion of $\mathbb{E} L_n$, and then use this result to recast one of the second order results of Wade and Xu [WX15a] in a stronger form. The expansion of $\mathbb{E} L_n$ can be compared with the expansions found by Grebenkov, Lanoiselée and Majumdar [GLM17], see Section 1.1.3 for details.

Finally, we provide an inequality for the same expansion for the diameter. The exact asymptotic result does not follow from the methods we employ here, and remains, as far as we know, an open problem. The second order results are known, in fact we prove them in Chapter 5; we dedicate a chapter to these results because they require a lengthier proof along the lines of the method Wade and Xu used to establish the perimeter length results.

In this section, we do not use any specific methods to obtain the results, we just make use of some classical probability theory, Cauchy's formula and some other geometrical facts[1].

2.1 Laws of large numbers

Throughout this chapter, we consider the walk with the notation as described at $(\mathbf{W}_\mu^2)$. Then our first result is the following law of large numbers for L_n.

Theorem 2.1.1. *Suppose that* $\mathbb{E}\,\|Z\| < \infty$. *Then*

$$\lim_{n\to\infty} n^{-1} L_n = 2\|\mu\|, \text{ a.s. and in } L^1.$$

On the other hand, if $\mathbb{E}\,\|Z\| = \infty$ *then* $\limsup_{n\to\infty} n^{-1} L_n = \infty$, *a.s.*

Remark 2.1.2. It is a natural question to ask whether, when $\mathbb{E}\,\|Z\| = \infty$, does it in fact hold that $\lim_{n\to\infty} n^{-1} L_n = \infty$? We note that the proof employed here does not directly answer this question, and yet neither have we found a counter example to this statement, so it remains an open problem.

[1]Based on work published in [MW18], the whole paper was joint work between the authors.

Similarly, we have a law of large numbers for D_n.

Theorem 2.1.3. *Suppose that* $\mathbb{E}\,\|Z\| < \infty$*. Then*

$$\lim_{n\to\infty} n^{-1} D_n = \|\mu\|, \text{ a.s. and in } L^1.$$

On the other hand, if $\mathbb{E}\,\|Z\| = \infty$ *then* $\limsup_{n\to\infty} n^{-1} D_n = \infty$*, a.s.*

In the case $\mu \neq \mathbf{0}$, Theorems 2.1.1 and 2.1.3 have the following immediate consequence.

Corollary 2.1.4. *Suppose that* $\mathbb{E}\,\|Z\| < \infty$ *and that* $\mu \neq \mathbf{0}$*. Then*

$$\lim_{n\to\infty} L_n / D_n = 2, \text{ a.s.}$$

Before we start on the proofs, we recall that *Cauchy's formula*, equation (1.1.1), can be stated in the following form (see e.g. equation (2.1) of [SS93]), for a finite point set $\{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_n\} \subset \mathbb{R}^2$, the perimeter length of $\operatorname{hull}\{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_n\}$ is given by

$$\int_0^{2\pi} \max_{0\le k\le n} (\mathbf{x}_k \cdot \mathbf{e}_\theta) \mathrm{d}\theta.$$

Proof of Theorem 2.1.1. Cauchy's formula applied to our random walk implies that

$$L_n = \int_0^{2\pi} \max_{0\le k\le n} (S_k \cdot \mathbf{e}_\theta) \mathrm{d}\theta. \tag{2.1.1}$$

First suppose that $\mathbb{E}\,\|Z\| < \infty$. Then the strong law of large numbers says that for any $\varepsilon > 0$ there exists N_ε with $\mathbb{P}(N_\varepsilon < \infty) = 1$ for which

$$\|S_n - n\mu\| < n\varepsilon, \text{ for all } n \ge N_\varepsilon. \tag{2.1.2}$$

Since $S_0 = \mathbf{0}$, taking $k = 0$ and $k = n$ in (2.1.1) and writing $x^+ := x\mathbf{1}\{x > 0\}$, we have

$$L_n \ge \int_0^{2\pi} (S_n \cdot \mathbf{e}_\theta)^+ \mathrm{d}\theta = 2\|S_n\|, \tag{2.1.3}$$

by Cauchy's formula for $\operatorname{hull}\{\mathbf{0}, S_n\}$. For $n \ge N_\varepsilon$ we have from (2.1.2) that

$$\|S_n\| \ge \|n\mu\| - \|S_n - n\mu\| \ge n\|\mu\| - n\varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, it follows that $\liminf_{n\to\infty} n^{-1} L_n \ge 2\|\mu\|$, a.s.

On the other hand, for any $\varepsilon > 0$, we have from (2.1.2) that

$$\begin{aligned}\max_{0\le k\le n}(S_k\cdot\mathbf{e}_\theta) &\le \max_{0\le k\le N_\varepsilon}(S_k\cdot\mathbf{e}_\theta) + \max_{N_\varepsilon\le k\le n}(S_k\cdot\mathbf{e}_\theta)\\ &\le \max_{0\le k\le N_\varepsilon}\|S_k\| + \max_{0\le k\le n}(k(\mu\cdot\mathbf{e}_\theta+\varepsilon))\\ &= \max_{0\le k\le N_\varepsilon}\|S_k\| + n(\mu\cdot\mathbf{e}_\theta+\varepsilon)^+.\end{aligned}$$

Let $A_\varepsilon := \{\theta\in[0,2\pi] : \mu\cdot\mathbf{e}_\theta > -\varepsilon\}$. Then

$$\int_0^{2\pi}(\mu\cdot\mathbf{e}_\theta+\varepsilon)^+\mathrm{d}\theta = \int_{A_\varepsilon}(\mu\cdot\mathbf{e}_\theta+\varepsilon)\mathrm{d}\theta \le \int_{A_\varepsilon}\mu\cdot\mathbf{e}_\theta\mathrm{d}\theta + 2\pi\varepsilon.$$

But

$$\begin{aligned}\int_{A_\varepsilon}\mu\cdot\mathbf{e}_\theta\mathrm{d}\theta &= \int_{A_0}\mu\cdot\mathbf{e}_\theta\mathrm{d}\theta + \int_{A_\varepsilon\setminus A_0}\mu\cdot\mathbf{e}_\theta\mathrm{d}\theta\\ &\le \int_0^{2\pi}(\mu\cdot\mathbf{e}_\theta)^+\mathrm{d}\theta + \|\mu\||A_\varepsilon\setminus A_0|.\end{aligned}$$

Hence, from (2.1.1) we obtain

$$L_n \le 2\pi\max_{0\le k\le N_\varepsilon}\|S_k\| + n\int_0^{2\pi}(\mu\cdot\mathbf{e}_\theta)^+\mathrm{d}\theta + 2\pi n\varepsilon + n\|\mu\||A_\varepsilon\setminus A_0|.$$

Since $\mathbb{P}(N_\varepsilon<\infty)=1$, it follows from Cauchy's formula for $\operatorname{hull}\{\mathbf{0},\mu\}$ that, a.s.,

$$\limsup_{n\to\infty} n^{-1}L_n \le 2\|\mu\| + 2\pi\varepsilon + \|\mu\||A_\varepsilon\setminus A_0|.$$

Since $\varepsilon>0$ was arbitrary, and $|A_\varepsilon\setminus A_0|\to 0$ as $\varepsilon\to 0$, we get $\limsup_{n\to\infty} n^{-1}L_n \le 2\|\mu\|$, a.s. Thus the almost sure convergence statement is established.

Moreover, from (2.1.1),

$$\begin{aligned}L_n &\le \int_0^{2\pi}\max_{0\le k\le n}\|S_k\|\mathrm{d}\theta\\ &\le 2\pi\max_{0\le k\le n}\sum_{j=1}^k\|Z_j\|\\ &\le 2\pi\sum_{j=1}^n\|Z_j\|.\end{aligned}$$

The strong law shows that, $n^{-1}\sum_{j=1}^n\|Z_j\| \xrightarrow{\text{a.s.}} \mathbb{E}\,\|Z\|<\infty$, while $\mathbb{E}(n^{-1}\sum_{j=1}^n\|Z_j\|) = \mathbb{E}\,\|Z\|$; hence Pratt's lemma, Theorem 1.3.10, implies that $n^{-1}L_n\to 2\|\mu\|$ in L^1.

Finally, suppose that $\mathbb{E}\,\|Z\| = \infty$. From (2.1.3), it suffices to show that

$$\limsup_{n\to\infty} n^{-1}\|S_n\| = \infty, \text{ a.s.}$$

To this end we follow [Gut05, p. 297]. First (see e.g. [Gut05, p. 75]) $\mathbb{E}\,\|Z\| = \infty$ implies that for any $c > 0$, we have $\sum_{n=1}^{\infty} \mathbb{P}(\|Z_n\| \geq cn) = \infty$, which, by the Borel–Cantelli lemma, implies that $\mathbb{P}(\|Z_n\| \geq cn \text{ i.o.}) = 1$. But $\|Z_n\| \leq \|S_n\| + \|S_{n-1}\|$, so it follows that $\mathbb{P}(\|S_n\| \geq cn/2 \text{ i.o.}) = 1$. In other words, $\limsup_{n\to\infty} n^{-1}\|S_n\| \geq c/2$, a.s., and, since $c > 0$ was arbitrary, we get the result. □

We now can use Theorem 2.1.1 to establish the law of large numbers for the diameter. In order to do so, we need the following observation about the relationship between L_n and D_n. Provided that $\mathbb{P}(Z = \mathbf{0}) < 1$, convexity implies that a.s., for all but finitely many n,

$$2 \leq L_n/D_n \leq \pi. \tag{2.1.4}$$

This result will appear at several times throughout this thesis, because the ratio itself gives some rough information about the shape of $\mathcal{H}_n$. In particular, the extrema in the inequality relate to certain shapes, specifically the line segment and shapes of constant width (such as the disc) respectively.

Proof of Theorem 2.1.3. From the definition of D_n and equation (2.1.4), we have $\|S_n\| \leq D_n \leq L_n/2$. Then we can apply the strong law for S_n, which implies that $n^{-1}\|S_n\| \to \|\mu\|$, and Theorem 2.1.1, to deduce that $n^{-1}D_n \to \|\mu\|$, a.s. Since $n^{-1}D_n \leq n^{-1}L_n/2$ we may again apply Pratt's lemma, see Theorem 1.3.10, to deduce the L^1 convergence. Finally, if $\mathbb{E}\,\|Z\| = \infty$ we use the bound $D_n \geq L_n/\pi$ from (2.1.4) and the final statement in Theorem 2.1.1 to deduce that $\limsup_{n\to\infty} n^{-1}D_n = \infty$, a.s. □

2.2 Case with drift extensions

Now we turn to the individual asymptotics for L_n and D_n in the case with non-zero drift. Recall that the behaviour of L_n was studied in [WX15a], where it was shown that if $\mathbb{E}(\|Z\|^2) < \infty$ and $\mu \neq \mathbf{0}$, then, as $n \to \infty$,

$$n^{-1/2}|L_n - \mathbb{E}\,L_n - 2(S_n - \mathbb{E}\,S_n) \cdot \hat{\mu}| \to 0, \text{ in } L^2. \tag{2.2.1}$$

We show that (2.2.1) may be recast in the following stronger form.

Theorem 2.2.1. *Suppose that* $\mathbb{E}(\|Z\|^2) < \infty$ *and* $\mu \neq \mathbf{0}$*. Then, as* $n \to \infty$*,*

$$n^{-1/2}|L_n - 2S_n \cdot \hat{\mu}| \to 0, \text{ in } L^2.$$

The following asymptotic expansion of $\mathbb{E}\, L_n$ is the key additional component in the proof of Theorem 2.2.1, and is of interest in its own right; its proof again uses the Spitzer–Widom formula (1.1.2).

Theorem 2.2.2. *Suppose that* $\mathbb{E}(\|Z\|^2) < \infty$ *and* $\mu \neq \mathbf{0}$*. Then, as* $n \to \infty$*,*

$$\mathbb{E}\, L_n = 2\|\mu\| n + \left(\frac{\sigma^2_{\mu_\perp}}{\|\mu\|} + o(1) \right) \log n.$$

Again, we can use this result, and one lemma that we describe below, to obtain an inequality for the equivalent expansion for $\mathbb{E}\, D_n$. We do not have a method for improving this result to be an asymptotic equality.

Lemma 2.2.3. *Suppose that* $\mathbb{E}(\|Z\|^2) < \infty$ *and* $\mu \neq \mathbf{0}$*. There exists* $C < \infty$ *such that*

$$0 \leq \mathbb{E}\, D_n - \|\mu\| n \leq C(1 + \log n), \text{ for all } n \geq 1.$$

We work towards a proof of Theorem 2.2.2.

We recall some notation and introduce some additional notation for these proofs. Write $X_n := S_n \cdot \hat{\mu}$ and $Y_n := S_n \cdot \hat{\mu}_\perp$, where $\hat{\mu}_\perp$ is any fixed unit vector orthogonal to μ. Then X_n and Y_n are one-dimensional random walks with increment distributions $Z \cdot \hat{\mu}$ and $Z \cdot \hat{\mu}_\perp$ respectively; note that $\mathbb{E}(Z \cdot \hat{\mu}) = \|\mu\|$, $\mathbb{E}(Z \cdot \hat{\mu}_\perp) = 0$, $\mathbb{V}\mathrm{ar}(Z \cdot \hat{\mu}) = \sigma^2_\mu$, and

$$\begin{aligned} \mathbb{V}\mathrm{ar}(Z \cdot \hat{\mu}_\perp) = \mathbb{E}[((Z - \mu) \cdot \hat{\mu}_\perp)^2] &= \mathbb{E}[\|Z - \mu\|^2] - \mathbb{E}[((Z - \mu) \cdot \hat{\mu})^2] \\ &= \sigma^2 - \sigma^2_\mu = \sigma^2_{\mu_\perp}. \end{aligned}$$

Also recall, for $x \in \mathbb{R}$ set $x^+ := x\mathbf{1}\{x > 0\}$, and also set $x^- = -x\mathbf{1}\{x < 0\}$.

Lemma 2.2.4. *Suppose that* $\mathbb{E}(\|Z\|^2) < \infty$ *and* $\mu \neq \mathbf{0}$*. Then* $\|S_n\| - |S_n \cdot \hat{\mu}|$ *is uniformly integrable.*

Proof. The central limit theorem shows that $n^{-1}Y_n^2 \xrightarrow{\mathrm{d}} \sigma^2_{\mu_\perp} \zeta^2$ where $\zeta \sim \mathcal{N}(0,1)$. Also, since $\mathbb{E}[Y_n^2] = n\sigma^2_{\mu_\perp}$, $n^{-1}\,\mathbb{E}(Y_n^2) \to \sigma^2_{\mu_\perp} = \mathbb{E}(\sigma^2_{\mu_\perp}\zeta^2)$. It is a fact that if $\theta, \theta_1, \theta_2, \ldots$

are $\mathbb{R}_+$-valued random variables with $\theta_n \xrightarrow{\mathrm{d}} \theta$, then $\mathbb{E}\,\theta_n \to \mathbb{E}\,\theta < \infty$ if and only if θ_n is uniformly integrable: see [Kal02, Lemma 4.11]. Hence we conclude that

$$n^{-1}Y_n^2 \text{ is uniformly integrable.} \tag{2.2.2}$$

Fix $\varepsilon > 0$. Let $\delta \in (0, \|\mu\|)$ to be chosen later. For ease of notation, write $T_n = \|S_n\| - |X_n|$. Then since $T_n \leq \|S_n\|$ and $|X_n| \leq \|S_n\|$, we have

$$\begin{aligned}\mathbb{E}\left[T_n \mathbf{1}\{T_n > M\}\mathbf{1}\{\|S_n\| \leq \delta n\}\right] &\leq \delta n \mathbb{P}(\|S_n\| \leq \delta n) \\ &\leq \delta n \mathbb{P}(|X_n| \leq \delta n) \\ &\leq \delta n \mathbb{P}(|X_n - \|\mu\| n| > (\|\mu\| - \delta)n).\end{aligned}$$

Since $\mathbb{E}\,X_n = n\|\mu\|$ and $\mathbb{V}\mathrm{ar}\,X_n = n\sigma_\mu^2$, Chebyshev's inequality then yields

$$\mathbb{E}\left[T_n \mathbf{1}\{T_n > M\}\mathbf{1}\{\|S_n\| \leq \delta n\}\right] \leq \delta n \frac{n\sigma_\mu^2}{(\|\mu\| - \delta)^2 n^2}.$$

It follows that, for suitable choice of δ (not depending on M) and any $M \in (0, \infty)$,

$$\sup_n \mathbb{E}\left[T_n \mathbf{1}\{T_n > M\}\mathbf{1}\{\|S_n\| \leq \delta n\}\right] \leq \varepsilon.$$

On the other hand, we use the fact that

$$0 \leq \|S_n\| - |X_n| = T_n = \frac{\|S_n\|^2 - X_n^2}{\|S_n\| + |X_n|} = \frac{Y_n^2}{\|S_n\| + |X_n|}. \tag{2.2.3}$$

Hence

$$\begin{aligned}\mathbb{E}\left[T_n \mathbf{1}\{T_n > M\}\mathbf{1}\{\|S_n\| > \delta n\}\right] &= \mathbb{E}\left[\tfrac{Y_n^2}{\|S_n\|+|X_n|}\mathbf{1}\left\{\tfrac{Y_n^2}{\|S_n\|+|X_n|} > M\right\}\mathbf{1}\{\|S_n\| > \delta n\}\right] \\ &\leq \frac{1}{\delta n}\mathbb{E}\left[Y_n^2 \mathbf{1}\{Y_n^2 > M\delta n\}\right].\end{aligned}$$

It follows that

$$\sup_n \mathbb{E}\left[T_n \mathbf{1}\{T_n > M\}\mathbf{1}\{\|S_n\| > \delta n\}\right] \leq \frac{1}{\delta}\sup_n \mathbb{E}\left[n^{-1}Y_n^2 \mathbf{1}\{n^{-1}Y_n^2 > M\delta\}\right],$$

which, for fixed δ, tends to 0 as $M \to \infty$ by (2.2.2).

Thus for any $\varepsilon > 0$ we have that $\sup_n \mathbb{E}\left[T_n \mathbf{1}\{T_n > M\}\right] \leq \varepsilon$, for all M sufficiently large, which completes the proof. □

Lemma 2.2.5. *Let $\xi, \xi_1, \xi_2, \ldots$ be i.i.d. random variables with $\mathbb{E}(\xi^2) < \infty$ and $\mathbb{E}\,\xi > 0$.*

Let $X_n = \sum_{k=1}^{n} \xi_k$*. Then* $\lim_{n\to\infty} \mathbb{E}\, X_n^- = 0$.

Proof. Let $\mathbb{E}\,\xi = m > 0$ and $\mathbb{V}\mathrm{ar}\,\xi = s^2 < \infty$. Fix $\varepsilon > 0$. Note that

$$\mathbb{E}\, X_n^- = \int_0^\infty \mathbb{P}(X_n^- > r)\mathrm{d}r = \int_0^{\varepsilon n} \mathbb{P}(X_n^- > r)\mathrm{d}r + \int_{\varepsilon n}^\infty \mathbb{P}(X_n^- > r)\mathrm{d}r.$$

Here we have that, by Chebyshev's inequality,

$$\mathbb{P}(X_n^- > r) \le \mathbb{P}(|X_n - mn| > mn + r) \le \frac{\mathbb{V}\mathrm{ar}\, X_n}{(mn+r)^2} = \frac{s^2 n}{(mn+r)^2}.$$

It follows that

$$\int_0^{\varepsilon n} \mathbb{P}(X_n^- > r)\mathrm{d}r \le s^2 n \int_0^{\varepsilon n} \frac{\mathrm{d}r}{(mn+r)^2} \le \frac{s^2\varepsilon}{m^2}. \tag{2.2.4}$$

For $B \in (0,\infty)$ let $\xi_k' := \xi_k \mathbf{1}\{|\xi_k| \le B\}$ and $\xi_k'' := \xi_k \mathbf{1}\{|\xi_k| > B\}$. Set $X_n' := \sum_{k=1}^n \xi_k'$ and $X_n'' := \sum_{k=1}^n \xi_k''$. By dominated convergence, we have that as $B \to \infty$, $\mathbb{E}\,\xi_1' \to m$, $\mathbb{V}\mathrm{ar}\,\xi_1' \to s^2$, $\mathbb{E}\,|\xi_1''| \to 0$, and $\mathbb{V}\mathrm{ar}\,\xi_1'' \to 0$, so in particular we may (and do) choose B large enough so that $\mathbb{E}\,\xi_1' > m/2$, $\mathbb{E}\,|\xi_1''| < \varepsilon/4$, and $\mathbb{V}\mathrm{ar}\,\xi_1'' < \varepsilon^2$.

Since $X_n = X_n' + X_n''$, for any $r > 0$ we have

$$\mathbb{P}(X_n < -r) \le \mathbb{P}(X_n' < -r/2) + \mathbb{P}(X_n'' < -r/2). \tag{2.2.5}$$

Here, since $\mathbb{E}((\xi_k')^4) \le B^4 < \infty$, it follows from Markov's inequality and the Marcinkiewicz–Zygmund inequality, Theorem 1.3.18, that for some constant $C < \infty$ (depending on B),

$$\mathbb{P}(X_n' < -r) \le \mathbb{P}(|X_n' - \mathbb{E}\, X_n'|^4 > (\mathbb{E}\, X_n' + r)^4) \le \frac{Cn^2}{((m/2)n + r)^4}.$$

So

$$\int_{\varepsilon n}^\infty \mathbb{P}(X_n' < -r/2)\mathrm{d}r \le 16Cn^2 \int_0^\infty \frac{\mathrm{d}r}{(mn+r)^4} = O(1/n). \tag{2.2.6}$$

On the other hand, by Chebyshev's inequality, for $r > (\varepsilon/4)n$,

$$\mathbb{P}(X_n'' < -r) \le \mathbb{P}(|X_n'' - \mathbb{E}\, X_n''| > \mathbb{E}\, X_n'' + r) \le \frac{\mathbb{V}\mathrm{ar}\, X_n''}{(r - (\varepsilon/4)n)^2} \le \frac{\varepsilon^2 n}{(r-(\varepsilon/4)n)^2}.$$

Hence

$$\int_{\varepsilon n}^\infty \mathbb{P}(X_n'' < -r/2) \le 4\varepsilon^2 n \int_{\varepsilon n}^\infty \frac{\mathrm{d}r}{(r-(\varepsilon/2)n)^2} = 8\varepsilon. \tag{2.2.7}$$

So from (2.2.5) with (2.2.6) and (2.2.7), we have

$$\limsup_{n\to\infty} \int_{\varepsilon n}^{\infty} \mathbb{P}(X_n < -r)\mathrm{d}r \le 8\varepsilon,$$

which combined with (2.2.4) implies that

$$\limsup_{n\to\infty} \mathbb{E}\, X_n^- \le \frac{s^2\varepsilon}{m^2} + 8\varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, the result follows. □

Using these two lemmas, we can establish the following result which is of some independent interest, and may be known, although we could find no reference.

Lemma 2.2.6. *Suppose that* $\mathbb{E}(\|Z\|^2) < \infty$ *and* $\mu \ne \mathbf{0}$*. Then*

$$0 \le \|S_n\| - S_n \cdot \hat{\mu} \to \frac{\sigma^2_{\mu_\perp}\zeta^2}{2\|\mu\|}, \ \textit{in } L^1, \ \textit{as } n \to \infty,$$

for $\zeta \sim \mathcal{N}(0,1)$*. In particular,*

$$0 \le \mathbb{E}\,\|S_n\| - \|\mu\|n = \frac{\sigma^2_{\mu_\perp}}{2\|\mu\|} + o(1), \ \textit{as } n \to \infty.$$

Proof. As above, for $x \in \mathbb{R}$ set $x^+ := x\mathbf{1}\{x > 0\}$, and also set $x^- = -x\mathbf{1}\{x < 0\}$. Then $x = x^+ - x^-$ and $|x| = x^+ + x^-$, so $x = |x| - 2x^-$; thus $|X_n| - 2X_n^- = X_n \le |X_n|$, and

$$0 \le \|S_n\| - |X_n| \le \|S_n\| - X_n = \|S_n\| - |X_n| + 2X_n^-; \tag{2.2.8}$$

in particular $\mathbb{E}\,\|S_n\| \ge \mathbb{E}\, X_n = \|\mu\|n$. Now, we have from (2.2.3) that

$$\|S_n\| - |X_n| = \frac{Y_n^2}{\|S_n\| + |X_n|} = \frac{n^{-1}Y_n^2}{n^{-1}\|S_n\| + n^{-1}|X_n|},$$

where $n^{-1}Y_n^2 \xrightarrow{\mathrm{d}} \sigma^2_{\mu_\perp}\zeta^2$ for $\zeta \sim \mathcal{N}(0,1)$, and, by the strong law of large numbers, both $n^{-1}\|S_n\|$ and $n^{-1}|X_n|$ tend to $\|\mu\|$ a.s. Hence $0 \le \|S_n\| - |X_n| \xrightarrow{\mathrm{d}} \frac{\sigma^2_{\mu_\perp}\zeta^2}{2\|\mu\|}$, and by Lemma 2.2.4 we conclude that $\|S_n\| - |X_n| \to \frac{\sigma^2_{\mu_\perp}\zeta^2}{2\|\mu\|}$ in L^1. Moreover, Lemma 2.2.5 shows that $X_n^- \to 0$ in L^1. Thus the result follows from (2.2.8). □

We are now, finally in a position to complete the proof of the recasting of Wade and Xu's result, Theorem 2.2.2, and then the proof of the asymptotic expansion of $\mathbb{E}\, L_n$, Theorem 2.2.1.

Proof of Theorem 2.2.2. From the Spitzer–Widom formula (1.1.2) and Lemma 2.2.6, we have

$$\mathbb{E}\,L_n = 2\sum_{k=1}^{n}\frac{1}{k}\left(\|\mu\|k + \frac{\sigma^2_{\mu_\perp}}{2\|\mu\|} + o(1)\right) = 2\|\mu\|n + \frac{\sigma^2_{\mu_\perp}}{\|\mu\|}\log n + o(\log n),$$

as claimed. □

Proof of Theorem 2.2.1. Theorem 2.2.2 shows that

$$n^{-1/2}|\,\mathbb{E}\,L_n - 2\,\mathbb{E}\,S_n \cdot \hat{\mu}| \to 0. \tag{2.2.9}$$

Then by the triangle inequality

$$n^{-1/2}|L_n - 2S_n \cdot \hat{\mu}| \le n^{-1/2}|L_n - \mathbb{E}\,L_n - 2(S_n - \mathbb{E}\,S_n)\cdot\hat{\mu}| + n^{-1/2}|\,\mathbb{E}\,L_n - 2\,\mathbb{E}\,S_n \cdot \hat{\mu}|,$$

which tends to 0 in L^2 by (2.2.1) and (2.2.9). □

It is now a simple exercise to obtain the proof of the inequality for $\mathbb{E}\,D_n$.

Proof of Lemma 2.2.3. The lower bound follows from Lemma 2.2.6 and the fact that $D_n \ge \|S_n\|$. The upper bound follows from the fact that $D_n \le L_n/2$ and the fact that, by Theorem 2.2.2, $\mathbb{E}\,L_n \le 2\|\mu\|n + C(1 + \log n)$. □

2.3 Application of results to our examples

Since this is the first time we see some simulations of our examples, we will start by showing the pictures of the random walks, their convex hulls, and the diameter as a line. The seed for each example is different but fixed throughout the thesis so that we are seeing the picture of the same walk which produces the ensuing simulations demonstrating the laws of large numbers. Here, all of the walks are 2-dimensional both for presentational purposes and because we have so far restricted ourselves to walks following the structure of $(\mathbf{W}^2_\mu)$. In each case, we have taken $n = 10^5$ which is sufficient to get an idea of the shape of the walk and see the results described in this chapter.

2.3.1 Random walk pictures

The first random walk is the simple symmetric random walk. Upon close inspection, the lattice structure can be seen to be exhibited by the horizontal or vertical lines of the walk's steps. The size and shape of the walk itself will be studied further in the next chapter, but one inference from the results above which is backed up by this picture is that the walk certainly does not seem to cover an area of the same scaling as n; in fact, it is totally contained in a 400×400 box despite having 10^5 steps.

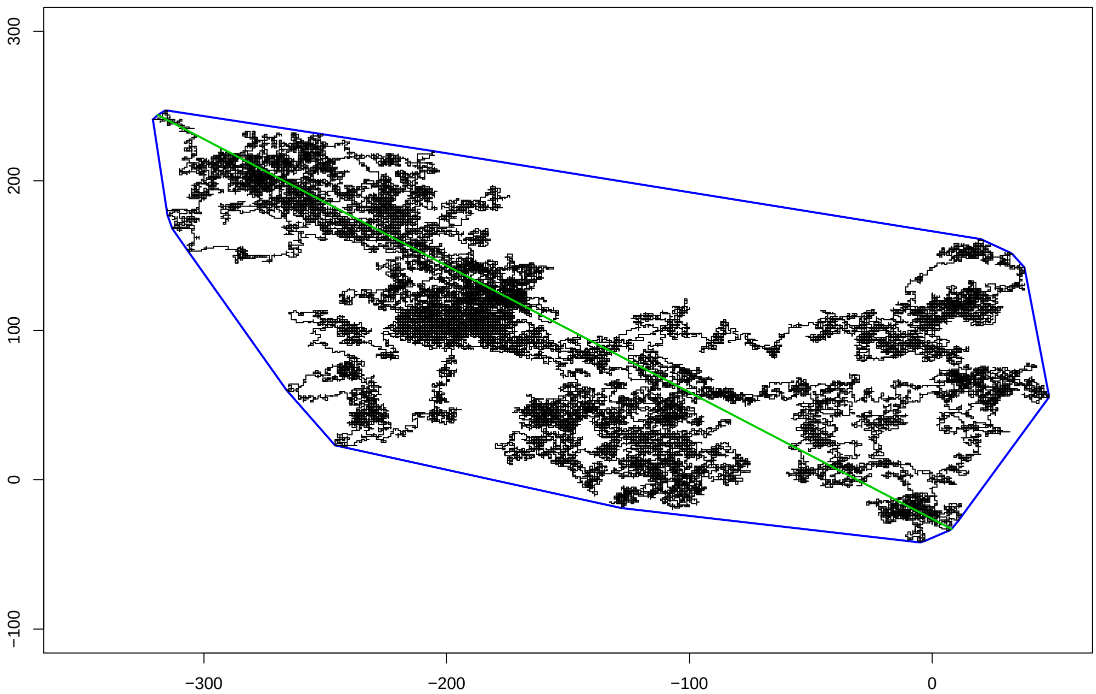


Figure 2.1: A simulation of the simple symmetric random walk with its convex hull and diameter highlighted.

Our second picture, is of the standard Normal random walk. The only significant difference from the first picture, that can be seen here is that this walk is supported on $\mathbb{R}^2$ not just $\mathbb{Z}^2$.

Then we have our first walk with drift, the *random walk with drift and all coordinates Normally distributed*, unit mean to the right. Clearly, we have a very different picture here. Note, the different scale on the horizontal axis and how there is much more of an

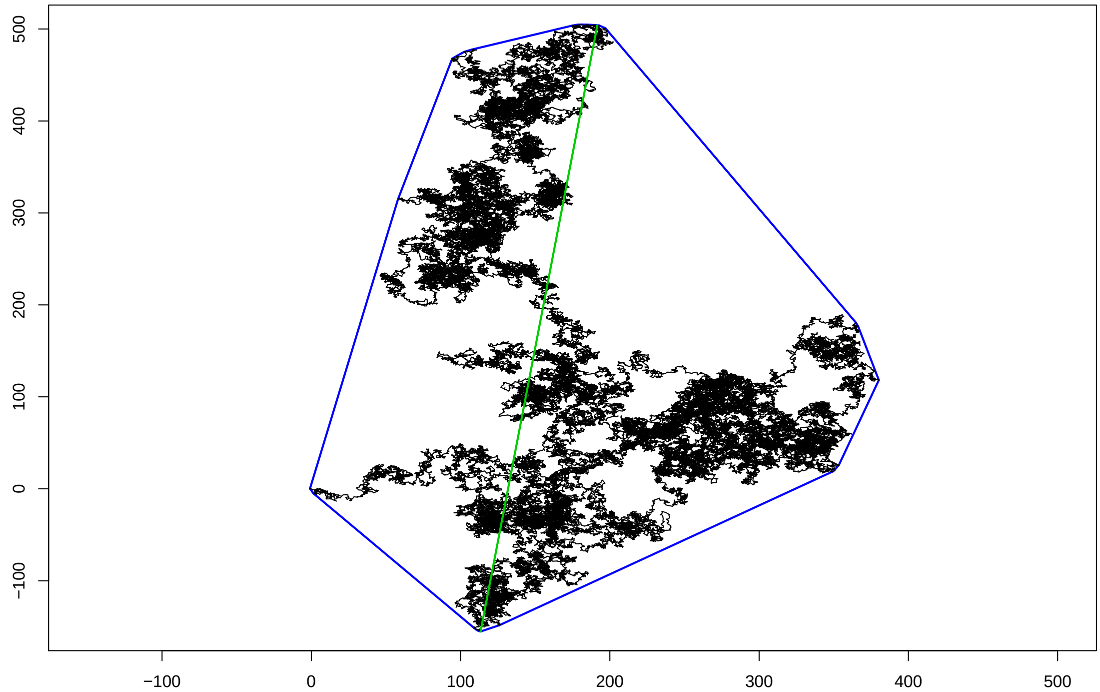


Figure 2.2: A simulation of the standard Normal random walk with its convex hull and diameter highlighted.

elongate shape than the first two pictures, but that the vertical axis is a similar scale to the zero drift walks.

Figure 2.3: A simulation of the random walk with drift and all coordinates Normally distributed, unit mean, with its convex hull and diameter highlighted.

For the walk with Normal increments but $\|\mu\| = 5$, the picture is not too different. As seems natural, the horizontal axis goes up to 5×10^5 compared to 10^5 from before. Finally, we have the walk with fixed drift. This walk also seems to be similar to the previous two with drift.

One might be tempted to mention the differing angles of the hulls, but the scale on the vertical axis must be considered when doing this. In reality, the differing angles are only natural variances in the walks, and we include a seemingly non-sensical plot of the walk where we have the vertical axis on the same scale as the horizontal one. Here, we see how the angle is only a feature due to the more detailed axis, and that the other two plots would look the same on this macroscopic scale. This also reinforces the idea that the convex hulls of the walks with drift really behave like a straight line in terms of their perimeter lengths and diameters.

Figure 2.4: A simulation of the random walk with drift and all coordinates Normally distributed, mean of length 5, with its convex hull and diameter highlighted.

Figure 2.5: A simulation of the random walk with drift and no variance in the first coordinate, unit mean, with its convex hull and diameter highlighted.

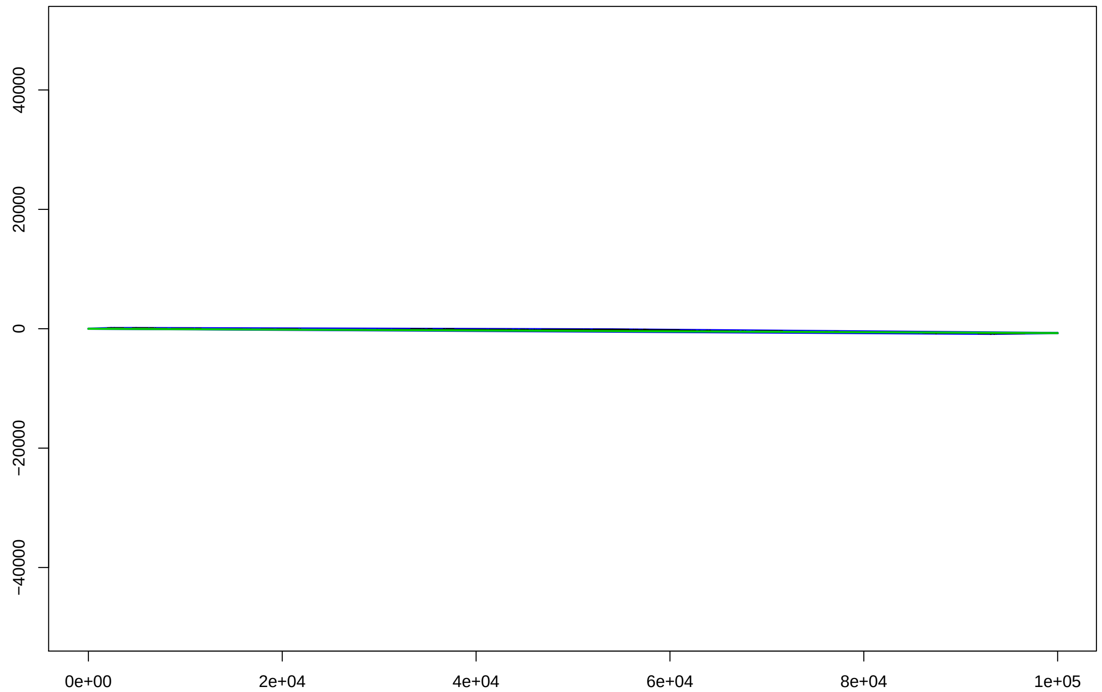


Figure 2.6: A simulation of the random walk with drift and no variance in the first coordinate, unit mean, with its convex hull and diameter highlighted, with both axes on the same scale.

2.3.2 Law of large numbers pictures

The first two of these pictures really tell the same story. The scaled perimeter lengths, in blue, and the scaled diameters, in green, converge to 0 as n tends to infinity. It is not surprising the perimeter length is larger than the diameter, as noted above $L_n \geq 2D_n$.

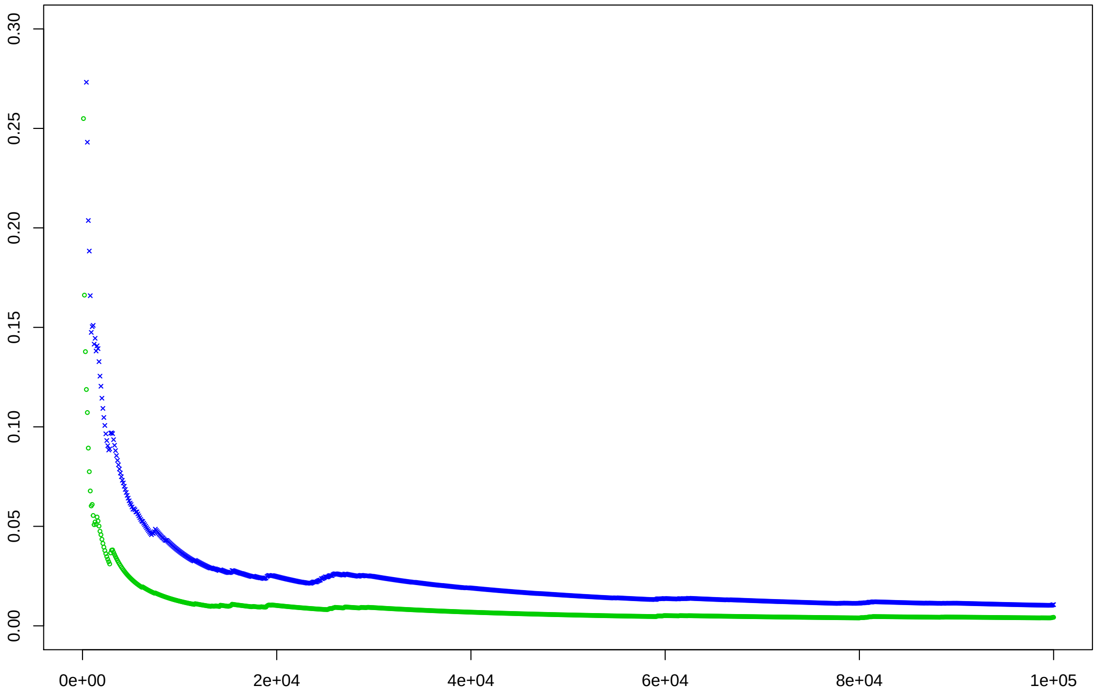


Figure 2.7: A demonstration of the law of large numbers applied to the simple symmetric random walk. The scaled perimeter length, in blue, and scaled diameter, in green, are plotted for the first 10^5 steps.

Then we see the theorem in action for the walks with drift. The perimeter length converging to $2\|\mu\|$ and the diameter to $\|\mu\|$. The convergence is so quick that the right hand side of both plots looks like two straight lines in all three cases.

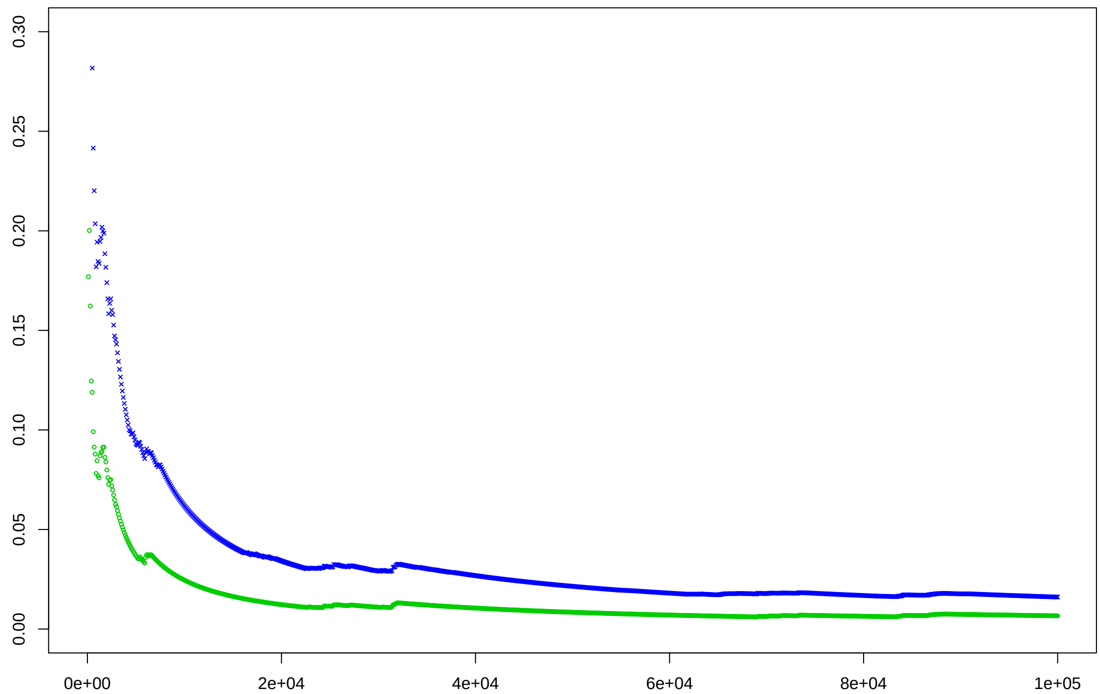


Figure 2.8: A demonstration of the law of large numbers applied to the standard Normal random walk. The scaled perimeter length, in blue, and scaled diameter, in green, are plotted for the first 10^5 steps.

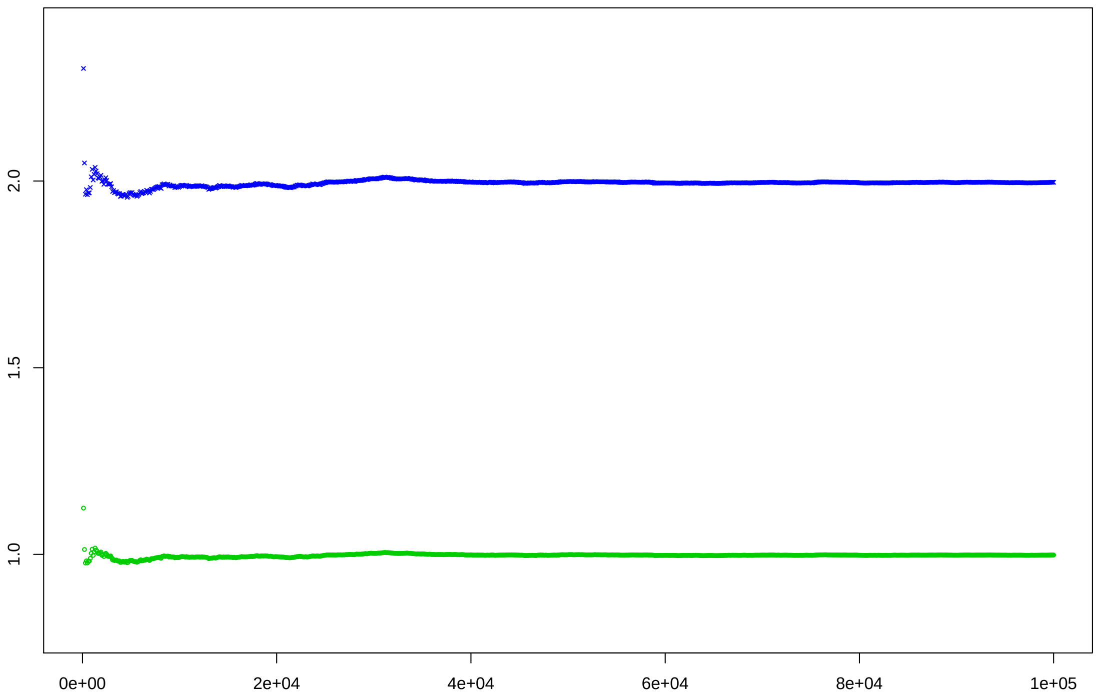


Figure 2.9: A demonstration of the law of large numbers applied to the random walk with Normal drift, unit mean. The scaled perimeter length, in blue, and scaled diameter, in green, are plotted for the first 10^5 steps.

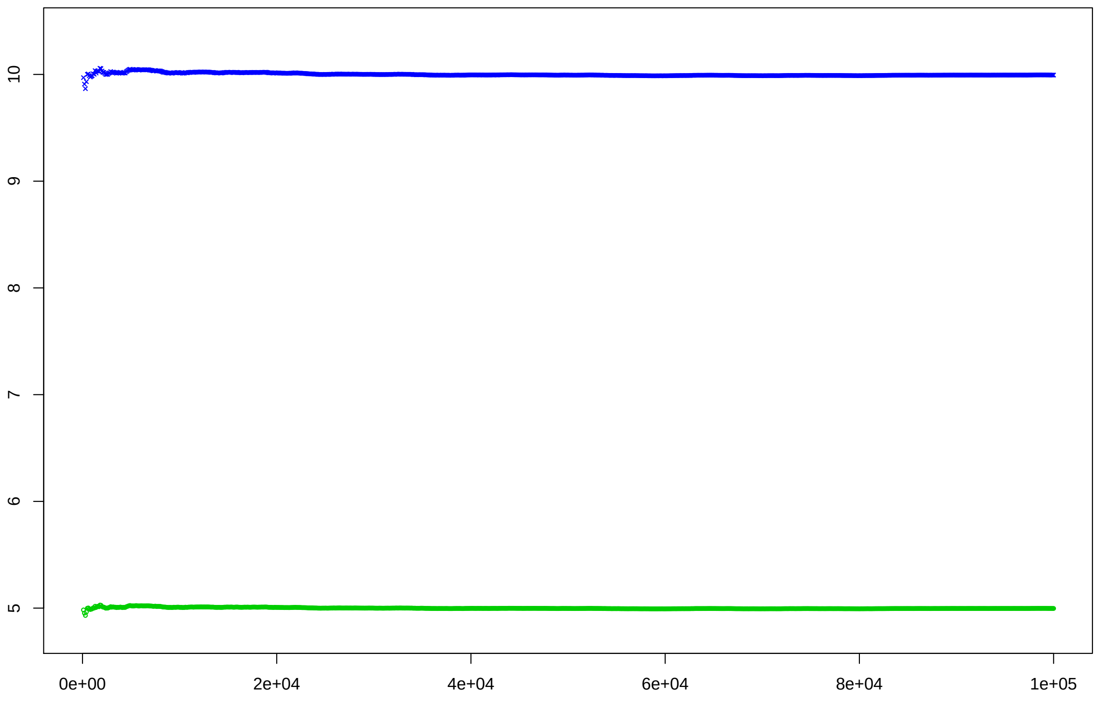


Figure 2.10: A demonstration of the law of large numbers applied to the random walk with Normal drift, mean of length 5. The scaled perimeter length, in blue, and scaled diameter, in green, are plotted for the first 10^5 steps.

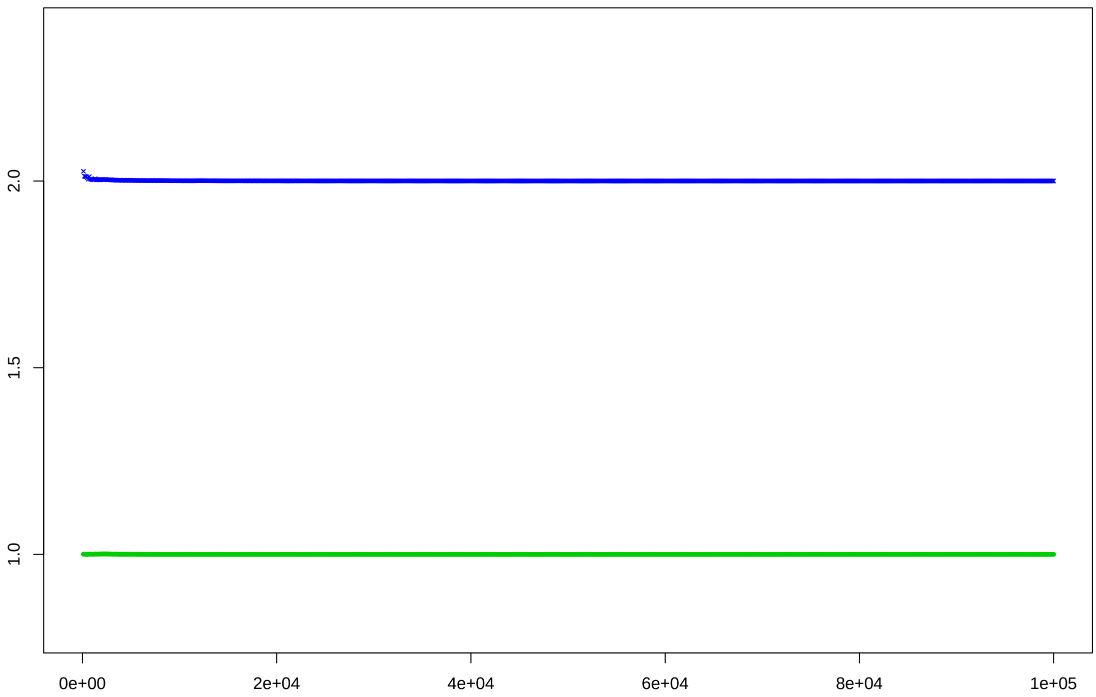


Figure 2.11: A demonstration of the law of large numbers applied to the random walk with fixed drift, unit mean. The scaled perimeter length, in blue, and scaled diameter, in green, are plotted for the first 10^5 steps.

2.3.3 Case with drift ratio simulation

We only show the results for the walks with drift here. The equivalent result for the zero drift case is the subject of Chapter 4.

First, we have the Normal unit drift, for which we show the result with the vertical axis scaled to see the whole range of possible values the ratio can take.

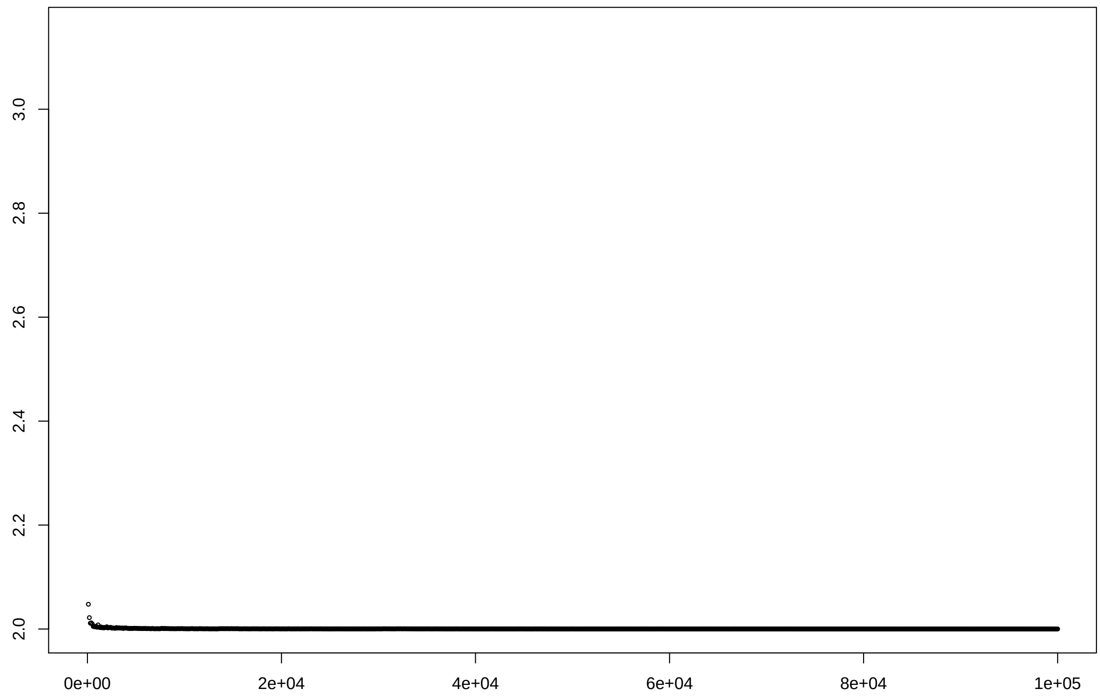


Figure 2.12: A demonstration of the law of large numbers for the ratio of the perimeter length and diameter applied to the random walk with drift and all coordinates Normally distributed, unit mean. The ratio is plotted for the first 10^5 steps.

Clearly, this scaling, although informative in some sense, is too zoomed out to see anything interesting, so we provide the same plot with the vertical axis only showing values near to 2.

Even on this scaling, the convergence seems to be fairly fast. We provide similarly scaled (but note, not exactly) plots for the other two walks with drift.

In all of the pictures, the convergence of L_n/D_n to 2 is supported by the plot.

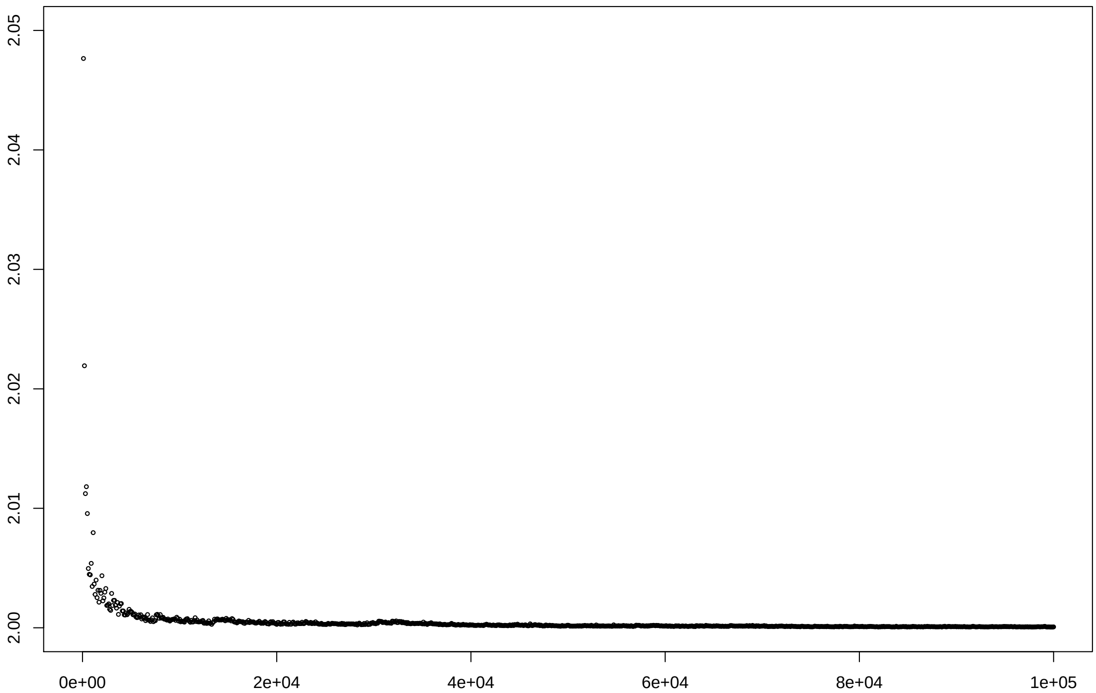


Figure 2.13: A demonstration of the law of large numbers for the ratio of the perimeter length and diameter applied to the random walk with drift and all coordinates Normally distributed, unit mean. The ratio is plotted for the first 10^5 steps, with the vertical axis showing values near to 2.

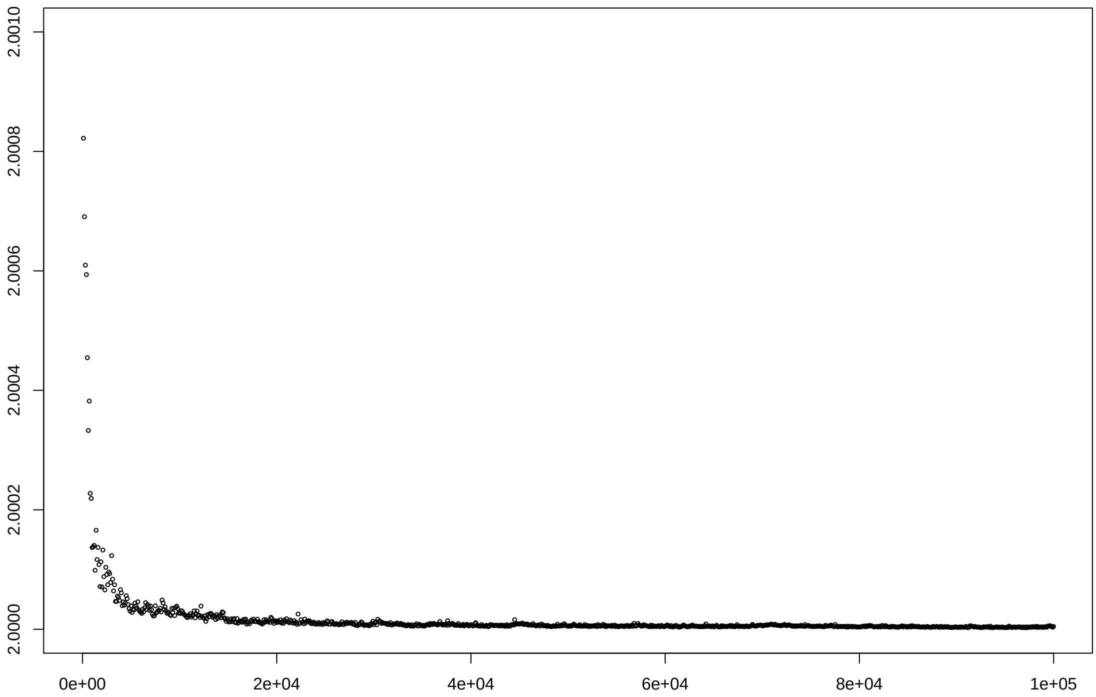


Figure 2.14: A demonstration of the law of large numbers for the ratio of the perimeter length and diameter applied to the random walk with drift and all coordinates Normally distributed, mean of length 5. The ratio is plotted for the first 10^5 steps, with the vertical axis showing values near to 2.

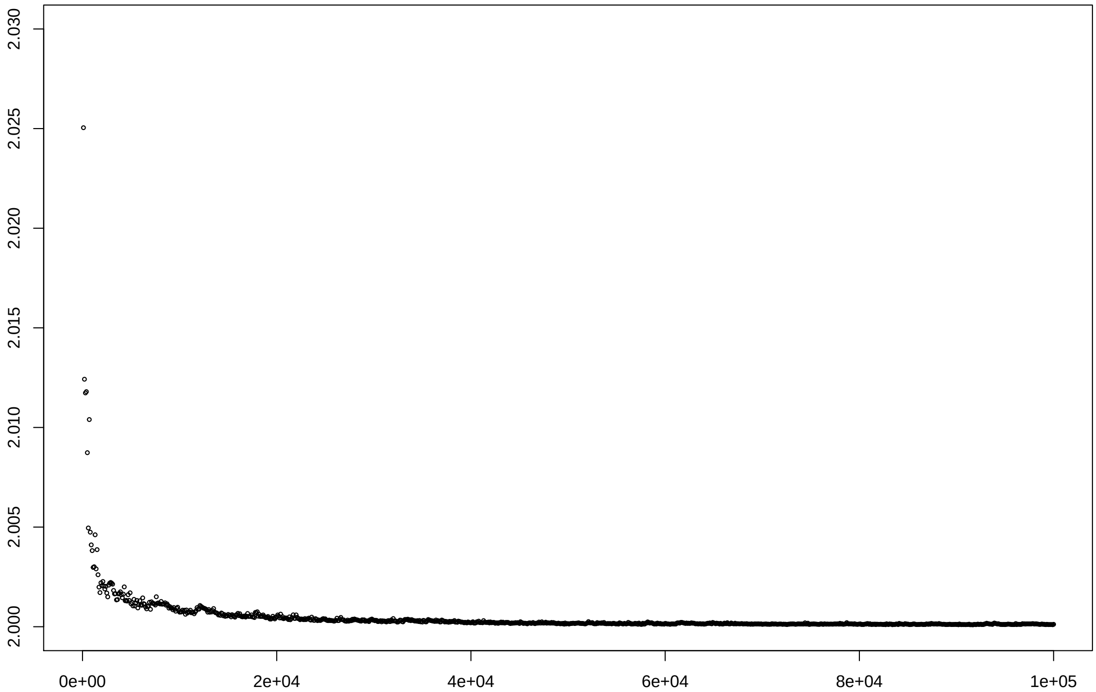


Figure 2.15: A demonstration of the law of large numbers for the ratio of the perimeter length and diameter applied to the random walk with drift and no variance in the first coordinate, mean of length 5. The ratio is plotted for the first 10^5 steps, with the vertical axis showing values near to 2.

Chapter 3

Functional limit approach

The results on the perimeter length and diameter in Chapter 2 can be interpreted as an indication that the convex hull exhibits particular shapes after the random walk takes a large number of steps. In this chapter, we consider this question of the shape of the random walk and in turn the convex hull, and how this can be extended to be informative about functionals of the convex hull, not least, returning to the perimeter length and diameter[1]. Here we do not restrict ourselves to 2 dimensions as before, so our functionals become more general too.

The heuristic idea underlying the functional limit theorems starts with the story of law of large numbers and central limit theorem, see Theorem 1.3.14 and Theorem 1.3.15. These early results refer to the sums of random variables, or the endpoint of our random walks, which were the first real quantities of interest in this area due to their application in the contexts of long run profit in gambling games and errors when sampling large amounts of data. The functional limit theorem extension of these results, presented in Billingsley [Bil99], considers the paths of the random walks not just the endpoints confirming the intuitive idea that the path moves linearly towards its endpoint, at least on the law of large numbers scaling.

As the pictures of our examples in Section 2.3 suggest, the case with drift and the zero drift case will, unsurprisingly, be seen to behave differently. Under the law of large

[1]All sections in this chapter except Section 3.4 are based on work published in [LMW18]. The theory and the maximum functional example were joint work with the other authors, but the arc-sine law and the convex hull material was written independently.

numbers scaling, the walk with drift converges to the unit vector in the direction of the mean whilst the zero-drift case is degenerate converging to the point at the origin (which itself could be considered as the degenerate unit vector in the direction of the origin). The zero-drift case attains a non-trivial limiting distribution under the central limit theorem scaling, for which it converges to Brownian motion in the sense of weak convergence of functions, to be formally defined later. Our contribution to the theory for the random walks is to show that these results extend to higher dimensions in the natural way.

In order to see the theory in action, we present the example of the maximum functional which also serves to demonstrate how the continuous mapping theorem can be used to determine information about functionals of the random walk. We then give another example, this time an original generalisation of the arcsine law to higher dimensions which states that the walk's direction has no limiting direction or subset of directions. This example is nicely coherent with our shape result in Chapter 4 which, loosely speaking, says that the random walk with zero-drift approximates any shape with unit diameter infinitely often, after appropriate scaling. If the random walk's direction had a limiting subset of the sphere as our arc-sine law rules out, then it would not be surprising if some shapes could not be well-approximated infinitely often because this would require increasingly large, and unlikely, jumps. Conversely, if the walk had a limiting shape, it would not be surprising to find that the walk's directions had a limiting set, or at least could not infinitely often spend almost every time point in a subset which would contradict the directions of points in the limiting shape. So although the arc-sine law relates to the walk and not the convex hull, they should not be seen as isolated results.

With the strategy understood and preliminary examples presented, we consider the point set of the random walk. Without even taking the convex hull of these points, we will already be in a position to study the diameter of the convex hull, because this coincides with the diameter of the point set. Next, using either this convergence of sets, or the trajectory convergence, we will extend the results to the convergence of the convex hulls and consider the convergence of the mean width, volume and surface area, all defined later on. Again, the results are quite different in the case with drift, which will see deterministic or trivial limits based on the convergence to the line segment in the direction of the mean, whilst the zero-drift case will see convergence to the random limits related to Brownian motion. It is also possible to see a non-trivial scaling of the case with drift to a Brownian limit, but this requires differential scaling of the walk with the direction of the mean scaled by n^{-1} and all orthogonal directions scaled by $n^{-1/2}$. Further details can be found in, and directly proceeding, Lemma 3.3.6.

Finally, motivated by the convergence to Brownian motion, we make some comments about known results relating to certain functionals of Brownian motion so that we can get a better intuition of what our zero drift results actually tell us. This will include a new result on the expected diameter of Brownian motion in 2 dimensions[2].

We conclude the chapter by showing some of the results in action through our simulated examples and discussing open problems in this area.

It is to be noted that the extension to higher dimensions of the results regarding the random walk theory are non-trivial, but the methods follow closely that of Billingsley, and so we only state the results in the main text without their proofs which are lengthy. We do this so that we can keep the focus and flow of the results related to convex hulls, however, for reference and completeness, the proofs and some of the surrounding discussion are attached in the appendix.

3.1 Random walk convergence

First, in order to answer the question of progress towards the endpoint, $n\mu$, it is necessary to define the trajectory of the walk. First, consider the discrete jump process of the partial sums with each time step rescaled by $1/n$, so the partial sums are indexed by times in the the interval $[0, 1]$, in fact they are at the times $\frac{k}{n}$ with $k = 0, 1, \dots, n$. However, we wish to consider a continuous-time trajectory, so we have two choices on how to fill in the gaps. Either we can say the walk moves linearly from each partial sum to the next, in which case the trajectory is

$$X_n(t) := n^{-1}\left(S_{\lfloor nt \rfloor} + (nt - \lfloor nt \rfloor) Z_{\lfloor nt \rfloor + 1}\right), \tag{3.1.1}$$

[2]Based on work published in [MX17] which was the product of joint work between the authors.

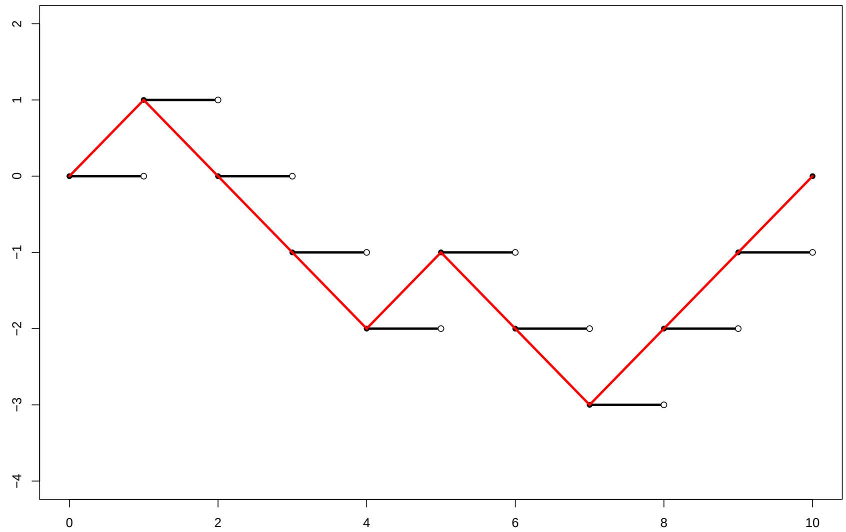


Figure 3.1: An example of a possible one-dimensional random walk plotted with time on the horizontal axis, with the two continuous-time trajectories we can create from it; the continuous interpolating $X_n(t)$ in red and the piecewise constant process $X'_n(t)$ in black.

or we can consider the trajectory where the walk 'stays still' and makes small jumps when it reaches the next time indexing a new partial sum, in which case we have

$$X'_n(t) := \frac{1}{n} S_{\lfloor nt \rfloor}. \tag{3.1.2}$$

In order to study convergence of these trajectories, we need to specify the metric spaces in which they live. Conveniently, we have defined three such spaces in Section 1.3, the continuous trajectories endowed with the supremum norm $(\mathcal{C}^d, \rho_\infty)$ and the Skrorokhod metric or Kolmogorov-Billingsley metric on $\mathcal{D}^d$, $(\mathcal{D}^d, \rho_S)$ or $(\mathcal{D}^d, \rho_S^\circ)$. The first space will be used to show the convergence of $X_n(t)$ which is itself a continuous function, and the latter two will be used when discussing $X'_n(t)$. It would not be unreasonable to question why we have two metrics for $X'_n(t)$, but the following result should make this a bit clearer.

Proposition 3.1.1 ([Kol56, Theorem 7]). *The metrics ρ_S° and ρ_S are equivalent. That is, for a sequence of functions $f, f_1, f_2, \ldots$ on $\mathcal{D}^d$, $\rho_S(f_n, f) \to 0$ as $n \to \infty$ if and only if $\rho_S^\circ(f_n, f) \to 0$ as $n \to \infty$.*

The fact that the metrics are equivalent means we can use either one to prove continuity of a functional on $\mathcal{D}^d$ and the result will hold for the other; we will use the metric which is simplest for each application. Likewise, almost-sure statements using one metric carry over to the other. Note also that as equivalent metrics, ρ_S and ρ_S° generate the same topology (open sets) on $\mathcal{D}^d$, and hence also the same Borel sets. Further motivation behind ρ_S° in particular is that, under this metric, $\mathcal{D}^d$ is both separable and complete, which is useful for the proofs of Donsker's theorem below. Further details on this subject are left to the appendix.

We now can state the important results regarding the almost sure convergence of our trajectories.

Theorem 3.1.2 (Functional law of large numbers [Whi02, p. 26])**.** *Consider the random walk trajectories as defined at* (3.1.1) *and* (3.1.2)*. Let* $I_\mu \in \mathcal{C}^d$ *be the function defined by* $I_\mu(t) := \mu t$ *for* $t \in [0, 1]$*.*

(a) We have $X_n \xrightarrow{\text{a.s.}} I_\mu$ *on* $(\mathcal{C}_0^d, \rho_\infty)$*.*

(b) We have $X'_n \xrightarrow{\text{a.s.}} I_\mu$ *on* $(\mathcal{D}_0^d, \rho_\infty)$*.*

Remark 3.1.3. By Lemma 1.3.2, part (b) also shows that $X'_n \xrightarrow{\text{a.s.}} I_\mu$ on $(\mathcal{D}_0^d, \rho_S)$ and Proposition 3.1.1 in turn shows that $X'_n \xrightarrow{\text{a.s.}} I_\mu$ on $(\mathcal{D}_0^d, \rho_S^\circ)$.

Alongside the convergence of the trajectories, we will need the following mapping theorem in order to extend the results to further functionals of the trajectories including our convex hull properties. First, note that, given two metric measure spaces $(S, \mathcal{S}, \rho)$ and $(S', \mathcal{S}', \rho')$ and a measurable function $h : S \to S'$, the set D_h of discontinuities of h satisfies $D_h \in \mathcal{S}$: see [Bil99, p. 243], and hence $\mathbb{P}(X \in D_h)$ is well defined.

Theorem 3.1.4 (Continuous mapping theorem for almost-sure convergence [Gut05, p. 244])**.** *Let* $X, X_1, X_2, \dots$ *be random variables on the probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *taking values in the metric measure space* $(S, \mathcal{S}, \rho)$*. Let* $(S', \mathcal{S}', \rho')$ *be another metric measure space, and let* $h : (S, \mathcal{S}, \rho) \to (S', \mathcal{S}', \rho')$ *be measurable. If* $X_n \xrightarrow{\text{a.s.}} X$ *and* $\mathbb{P}(X \in D_h) = 0$*, then* $h(X_n) \xrightarrow{\text{a.s.}} h(X)$*.*

In the zero drift case, we need a new scaling, but still maintain the two different trajectories. Precisely, for $n \in \mathbb{N}$ and $t \in [0,1]$ we define

$$\begin{aligned} Y_n(t) &:= \frac{1}{\sqrt{n}} \left(S_{\lfloor nt \rfloor} + (nt - \lfloor nt \rfloor) \xi_{\lfloor nt \rfloor + 1} \right); \\ Y_n'(t) &:= \frac{1}{\sqrt{n}} S_{\lfloor nt \rfloor}. \end{aligned} \tag{3.1.3}$$

Here $Y_n \in \mathcal{C}_0^d$ and $Y_n' \in \mathcal{D}_0^d$. Then, recalling b_d is a standard d-dimensional Brownian motion, we can now state the weak convergence result for our zero-drift trajectories. The one-dimensional case was first proven in [Don51], and further discussion can be found in, for example, [Bil99; Kal02]. We also point the reader to [EK09, §5] for a comprehensive discussion of both d-dimensional Brownian motion and the steps leading to this result.

Theorem 3.1.5 (Donsker's theorem)**.** *Suppose that we have a random walk as defined at* $(\mathbf{W}_\mu)$ *with* $\mu = \mathbf{0}$ *and satisfying* $(\mathbf{V})$.

(a) We have $Y_n \Rightarrow \Sigma^{1/2} b_d$ *in the sense of weak convergence on* $(\mathcal{C}_0^d, \rho_\infty)$.

(b) We have $Y_n' \Rightarrow \Sigma^{1/2} b_d$ *in the sense of weak convergence on* $(\mathcal{D}_0^d, \rho_S)$.

As with the almost-sure convergence result, we will need a mapping theorem in order to extend our results to functionals of the random walk.

Theorem 3.1.6 (Continuous mapping theorem for weak convergence [Bil99, p. 20])**.** *Let* $P, P_1, P_2, \ldots$ *be a sequence of probability measures on a metric measure space* $(S, \mathcal{S}, \rho)$. *Let* $(S', \mathcal{S}', \rho')$ *be another metric measure space, and let* $h : (S, \mathcal{S}, \rho) \to (S', \mathcal{S}', \rho')$ *be measurable. For each* n, *we define* $P_n h^{-1}$, *a probability measure on* $(S', \mathcal{S}', \rho')$ *by* $P_n h^{-1}(A) = P_n(h^{-1}(A))$ *for* $A \in \mathcal{S}'$. *If* $P_n \Rightarrow P$ *and* $P(D_h) = 0$, *then* $P_n h^{-1} \Rightarrow P h^{-1}$.

Corollary 3.1.7. *If* $X_n \Rightarrow X$ *and* $\mathbb{P}(X \in D_h) = 0$, *then* $h(X_n) \Rightarrow h(X)$.

Remark 3.1.8. Part (b) of Theorem 3.1.5 is stated for the space $(\mathcal{D}_0^d, \rho_S)$, but weak convergence on $(\mathcal{D}_0^d, \rho_S)$ is equivalent to weak convergence on $(\mathcal{D}_0^d, \rho_S^\circ)$. To see this, recall the definition of weak convergence from Definition 1.3.7 with the probability measure discussion above there, and note Proposition 3.1.1 which tells us that a continuous function f under one metric is continuous under the other. Thus, the set of bounded continuous functions is the same in both metric spaces and so weak convergence must be equivalent.

3.1.1 The maximum functional

As a first example of the theory developed above, we consider a d-dimensional version of the *maximum functional* $M : \mathcal{M}^d \to \mathbb{R}$ defined by $M(f) := \sup_{0 \le t \le 1} \|f(t)\|$, where we recall $\mathcal{M}^d$ is the set of trajectories, see Section 1.3. Note that $|M(f)| \le \|f\|_\infty$. The next result shows that M is a continuous map from $(\mathcal{M}^d, \rho_\infty)$ to $(\mathbb{R}, \rho_E)$ and also a continuous map from $(\mathcal{M}^d, \rho_S)$ to $(\mathbb{R}, \rho_E)$.

Theorem 3.1.9. *For any $f, g \in \mathcal{M}^d$ we have $|M(f) - M(g)| \le \rho_S(f, g) \le \rho_\infty(f, g)$.*

Proof. Take $f, g \in \mathcal{M}^d$, and suppose without loss of generality that $\sup_{s\in[0,1]} \|f(s)\| \ge \sup_{t\in[0,1]} \|g(t)\|$. Recall Λ is the set of $\lambda : [0, 1] \to [0, 1]$ that are strictly increasing and surjective. For any $\lambda \in \Lambda'$,

$$\begin{aligned} |M(f) - M(g)| &= \sup_{s\in[0,1]} \|f(s)\| - \sup_{t\in[0,1]} \|g(t)\| \\ &= \sup_{s\in[0,1]} \|f(s)\| - \sup_{t\in[0,1]} \|g \circ \lambda(t)\|, \end{aligned}$$

since $\lambda[0, 1] = [0, 1]$. Hence

$$\begin{aligned} |M(f) - M(g)| &= \sup_{s\in[0,1]} \left(\|f(s)\| - \sup_{t\in[0,1]} \|g \circ \lambda(t)\| \right) \\ &\le \sup_{s\in[0,1]} (\|f(s)\| - \|g \circ \lambda(s)\|) \\ &\le \sup_{s\in[0,1]} \|f(s) - g \circ \lambda(s)\| \\ &= \|f - g \circ \lambda\|_\infty \\ &\le \|\lambda - I\|_\infty \vee \|f - g \circ \lambda\|_\infty. \end{aligned}$$

We therefore have that

$$|M(f) - M(g)| \le \inf_{\lambda\in\Lambda} \{\|\lambda - I\|_\infty \vee \|f - g \circ \lambda\|_\infty\} \le \rho_S(f, g).$$

Lemma 1.3.2 completes the proof. □

Since we have shown the maximum functional is continuous, we can also apply the mapping theorem to the functional law of large numbers, to obtain the following result.

Theorem 3.1.10. *Consider the random walk defined at* $(\mathbf{W}_\mu)$*. Then, as* $n \to \infty$*,*

$$\frac{1}{n} \max_{0 \le k \le n} \|S_k\| \xrightarrow{\text{a.s.}} \|\mu\|.$$

Proof. Let $X_n'(t)$ be as defined at (3.1.2). The functional strong law of large numbers, Theorem 3.1.2, says that $X_n' \xrightarrow{\text{a.s.}} I_\mu$ on $(\mathcal{D}^d, \rho_\infty)$, while Theorem 3.1.9 says that M is continuous. Thus the mapping theorem, Theorem 3.1.4, implies that $M(X_n') \xrightarrow{\text{a.s.}} M(I_\mu)$ on $(\mathbb{R}, \rho_E)$. But $M(X_n') = n^{-1} \max_{0 \le k \le n} \|S_k\|$ and $M(I_\mu) = \|\mu\|$, giving the result. □

Further to the law of large numbers scaling result, we can apply the mapping theorem to the functional central limit theorem to establish the following result in the case of zero drift.

Theorem 3.1.11. *Suppose that we have a random walk as defined at* $(\mathbf{W}_\mu)$ *with* $\mu = \mathbf{0}$*, and satisfying* $(\mathbf{V})$*. Then as* $n \to \infty$*,*

$$\frac{1}{\sqrt{n}} \max_{0 \le k \le n} \|S_k\| \xrightarrow{\text{d}} \sup_{t \in [0,1]} \|\Sigma^{1/2} b_d(t)\|.$$

Proof. Donsker's theorem, Theorem 3.1.5, together with the mapping theorem, Corollary 3.1.7, and continuity of the function $M : (\mathcal{D}^d, \rho_S) \to (\mathbb{R}_+, \rho_E)$, Theorem 3.1.9, shows that

$$M(Y_n') = \sup_{t \in [0,1]} \|Y_n'(t)\| \xrightarrow{\text{d}} M(\Sigma^{1/2} b_d) = \sup_{t \in [0,1]} \|\Sigma^{1/2} b_d(t)\|.$$

But we have that $\sup_{t \in [0,1]} \|Y_n'(t)\| = n^{-1/2} \max\{\|S_0\|, \|S_1\|, \dots, \|S_n\|\}$, completing the proof. □

Remark 3.1.12. In the case $\Sigma = I_d$, the d-dimensional identity matrix, the right hand side of Theorem 3.1.11 is concerned with the maximum of a d-dimensional Bessel process. In the case $d = 1$, the maximum functional would more naturally be $M(f) := \sup_{0 \le t \le 1} f(t)$, which would give different results. This functional was presented in

[LMW18]. In this case, the distribution of $\sup_{t\in[0,1]} b(t)$ can be determined by the reflection principle for Brownian motion, and so Theorem 3.1.11 would give us the limiting distribution for $\max_{1\le k\le n} S_k/\sqrt{n}$: see [Bil99, pp. 91–93].

3.1.2 Generalisation of the arcsine law

We now turn to our second example. The classical arcsine law states the following [Fel68, p. 82], first established for the simple symmetric random walk.

Theorem 3.1.13. *If $0 < \gamma < 1$, the probability that an n-step simple symmetric random walk spends less than γn time on the positive side tends to $2\pi^{-1} \arcsin \sqrt{\gamma}$ as $n \to \infty$.*

Discussion of this result and its connection to other functionals in the one-dimensional case can be found at [Bil99, pp. 97–101]. We wish to extend the result to higher dimensions, which requires a generalisation of the functional itself. In [BD88] the functional which generalises 'time on the positive side' to 'time in the positive quadrant' is considered and shown not to follow an arc-sine distribution by comparison of moments. The generalisation that we will consider is $\pi_n(A)$, defined to be the proportion of time the normalised walk spends in a given subset of the sphere. Formally, recall $\hat{\mathbf{x}} := \mathbf{x}/\|\mathbf{x}\|$ for $\mathbf{x} \neq \mathbf{0}$ and $\hat{\mathbf{0}} := \mathbf{0}$, then, for a measurable set $A \subseteq \mathbb{S}^{d-1}$,

$$\pi_n(A) := \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\{\hat{S}_i \in A\}.$$

Theorem 3.1.14. *Suppose that we have a random walk as defined at ($\mathbf{W}_\mu$) with $\mu = \mathbf{0}$, and satisfying ($\mathbf{V}$). Let $\hat{b}_d^\Sigma(t) := \Sigma^{1/2} b_d(t)/\|\Sigma^{1/2} b_d(t)\|$, the d-dimensional Brownian motion projected onto the sphere and $A \subseteq \mathbb{S}^{d-1}$ with $\mu_{d-1}(\partial A) = 0$, where μ_{d-1} here denotes Haar measure on $\mathbb{S}^{d-1}$. Then as $n \to \infty$,*

$$\pi_n(A) \xrightarrow{\mathrm{d}} \int_0^1 \mathbf{1}\{\hat{b}_d^\Sigma(t) \in A\} \mathrm{d}t.$$

For further details on Haar measures, see [Hal50, §58]. As with all our examples, we must prove the continuity of the functional in order to complete the proof. First, for measurable $A \subseteq \mathbb{S}^{d-1}$ and $f \in \mathcal{D}^d$, define

$$\varpi_A(f) := \int_0^1 \mathbf{1}\left\{\widehat{f(t)} \in A\right\} \mathrm{d}t.$$

Note that $\pi_n(A) = \varpi_A(Y_n')$.

Lemma 3.1.15. *Fix a measurable $A \subseteq \mathbb{S}^{d-1}$. Then, as a function from $(\mathcal{D}^d, \rho_S)$ to $([0,1], \rho_E)$, $f \mapsto \varpi_A(f)$ is continuous on the set*

$$F_A := \left\{ f \in \mathcal{D}^d : \int_0^1 \mathbf{1}\left\{\widehat{f(t)} \in \{\mathbf{0}\} \cup \partial A\right\} \mathrm{d}t = 0 \right\}.$$

Proof. Since ρ_S and ρ_S° are equivalent (see Proposition 3.1.1), it suffices to work with the latter. For $f \in \mathcal{D}^d$ define for all measurable $B \subseteq \mathbb{R}^d$,

$$\nu_f(B) := \int_0^1 \mathbf{1}\{f(t) \in B\} \mathrm{d}t.$$

Note that ν_f is a finite measure on $\mathbb{R}^d$. Now let $\tilde{A} = \{rx : x \in A,\, r > 0\}$, then $\varpi_A(f) = \nu_f(\tilde{A})$. Take $f, g \in \mathcal{D}^d$ and suppose, without loss of generality, that $\nu_f(\tilde{A}) \geq \nu_g(\tilde{A})$, let $\tilde{A}^\varepsilon = \{x \in \mathbb{R}^d : \rho_E(x, \tilde{A}) \leq \varepsilon\}$ and let $\tilde{A}_\varepsilon = \{x \in \mathbb{R}^d : \rho_E(x, \mathbb{R}^d \setminus \tilde{A}) \geq \varepsilon\}$, then $\tilde{A}_\varepsilon \subseteq \tilde{A} \subseteq \tilde{A}^\varepsilon$ and

$$|\varpi_A(f) - \varpi_A(g)| = \nu_f(\tilde{A}) - \nu_g(\tilde{A}) = \nu_f(\tilde{A}_\varepsilon) - \nu_g(\tilde{A}^\varepsilon) + \nu_f(\tilde{A} \setminus \tilde{A}_\varepsilon) + \nu_g(\tilde{A}^\varepsilon \setminus \tilde{A}).$$

If $f, g \in F_A$ then since $x \in \partial\tilde{A}$ implies $\hat{x} \in \{\mathbf{0}\} \cup \partial A$, we have that as $\varepsilon \to 0$, by continuity of measures along monotone limits, $\nu_f(\tilde{A} \setminus \tilde{A}_\varepsilon) \to \nu_f(\partial\tilde{A}) = 0$, and $\nu_g(\tilde{A}^\varepsilon \setminus \tilde{A}) \to \nu_g(\partial\tilde{A}) = 0$. Moreover, we can use the change of variable $t = \lambda(s)$ in the ν_g-integral, where $\lambda \in \Lambda$, in order to appeal to Lemma A.1.3. As in that lemma, we note that, by Lebesgue's theorem on the differentiability of monotone functions, see [KF12, p. 321], $\lambda'(t)$ exists almost everywhere on $t \in (0,1)$, and so for any $\delta > 0$

$$\begin{aligned}
\nu_f(\tilde{A}_\varepsilon) - \nu_g(\tilde{A}^\varepsilon) &= \int_0^1 \mathbf{1}\{f(t) \in \tilde{A}_\varepsilon\}\mathrm{d}t - \int_0^1 \mathbf{1}\{g(t) \in \tilde{A}^\varepsilon\}\mathrm{d}t \\
&\leq \int_0^1 \mathbf{1}\{f(t) \in \tilde{A}_\varepsilon\}\mathrm{d}t - \int_0^1 \mathbf{1}\{g(t) \in \tilde{A}^\varepsilon, \lambda'(t) \text{ exists}\}\mathrm{d}t \\
&= \int_0^1 \mathbf{1}\{f(t) \in \tilde{A}_\varepsilon\}\mathrm{d}t - \int_0^1 \lambda'(s)\mathbf{1}\{g(\lambda(s)) \in \tilde{A}^\varepsilon, \lambda'(s) \text{ exists}\}\mathrm{d}s, \\
&\leq \gamma + \int_0^1 \mathbf{1}\{f(t) \in \tilde{A}_\varepsilon,\, g(\lambda(t)) \notin \tilde{A}^\varepsilon, \lambda'(t) \text{ exists}\}\mathrm{d}t \\
&\qquad - \int_0^1 \mathbf{1}\{f(t) \notin \tilde{A}_\varepsilon,\, g(\lambda(t)) \in \tilde{A}^\varepsilon, \lambda'(t) \text{ exists}\}\mathrm{d}t
\end{aligned}$$

$$
\begin{aligned}
&- \int_0^1 (\lambda'(s) - 1)\mathbf{1}\{g(\lambda(s)) \in \tilde{A}^{\varepsilon}, \lambda'(s) \text{ exists}\}\mathrm{d}s,\\
&\leq \gamma + \int_0^1 \mathbf{1}\{f(t) \in \tilde{A}_{\varepsilon},\, g(\lambda(t)) \notin \tilde{A}^{\varepsilon}\}\mathrm{d}t\\
&\quad + \int_0^1 \|\lambda'(s) - 1\|\mathbf{1}\{\lambda'(s) \text{ exists}\}\mathrm{d}s.
\end{aligned}
$$

Here we have that

$$\int_0^1 \mathbf{1}\left\{f(t) \in \tilde{A}_{\varepsilon},\, g(\lambda(t)) \notin \tilde{A}^{\varepsilon}\right\}\mathrm{d}t \leq \mathbf{1}\left\{\|f - g \circ \lambda\| \geq 2\varepsilon\right\}.$$

In particular, given $f \in F_A$ and $\varepsilon > 0$, we can choose δ sufficiently small so that any g with $\rho^{\circ}_S(f, g) < \delta$ has a λ for which, by Lemma A.1.3, $\int_0^1 \|\lambda'(s)-1\|\mathbf{1}\{\lambda'(s) \text{ exists}\}\mathrm{d}s \leq c(\lambda) < \varepsilon$ and $\|f - g \circ \lambda\| < \varepsilon$. Hence, since $\gamma > 0$ and $\varepsilon > 0$ were arbitrary, $|\varpi_A(f) - \varpi_A(g)| \to 0$ as $\rho^{\circ}_S(f, g) \to 0$. □

Proof of Theorem 3.1.14. Fix $A \subseteq \mathbb{S}^{d-1}$ with $\mu_{d-1}(\partial A) = 0$. Then we need to show that F_A has measure 1 under the law of $\Sigma^{1/2}b_d$.

First, note that $\mu_{d-1}(\partial A) = 0$ implies $\mu_{d-1}(\{\mathbf{0}\} \cup \partial A) = 0$ and $\mu_d(\{\mathbf{0}\} \cup \partial\tilde{A}) = 0$. Then an application of Fubini's theorem [Dur10, p. 37] and using the fact that the expectation of an indicator function is a probability, we get,

$$\mathbb{E}\int_0^1 \mathbf{1}\{\hat{b}^{\Sigma}_d(t) \in (\{\mathbf{0}\} \cup \partial A)\}\mathrm{d}t = \int_0^1 \mathbb{P}(\Sigma^{1/2}b_d(t) \in (\{\mathbf{0}\} \cup \partial\tilde{A}))\mathrm{d}t = 0,$$

where the last equality follows from the fact that $\mathbb{P}(\Sigma^{1/2}b_d(t) \in (\{\mathbf{0}\} \cup \partial\tilde{A})) = \mathbb{P}(X \in (\{\mathbf{0}\} \cup \partial\tilde{A}))$ where $X \sim \mathcal{N}(0, \Sigma t)$, the d-dimensional Normal distribution.

Now, since $\int_0^1 \mathbf{1}\{\hat{b}^{\Sigma}_d(t) \in (\{\mathbf{0}\} \cup \partial A)\}\mathrm{d}t \geq 0$, it follows (see [Wil91, p. 51]) that

$$\int_0^1 \mathbf{1}\{\hat{b}^{\Sigma}_d(t) \in (\{\mathbf{0}\} \cup \partial A)\}\mathrm{d}t = 0 \text{ a.s.}$$

which gives that F_A has measure 1 under the law of $\Sigma^{1/2}b_d$. Therefore, using Donsker's theorem, Theorem 4.9(b), which states $Y'_n \Rightarrow \Sigma^{1/2}b_d$, the continuous mapping theorem, Theorem 4.5, and Lemma 3.1.15, we have $\pi_n(A) = \varpi_A(Y'_n) \xrightarrow{\mathrm{d}} \varpi_A(\Sigma^{1/2}b_d)$. □

In particular, we can use this result to determine that there is no almost sure limit for the proportion of time spent in any non-trivial set.

Corollary 3.1.16. *For any set* $A \subseteq \mathbb{S}^{d-1}$ *with* $0 < \mu_{d-1}(A) < \mu_{d-1}(\mathbb{S}^{d-1})$ *and* $\mu_{d-1}(\partial A) = 0$,

$$\liminf_{n\to\infty} \pi_n(A) = 0 \ a.s. \qquad \text{and} \qquad \limsup_{n\to\infty} \pi_n(A) = 1 \ a.s.$$

Proof. We will use the Hewitt-Savage zero-one law [Dur10, p. 180]. In order to do so, we need to show that $\limsup_{n\to\infty} \pi_n(A)$ and $\liminf_{n\to\infty} \pi_n(A)$ are exchangeable random variables. For this, note

$$\begin{aligned}\limsup_{n\to\infty} \pi_n(A) &= \limsup_{n\to\infty} \left(\frac{1}{n} \sum_{i=1}^{k} \mathbf{1}\{\hat{S}_i \in A\} + \frac{1}{n} \sum_{i=k+1}^{n} \mathbf{1}\{\hat{S}_i \in A\} \right) \\ &= \limsup_{n\to\infty} \frac{1}{n} \sum_{i=k+1}^{n} \mathbf{1}\{\hat{S}_i \in A\} \quad \text{a.s.}\end{aligned}$$

which clearly does not depend on the order of the first k increments, and since k was arbitrary, it is clearly exchangeable. The exact same argument is true for the lim inf as well.

Thus, it will be sufficient to show that $\mathbb{P}(\limsup_{n\to\infty} \pi_n(A) \geq 1-\varepsilon) > 0$ for any $\varepsilon > 0$, and $\mathbb{P}(\liminf_{n\to\infty} \pi_n(A) \leq \varepsilon) > 0$ for any $\varepsilon > 0$. For the former, note

$$\begin{aligned}\mathbb{P}(\limsup_{n\to\infty} \pi_n(A) \geq 1-\varepsilon) &\geq \mathbb{P}(\pi_n(A) > 1-\varepsilon \text{ i.o.}) \\ &\geq \mathbb{P}(\cap_{n=1}^{\infty} \cup_{m\geq n} \{\pi_m(A) > 1-\varepsilon\}) \\ &= \lim_{n\to\infty} \mathbb{P}(\cup_{m\geq n}\{\pi_m(A) > 1-\varepsilon\}) \\ &\geq \lim_{n\to\infty} \mathbb{P}(\pi_n(A) > 1-\varepsilon). \end{aligned} \tag{3.1.4}$$

Then Theorem 3.1.14 states that $\pi_n(A) \xrightarrow{\text{d}} \int_0^1 \mathbf{1}\{\hat{b}_d^{\Sigma}(t) \in A\}\text{d}t$ so for all but countably many $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}(\pi_n(A) > 1-\varepsilon) = \mathbb{P}\left(\int_0^1 \mathbf{1}\{\hat{b}_d^{\Sigma}(t) \in A\}\text{d}t > 1-\varepsilon \right).$$

Now, recall $\text{int}\, A := A \setminus \partial A$ is the interior of A, which is an open set, see for example [Kel75, pp. 44–46]. By the assumptions $\mu_{d-1}(A) > 0$ and $\mu_{d-1}(\partial A) = 0$ it follows that $\mu_{d-1}(\text{int}\, A) > 0$ and so the interior is non-empty. Since the interior is an open, non-empty subset of A, it follows that there exists at least one ball, call it A_ε, with radius $\varepsilon > 0$ such that $A_\varepsilon \subseteq A$.

Then it is easy to see that there is positive probability that $\hat{b}_d^\Sigma(t)$ stays in A_ε for all $t \in [\varepsilon, 1]$ for any $\varepsilon > 0$ (allowing the path to move away from 0). Thus, combining this with (3.1.4), we have

$$\begin{aligned}\mathbb{P}\left(\limsup_{n\to\infty} \pi_n(A) \geq 1-\varepsilon\right) &\geq \lim_{n\to\infty} \mathbb{P}(\pi_n(A) \geq 1-\varepsilon) \\ &= \mathbb{P}\left(\int_0^1 \mathbf{1}\{\hat{b}_d^\Sigma(t) \in A\}\mathrm{d}t > 1-\varepsilon\right) > 0\end{aligned}$$

for any $\varepsilon > 0$, and the Hewitt-Savage zero-one law gives us $\limsup_{n\to\infty} \pi_n(A) = 1$ a.s. as required.

Finally, note that $\pi_n(A^c) \leq 1 - \pi_n(A)$ (the inequality is due to possible visits to 0) and since $\mu_{d-1}(A) + \mu_{d-1}(A^c) = \mu_{d-1}(\mathbb{S}^{d-1})$, we get $0 < \mu_{d-1}(A^c) < \mu_{d-1}(\mathbb{S}^{d-1})$. Also, $\partial A = \partial A^c$, see for example [Kel75, p. 46], so $\mu_{d-1}(\partial A^c) = \mu_{d-1}(\partial A) = 0$. Thus, the conditions of the previous calculation are in fact satisfied for A^c, so $\liminf_{n\to\infty} \pi_n(A) \leq 1 - \limsup_{n\to\infty} \pi_n(A^c) = 1 - 1 = 0$ a.s. which completes the proof. □

3.2 Random walk point set convergence

3.2.1 Hausdorff distance

In this section, we will turn to the set convergence of the random walk points which is our first step towards the results directly related to the convex hull, but first, we need to add to our previously described metric spaces with another space of sets on which this convergence should take place. The purpose of this subsection is to set this up.

Let $\mathfrak{S}_0^d$ denote the collection of bounded subsets of $\mathbb{R}^d$ containing $\mathbf{0}$. Let $\mathfrak{K}_0^d$ denote the set of compact subsets of $\mathbb{R}^d$ containing $\mathbf{0}$. The set of random walk points do not form a convex set, so we will work with these more general spaces in this section. Recall that $A^\varepsilon = \{\mathbf{x} \in \mathbb{R}^d : \rho_E(\mathbf{x}, A) \leq \varepsilon\}$ and $\rho_E(\mathbf{x}, A)$ is the distance between a point $\mathbf{x}$ and a set A. Recall also the definitions of the Hausdorff distance from (1.3.1) and (1.3.2) which can be applied to $\mathfrak{S}_0^d$ and $\mathfrak{K}_0^d$, however note that ρ_H is a metric on $\mathfrak{K}_0^d$ but on $\mathfrak{S}_0^d$, ρ_H is a only a pseudometric, since while the triangle inequality still holds, $\rho_H(A, B) = 0$ does not imply $A = B$ (e.g. take an open set A and take B to be its closure; see Lemma 3.2.2 below). Thus convergence must take place in $(\mathfrak{K}_0^d, \rho_H)$.

We need the following observations about the Hausdorff distance.

Lemma 3.2.1. *Consider functions $f, g \in \mathcal{M}_0^d$. Then $f[0,1], g[0,1] \in \mathfrak{S}_0^d$ and*

$$\rho_H(f[0,1], g[0,1]) \leq \rho_S(f,g) \leq \rho_\infty(f,g).$$

Proof. Recall Λ is the set of $\lambda : [0,1] \to [0,1]$ that are strictly increasing and surjective. Then by (1.3.1),

$$\begin{aligned}
\rho_H(f[0,1], g[0,1]) &= \sup_{t\in[0,1]} \rho_E(f(t), g[0,1]) \vee \sup_{t\in[0,1]} \rho_E(g(t), f[0,1]) \\
&= \sup_{t\in[0,1]} \inf_{s\in[0,1]} \|f(t) - g(s)\| \vee \sup_{t\in[0,1]} \inf_{s\in[0,1]} \|g(t) - f(s)\| \\
&= \sup_{t\in[0,1]} \inf_{s\in[0,1]} \|f(t) - g\circ\lambda(s)\| \vee \sup_{t\in[0,1]} \inf_{s\in[0,1]} \|g\circ\lambda(t) - f(s)\|,
\end{aligned}$$

for any $\lambda \in \Lambda'$. Using the fact that for any $h \in \mathcal{M}_0^d$ and any $t \in [0,1]$, $\inf_{s\in[0,1]} h(s) \leq h(t)$, we get

$$\rho_H(f[0,1], g[0,1]) \leq \sup_{t\in[0,1]} \|f(t) - g\circ\lambda(t)\|,$$

for any $\lambda \in \Lambda'$, and hence

$$\rho_H(f[0,1], g[0,1]) \leq \inf_{\lambda\in\Lambda'} \|f - g\circ\lambda\|_\infty.$$

It follows that $\rho_H(f[0,1], g[0,1]) \leq \rho_S(f,g)$, and Lemma 1.3.2 completes the proof. □

Note that if $f \in \mathcal{C}_0^d$ then $f[0,1]$ is the continuous image of a compact set, containing $f(0) = 0$, and hence $f[0,1] \in \mathfrak{K}_0^d$. Thus Lemma 3.2.1 shows that $f \mapsto f[0,1]$ is a continuous map from $(\mathcal{C}_0^d, \rho_\infty)$ to $(\mathfrak{K}_0^d, \rho_H)$. For $f \in \mathcal{D}_0^d$, we need to work instead with $\operatorname{cl} f[0,1]$. We need the following simple fact.

Lemma 3.2.2. *For any $A, B \in \mathfrak{S}_0^d$,*

$$\rho_H(\operatorname{cl} A, B) = \rho_H(A, B).$$

Proof. Clearly $A \subseteq \operatorname{cl} A$, so

$$\sup_{x\in\operatorname{cl} A} \rho_E(x, B) \geq \sup_{x\in A} \rho_E(x, B). \tag{3.2.1}$$

For any $z \in \operatorname{cl} A$, there exist $z_n \in A$ such that $z_n \to z$; by continuity, $\rho_E(z_n, B) \to \rho_E(z, B)$. Also, since $z_n \in A$, it is clear that $\rho_E(z_n, B) \leq \sup_{x\in A} \rho_E(x, B)$, which gives

$\rho_E(z, B) \leq \sup_{x \in A} \rho_E(x, B)$, and hence

$$\sup_{z \in \mathrm{cl}\, A} \rho_E(z, B) \leq \sup_{x \in A} \rho_E(x, B). \tag{3.2.2}$$

Combining (3.2.1) and (3.2.2) shows that $\sup_{x \in \mathrm{cl}\, A} \rho_E(x, B) = \sup_{x \in A} \rho_E(x, B)$.

Since $A \subseteq \mathrm{cl}\, A$ we have $\rho_E(y, \mathrm{cl}\, A) \leq \rho_E(y, A)$ for all $y \in B$. For any $z \in \mathrm{cl}\, A$, there exist $z_n \in A$ such that $z_n \to z$. Then

$$\rho_E(y, z) = \lim_{n \to \infty} \rho_E(y, z_n) \geq \rho_E(y, A),$$

so that for any $y \in B$,

$$\rho_E(y, \mathrm{cl}\, A) = \inf_{z \in \mathrm{cl}\, A} \rho_E(y, z) \geq \rho_E(y, A).$$

Hence $\rho_E(y, \mathrm{cl}\, A) = \rho_E(y, A)$ for any $y \in B$, so the result follows from (1.3.1). □

Combining the preceding two lemmas gives the following result, which shows that $f \mapsto \mathrm{cl}\, f[0,1]$ is a continuous map from $(\mathcal{D}_0^d, \rho_S)$ to $(\mathfrak{K}_0^d, \rho_H)$.

Corollary 3.2.3. *Consider functions $f, g \in \mathcal{D}_0^d$. Then $\mathrm{cl}\, f[0,1], \mathrm{cl}\, g[0,1] \in \mathfrak{K}_0^d$ and*

$$\rho_H(\mathrm{cl}\, f[0,1], \mathrm{cl}\, g[0,1]) \leq \rho_S(f, g) \leq \rho_\infty(f, g).$$

Proof. First note that $\{f(x) : x \in [0,1]\}$ is contained in the closed Euclidean ball centred at the origin with radius $\|f\|_\infty$, which is finite for $f \in \mathcal{D}_0^d$ [Bil99, p. 121]. Thus if $f, g \in \mathcal{D}_0^d$, then $f[0,1], g[0,1]$ are bounded, and hence their closures are compact. We use Lemma 3.2.2 twice to see $\rho_H(\mathrm{cl}\, f[0,1], \mathrm{cl}\, g[0,1]) = \rho_H(f[0,1], g[0,1])$, and the result then follows from Lemma 3.2.1. □

3.2.2 Point set convergence

Now we can present our limit theorems for the set $\{S_0, S_1, \ldots, S_n\}$. First we state a law of large numbers. Recall that $I_\mu(t) = \mu t$.

Theorem 3.2.4. *Suppose that we have a random walk as defined at $(\mathbf{W}_\mu)$. Then, as elements of $(\mathfrak{K}_0^d, \rho_H)$,*

$$n^{-1}\{S_0, S_1, \ldots, S_n\} \xrightarrow{\text{a.s.}} I_\mu[0,1].$$

Proof. The functional law of large numbers, Theorem 3.1.2(b), shows that $X_n' \xrightarrow{\text{a.s.}} I_\mu$ on $(\mathcal{D}_0^d, \rho_\infty)$. Corollary 3.2.3 shows that $f \mapsto \operatorname{cl} f[0,1]$ is continuous from $(\mathcal{D}_0^d, \rho_\infty)$ to $(\mathfrak{K}_0^d, \rho_H)$, so the mapping theorem, Theorem 3.1.4, shows that $\operatorname{cl} X_n'[0,1] \xrightarrow{\text{a.s.}} \operatorname{cl} I_\mu[0,1]$; note that $\operatorname{cl} X_n'[0,1] = X_n'[0,1] = n^{-1}\{S_0, S_1, \dots, S_n\}$ and $\operatorname{cl} I_\mu[0,1] = I_\mu[0,1]$. □

Theorem 3.2.5. *Suppose that we have a random walk as defined at* $(\mathbf{W}_\mu)$ *with* $\mu = \mathbf{0}$ *and satisfying* $(\mathbf{V})$*. Then, as elements of* $(\mathfrak{K}_0^d, \rho_H)$*,*

$$n^{-1/2}\{S_0, S_1, \dots, S_n\} \Rightarrow \Sigma^{1/2} b_d[0,1].$$

Proof. Donsker's theorem, Theorem 3.1.5(b), shows that $Y_n' \Rightarrow \Sigma^{1/2} b_d$ on $(\mathcal{D}_0^d, \rho_S)$. Corollary 3.2.3 shows that $f \mapsto \operatorname{cl} f[0,1]$ is continuous from $(\mathcal{D}_0^d, \rho_S)$ to $(\mathfrak{K}_0^d, \rho_H)$, so the mapping theorem, Theorem 3.1.6, shows that $\operatorname{cl} Y_n'[0,1] \Rightarrow \operatorname{cl} \Sigma^{1/2} b_d[0,1]$; note that $\operatorname{cl} Y_n'[0,1] = Y_n'[0,1] = n^{-1/2}\{S_0, S_1, \dots, S_n\}$ and $\operatorname{cl} \Sigma^{1/2} b_d[0,1] = \Sigma^{1/2} b_d[0,1]$. □

3.2.3 Diameter of random walks

As a first application of the results of this section (we see another application in Section 3.3), we consider the diameter of the random walk, as defined at (1.3.7).

The following is a generalisation to d-dimensions of the 2-dimensional almost-sure result contained in [MW18, Theorem 1.3].

Theorem 3.2.6. *Suppose that we have a random walk as defined at* $(\mathbf{W}_\mu)$*.*

(a) $n^{-1} D_n \xrightarrow{\text{a.s.}} \|\mu\|$ *as* $n \to \infty$*.*

(b) If $\mu = \mathbf{0}$ *and* $(\mathbf{V})$ *holds, then* $n^{-1/2} D_n \xrightarrow{\text{d}} \operatorname{diam}(\Sigma^{1/2} b_d[0,1])$ *as* $n \to \infty$*.*

The theorem rests on the following result, which shows that $A \mapsto \operatorname{diam} A$ is continuous from $(\mathfrak{K}_0^d, \rho_H)$ to $(\mathbb{R}_+, \rho_E)$ which can also be found at [MW18, Lemma 3.5].

Lemma 3.2.7. *For any* $A, B \in \mathfrak{S}_0^d$*,*

$$|\operatorname{diam} A - \operatorname{diam} B| \le 2\rho_H(A, B).$$

Proof. Let $\rho_H(A, B) = r$. From (1.3.2) we have that for any $\mathbf{x}_1, \mathbf{x}_2 \in A$ and any $s > r$, there exist $\mathbf{y}_1, \mathbf{y}_2 \in B$ such that $\rho_E(\mathbf{x}_i, \mathbf{y}_i) \leq s$. Then,

$$\rho_E(\mathbf{x}_1, \mathbf{x}_2) \leq \rho_E(\mathbf{x}_1, \mathbf{y}_1) + \rho_E(\mathbf{y}_1, \mathbf{y}_2) + \rho_E(\mathbf{y}_2, \mathbf{x}_2) \leq 2s + \operatorname{diam} B.$$

Hence

$$\operatorname{diam} A = \sup_{\mathbf{x}_1, \mathbf{x}_2 \in A} \rho_E(\mathbf{x}_1, \mathbf{x}_2) \leq 2s + \operatorname{diam} B,$$

and since $s > r$ was arbitrary we get $\operatorname{diam} A - \operatorname{diam} B \leq 2r$. Similarly, $\operatorname{diam} B - \operatorname{diam} A \leq 2r$, giving the result. □

Proof of Theorem 3.2.6. For part (a), we have from the law of large numbers for sets, Theorem 3.2.4, that $n^{-1}\{S_0, S_1, \dots, S_n\} \xrightarrow{\text{a.s.}} I_\mu[0, 1]$ on $(\mathfrak{K}_0^d, \rho_H)$, while Lemma 3.2.7 shows that $A \mapsto \operatorname{diam} A$ is continuous from $(\mathfrak{K}_0^d, \rho_H)$ to $(\mathbb{R}_+, \rho_E)$. Thus the mapping theorem, Theorem 3.1.4, yields $n^{-1} D_n \xrightarrow{\text{a.s.}} \operatorname{diam}(I_\mu[0, 1]) = \|\mu\|$.

For part (b), we have from the central limit theorem for sets, Theorem 3.2.5, that $n^{-1/2}\, \{S_0, S_1, \dots, S_n\} \Rightarrow \Sigma^{1/2} b_d[0, 1]$ on $(\mathfrak{K}_0^d, \rho_H)$. Lemma 3.2.7 together with the mapping theorem, Theorem 3.1.6, yield the result. □

3.3 Convergence of convex hulls

3.3.1 Trajectories and hulls

We use the notation for sets of subsets of $\mathbb{R}^d$ and for the Hausdorff distance ρ_H from Section 3.2.1 and equations (1.3.1) and (1.3.2). We need the following result.

Lemma 3.3.1. *For any $A, B \in \mathfrak{S}_0^d$,*

$$\rho_H(\operatorname{hull} A, \operatorname{hull} B) \leq \rho_H(A, B).$$

Proof. Note Carathéodory's theorem: for any $\mathbf{x} \in \operatorname{hull} A$ there exist finitely many points $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n \in A$ and $\lambda_1, \lambda_2, \dots, \lambda_n$ with $\lambda_i \geq 0$, $\sum_{i=1}^n \lambda_i = 1$, for which $\mathbf{x} = \sum_{i=1}^n \lambda_i \mathbf{x}_i$ (see e.g. [Gru07, p. 42]). Let $r := \rho_H(A, B)$. For any $s > r$, we have from (1.3.2) that for each $\mathbf{x}_i \in A$ there exists $\mathbf{y}_i \in B$ such that $\rho_E(\mathbf{x}_i, \mathbf{y}_i) \leq s$. Now consider $\mathbf{y} = \sum_{i=1}^{n} \lambda_i \mathbf{y}_i \in \operatorname{hull} B$. Then

$$\rho_E(\mathbf{x}, \mathbf{y}) \leq \sum_{i=1}^{n} \lambda_i \rho_E(\mathbf{x}_i, \mathbf{y}_i) \leq s.$$

This calculation implies that $\operatorname{hull} A \subseteq (\operatorname{hull} B)^s$, and by a similar argument we get $\operatorname{hull} B \subseteq (\operatorname{hull} A)^s$. With (1.3.2) we get $\rho_H(\operatorname{hull} A, \operatorname{hull} B) \leq s$. Since $s > r$ was arbitrary, the result follows. □

Let $\mathfrak{C}_0^d$ denote the set convex compact subsets of $\mathbb{R}^d$ containing $\mathbf{0}$. For $A \in \mathfrak{C}_0^d$, we define the *support function* of A by

$$h_A(\mathbf{x}) := \sup_{\mathbf{y} \in A} (\mathbf{x} \cdot \mathbf{y}), \text{ for any } \mathbf{x} \in \mathbb{R}^d. \tag{3.3.1}$$

Then for $A, B \in \mathfrak{C}_0^d$ we have another equivalent description of $\rho_H(A, B)$ (see e.g. [Gru07, p. 84]):

$$\rho_H(A, B) = \sup_{\mathbf{e} \in \mathbb{S}^{d-1}} |h_A(\mathbf{e}) - h_B(\mathbf{e})|. \tag{3.3.2}$$

Given $f \in \mathcal{D}_0^d$, we have $\operatorname{cl} f[0,1]$ is compact and contains $\mathbf{0} = f(0)$. A theorem of Carathéodory [Gru07, p. 44] says that if A is compact then so is $\operatorname{hull} A$; hence $\operatorname{hull} \operatorname{cl} f[0,1]$ is compact. Moreover, we have that $\operatorname{hull} \operatorname{cl} A = \operatorname{cl} \operatorname{hull} A$ [Gru07, p. 45]. Hence if $f \in \mathcal{D}_0^d$ then $\operatorname{cl} \operatorname{hull} f[0,1] \in \mathfrak{C}_0^d$. Of course, if $f \in \mathcal{C}_0^d$ then $f[0,1]$ and hence $\operatorname{hull} f[0,1]$ is already compact. The following result shows that $f \mapsto \operatorname{cl} \operatorname{hull} f[0,1]$ is a continuous map from $(\mathcal{D}_0^d, \rho_S)$ to $(\mathfrak{C}_0^d, \rho_H)$. This fact is also found as Lemma 5.1 in the recent paper of Molchanov and Wespi [MW16].

Lemma 3.3.2. *Consider two functions $f, g \in \mathcal{M}_0^d$. Then,*

$$\rho_H\left(\operatorname{cl} \operatorname{hull} f[0,1], \operatorname{cl} \operatorname{hull} g[0,1]\right) \leq \rho_S(f, g).$$

Proof. First, Lemma 3.2.2 (twice) and Lemma 3.3.1 yield

$$\rho_H\left(\operatorname{cl} \operatorname{hull} f[0,1], \operatorname{cl} \operatorname{hull} g[0,1]\right) = \rho_H\left(\operatorname{hull} f[0,1], \operatorname{hull} g[0,1]\right) \leq \rho_H(f[0,1], g[0,1]).$$

Lemma 3.2.1 completes the proof. □

3.3.2 Limit theorems for convex hulls

The following is our law of large numbers for the convex hull.

Theorem 3.3.3. *Suppose that we have a random walk as defined at* $(\mathbf{W}_\mu)$*. Then, as elements of* $(\mathfrak{C}_0^d, \rho_H)$*,*

$$n^{-1} \operatorname{hull}\{S_0, \dots, S_n\} \xrightarrow{\text{a.s.}} I_\mu[0,1].$$

Proof. Theorem 3.2.4 states $n^{-1}\{S_0, \dots, S_n\} \xrightarrow{\text{a.s.}} I_\mu[0,1]$ on $(\mathfrak{K}_0^d, \rho_H)$. Lemma 3.3.1 shows that $A \mapsto \operatorname{hull} A$ is a continuous map from $(\mathfrak{K}_0^d, \rho_H)$ to $(\mathfrak{C}_0^d, \rho_H)$, so the mapping theorem, Theorem 3.1.4, implies that $\operatorname{hull} n^{-1}\{S_0, \dots, S_n\} \xrightarrow{\text{a.s.}} \operatorname{hull} I_\mu[0,1]$. Here $\operatorname{hull} I_\mu[0,1] = I_\mu[0,1]$, and, since the convex hull is preserved under scaling, $\operatorname{hull} n^{-1}\{S_0, \dots, S_n\} = n^{-1} \operatorname{hull}\{S_0, \dots, S_n\}$. □

Next we state the accompanying central limit theorem. Let $h_d := \operatorname{hull} b_d[0,1]$, the convex hull of d-dimensional Brownian motion run for unit time.

Theorem 3.3.4. *Suppose that we have a random walk as defined at* $(\mathbf{W}_\mu)$ *with* $\mu = \mathbf{0}$ *and satisfying* $(\mathbf{V})$*. Then, as elements of* $(\mathfrak{C}_0^d, \rho_H)$*,*

$$n^{-1/2} \operatorname{hull}\{S_0, \dots, S_n\} \Rightarrow \Sigma^{1/2} h_d.$$

Proof. Theorem 3.2.5 states $n^{-1/2}\{S_0, \dots, S_n\} \Rightarrow \Sigma^{1/2} b_d[0,1]$ on $(\mathfrak{K}_0^d, \rho_H)$. Lemma 3.3.1 shows that $A \mapsto \operatorname{hull} A$ is a continuous map from $(\mathfrak{K}_0^d, \rho_H)$ to $(\mathfrak{C}_0^d, \rho_H)$, so the mapping theorem, Theorem 3.1.6, implies that $\operatorname{hull} n^{-1/2}\{S_0, \dots, S_n\} \Rightarrow \operatorname{hull} \Sigma^{1/2} b_d[0,1]$. Since the convex hull is preserved under affine transformations, $\operatorname{hull} \Sigma^{1/2} b_d[0,1] = \Sigma^{1/2} \operatorname{hull} b_d[0,1]$. □

Remark 3.3.5. Alternatively, we could obtain Theorems 3.3.3 and 3.3.4 directly from the functional law of large numbers, Theorem 3.1.2, and Donsker's theorem, Theorem 3.1.5, using Lemma 3.3.2.

Suppose now $d \geq 2$. To obtain second-order results in the case where $\mu \neq \mathbf{0}$, an additional scaling limit is required. Let $\{\mathbf{e}_1, \dots, \mathbf{e}_d\}$ be the standard orthonormal basis of $\mathbb{R}^d$, and supposing that $\mu \neq \mathbf{0}$, let $\{\mathbf{u}_1, \dots, \mathbf{u}_d\}$ be another orthonormal basis of $\mathbb{R}^d$ with $\mathbf{u}_1 = \hat{\mu}$. Then we transform Z into Z', a vector with components in the standard basis, by taking

$$Z' = (Z'_1, Z'_2, \dots, Z'_d) := (Z \cdot \mathbf{u}_1, Z \cdot \mathbf{u}_2, \dots, Z \cdot \mathbf{u}_d),$$

and consider $Z'_\perp := (Z'_2, \ldots, Z'_d)$. Note that, since $\mathbb{E}\, Z \cdot \mathbf{u}_k = \mu \cdot \mathbf{u}_k = 0$ for $k \neq 1$, we have $\mathbb{E}\, Z'_\perp = 0$. Then set

$$\Sigma_{\mu_\perp} := \mathbb{E}[Z'_\perp (Z'_\perp)^\top]. \tag{3.3.3}$$

This defines a $(d-1)$-dimensional covariance matrix, describing the covariances of the process projected onto the hyperplane orthogonal to the mean vector. Note that $\Sigma_{\mu_\perp}$ is non-negative definite and hence it has a unique non-negative definite symmetric square root matrix $\Sigma_{\mu_\perp}^{1/2}$. It will be useful to have notation for $\Sigma_{\mu_\perp}^{1/2}$ extended back to a d-dimensional matrix which we will denote as $\tilde{\Sigma}_{\mu_\perp}^{1/2}$, specifically we define

$$\tilde{\Sigma}_{\mu_\perp}^{1/2} := \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & \Sigma_{\mu_\perp}^{1/2} & \\ 0 & & & \end{pmatrix}. \tag{3.3.4}$$

We will need a new weak convergence result and as we took a mapping of the increments above, we need to define a different mapping for the walk process itself, for which we use a d-dimensional analogue of that used in [WX15b]. Namely, for $n \in \mathbb{N}$, define $\psi_{n,\mu} : \mathbb{R}^d \to \mathbb{R}^d$ by the image of $\mathbf{x} \in \mathbb{R}^d$ in Cartesian components:

$$\psi_{n,\mu}(\mathbf{x}) = \left(\frac{\mathbf{x} \cdot \mathbf{u}_1}{n\|\mu\|}, \frac{\mathbf{x} \cdot \mathbf{u}_2}{\sqrt{n}}, \ldots, \frac{\mathbf{x} \cdot \mathbf{u}_d}{\sqrt{n}} \right),$$

where $\{\mathbf{u}_1, \ldots, \mathbf{u}_d\}$ is the orthonormal basis defined above. We extend this, and subsequent similar notation, to sets in the usual way, $\psi_{n,\mu}(A) = \{\psi_{n,\mu}(\mathbf{x}) : \mathbf{x} \in A\}$. This mapping has an effect which is the natural extension of its 2-dimensional equivalent, rotating $\mathbb{R}^d$ mapping $\hat{\mu}$ to the unit vector in the horizontal direction, and scaling space with a horizontal shrinking factor of $\|\mu\| n$, but now also a factor of $\sqrt{n}$ in all $d-1$ directions orthogonal to the horizontal.

We will also need some notation for the first component of the mapping, and the $d-1$ vector containing the elements orthogonal to the mean, so we define the following:

$$\psi^1_{n,\mu}(\mathbf{x}) := \frac{\mathbf{x} \cdot \mathbf{u}_1}{n\|\mu\|} \qquad \text{and} \qquad \psi^\perp_{n,\mu}(\mathbf{x}) := \left(\frac{\mathbf{x} \cdot \mathbf{u}_2}{\sqrt{n}}, \ldots, \frac{\mathbf{x} \cdot \mathbf{u}_d}{\sqrt{n}} \right).$$

Naturally, we also need to define a new limiting process which combines the drift with

Brownian motion in a time-space way. We denote this $\tilde{b}_d(t)$, which is defined as

$$\tilde{b}_d(t) = (t, b_{d-1}(t)), \text{ for } t \in [0,1], \tag{3.3.5}$$

where we use the notation b_{d-1} to be clear that we mean $(d-1)$-dimensional Brownian motion. We use the notation $\tilde{h}_d^{\Sigma} := \operatorname{hull} \tilde{\Sigma}_{\mu_\perp}^{1/2} \tilde{b}_d[0,1]$, the hull of $\tilde{\Sigma}_{\mu_\perp}^{1/2} \tilde{b}_d$ run for unit time.

Lemma 3.3.6. *Suppose that we have a random walk as defined at* $(\mathbf{W}_\mu)$ *with* $\mu \neq \mathbf{0}$ *and satisfying* $(\mathbf{V})$*. Then, as* $n \to \infty$*, as elements of* $(\mathfrak{C}_0^d, \rho_H)$*,*

$$\psi_{n,\mu}(\operatorname{hull}\{S_0, S_1, \dots, S_n\}) \Rightarrow \tilde{h}_d^{\Sigma}.$$

Proof. First, note that, since $\psi_{n,\mu}$ is an affine transformation, we have

$$\psi_{n,\mu}(\operatorname{hull}\{S_0, \dots, S_n\}) = \operatorname{hull}\left(\psi_{n,\mu}(\{S_0, \dots, S_n\})\right).$$

Noting that $A \mapsto \operatorname{hull} A$ is continuous from $(\mathfrak{K}_0^d, \rho_H)$ to $(\mathfrak{C}_0^d, \rho_H)$ by Lemma 3.3.1, the continuous mapping theorem, Theorem 3.1.6, means it is sufficient to show

$$\psi_{n,\mu}(\{S_0, \dots, S_n\}) \Rightarrow \tilde{\Sigma}_{\mu_\perp}^{1/2} \tilde{b}_d[0,1] \text{ on } (\mathfrak{K}_0^d, \rho_H). \tag{3.3.6}$$

In order to show this, we first define a new unscaled trajectory as $W_n'(t) := S_{\lfloor nt \rfloor}$. Then we will show that,

$$\psi_{n,\mu}(W_n') \Rightarrow \tilde{\Sigma}_{\mu_\perp}^{1/2} \tilde{b}_d, \text{ on } (\mathcal{D}_0^d, \rho_S). \tag{3.3.7}$$

First, recall Theorem 1.3.13: if X_n, Y_n, and X are elements of a metric space (S, ρ), such that $X_n \Rightarrow X$ and $\rho(X_n, Y_n) \xrightarrow{\text{P}} 0$, then $Y_n \Rightarrow X$. Taking $X_n = (I, \psi_{n,\mu}^{\perp}(W_n'))$ where we recall I is the identity map on $[0,1]$, $Y_n = \psi_{n,\mu}(W_n')$ and $X = \tilde{\Sigma}_{\mu_\perp}^{1/2} \tilde{b}_d$, all elements of $(\mathcal{D}_0^d, \rho_S)$ it suffices to show that

$$\rho_S(\psi_{n,\mu}(W_n'), (I, \psi_{n,\mu}^{\perp}(W_n'))) \xrightarrow{\text{P}} 0, \tag{3.3.8}$$

and

$$(I, \psi_{n,\mu}^{\perp}(W_n')) \Rightarrow \tilde{\Sigma}_{\mu_\perp}^{1/2} \tilde{b}_d(t), \text{ on } (\mathcal{D}_0^d, \rho_S). \tag{3.3.9}$$

To prove (3.3.8), notice that $\psi_{n,\mu}^{1}(W_n')$ is the piecewise constant trajectory of a one-dimensional walk with $\|\mu\| > 0$ now normalised by $\|\mu\|^{-1} n^{-1}$, so Theorem 3.1.2 applies and we have

$$\lim_{n\to\infty} \psi^1_{n,\mu}(W'_n) = I \quad \text{a.s.} \tag{3.3.10}$$

Using Lemma 1.3.2 it becomes a simple exercise to see that, for $f \in \mathcal{C}_0^{d-1}$ and $g, h \in \mathcal{C}_0$ we have $\rho_S((f,g),(f,h)) \leq \rho_\infty((f,g),(f,h)) = \rho_\infty(g,h)$, which shows that (3.3.10) implies (3.3.8).

For (3.3.9), note $\psi^\perp_{n,\mu} W'_n$ is the piecewise constant trajectory of a $(d-1)$-dimensional walk with $\mu = \mathbf{0}$, normalised by $n^{-1/2}$ so Theorem 3.1.5 gives

$$\psi^\perp_{n,\mu}(W'_n) \Rightarrow \Sigma^{1/2}_{\mu_\perp} b_{d-1} \text{ on } (\mathcal{D}_0^{d-1}, \rho_S).$$

This implies that, for all bounded, continuous $f : \mathcal{D}_0^{d-1} \to \mathbb{R}$,

$$\mathbb{E}[f(\psi^\perp_{n,\mu}(W'_n))] \to \mathbb{E}[f(\Sigma^{1/2}_{\mu_\perp} b_{d-1})], \text{ as } n \to \infty. \tag{3.3.11}$$

Now consider $\mathbb{E}[g(I, \psi^\perp_{n,\mu}(W'_n)]$ for any bounded, continuous $g : \mathcal{D}_0^d \to \mathbb{R}$. Then, since I is a non-random function, there exists a function f, defined such that $f(\cdot) = g(I, \cdot)$ which is itself bounded and continuous on $\mathcal{D}_0^{d-1}$. By (3.3.11), it follows that

$$\mathbb{E}[g(I, \psi^\perp_{n,\mu}(W'_n))] = \mathbb{E}[f(\psi^\perp_{n,\mu}(W'_n))] \to \mathbb{E}[f(\Sigma^{1/2}_{\mu_\perp} b_{d-1})] = \mathbb{E}[g(I, \Sigma^{1/2}_{\mu_\perp} b_{d-1})],$$

and noting $g(I, \Sigma^{1/2}_{\mu_\perp} b_{d-1}) = g(\tilde{\Sigma}^{1/2}_{\mu_\perp} \tilde{b}_d)$, we have proven (3.3.9) and hence (3.3.7).

The final step is to notice that Corollary 3.2.3 shows that $f \mapsto \operatorname{cl} f[0,1]$ is continuous from $(\mathcal{D}_0^d, \rho_S)$ to $(\mathfrak{K}_0^d, \rho_H)$, so the mapping theorem, Theorem 3.1.6, with (3.3.7) shows that $\operatorname{cl} \psi_{n,\mu}(W_n[0,1]) \Rightarrow \operatorname{cl} \tilde{\Sigma}^{1/2}_{\mu_\perp} \tilde{b}_d[0,1]$. Observing that $\operatorname{cl} \psi_{n,\mu}(W_n[0,1]) = \psi_{n,\mu}(\{S_0, \ldots, S_n\})$ and $\operatorname{cl} \tilde{\Sigma}^{1/2}_{\mu_\perp} \tilde{b}_d[0,1] = \tilde{\Sigma}^{1/2}_{\mu_\perp} \tilde{b}_d[0,1]$, we have proven (3.3.6) and so the proof is complete. □

3.3.3 Applications to functionals of convex hulls

We consider three functionals defined on non-empty convex compact sets. First, let $\mathcal{W} : \mathfrak{C}_0^d \to \mathbb{R}_+$ denote the *mean width* defined by

$$\mathcal{W}(A) := \int_{\mathbb{S}^{d-1}} h_A(\mathbf{e}) \mathrm{d}\mathbf{e},$$

where h_A is the support function of A as defined at (3.3.1). Define the *volume* functional by

$$\mathcal{V}(A) := \mu_d(A),$$

the d-dimensional Lebesgue measure of A. Also we follow Gruber [Gru07, p. 104] and define the *surface area* functional by

$$\mathcal{S}(A) := \lim_{\lambda \downarrow 0} \left(\frac{\mathcal{V}(A^\lambda) - \mathcal{V}(A)}{\lambda} \right);$$

which was a definition originally suggested by Minkowski; the limit exists by the Steiner formula of integral geometry [Gru07, Theorem 6.6] which states, for $A \in \mathfrak{C}^d$,

$$\mathcal{V}(A^\lambda) = \mu_d(A^\lambda) = \sum_{i=0}^{d} \binom{d}{i} Q_i(A)\lambda^i, \tag{3.3.12}$$

where $\binom{x}{y}$ is the binomial coefficient with the convention $\binom{x}{0} = 1$, and $Q_i(A)$ are the quermassintegrals of A.

For the random walk, we use the notation

$$\mathcal{W}_n := \mathcal{W}(\text{hull}\{S_0, \dots, S_n\}); \quad \mathcal{V}_n := \mathcal{V}(\text{hull}\{S_0, \dots, S_n\}); \quad \mathcal{S}_n := \mathcal{S}(\text{hull}\{S_0, \dots, S_n\}).$$

We first investigate basic continuity properties of these functionals. We define the Euler gamma function by

$$\Gamma(t) := \int_0^\infty x^{t-1} \mathrm{e}^{-x} \mathrm{d}x, \text{ for } t > 0.$$

Lemma 3.3.7. *Suppose that $A, B \in \mathfrak{C}_0^d$. Then*

$$\rho_E(\mathcal{W}(A), \mathcal{W}(B)) \leq 2\rho_H(A, B); \tag{3.3.13}$$

$$\rho_E(\mathcal{S}(A), \mathcal{S}(B)) \leq 2(d-1)(\text{diam}(B) + \rho_H(A, B))^{d-2} \rho_H(A, B); \tag{3.3.14}$$

$$\begin{aligned} \rho_E(\mathcal{V}(A), \mathcal{V}(B)) \leq{}& 2\pi^{d-1} \rho_H(A, B)^d \\ &+ \max_{S \in \{A, B\}} \left(\mathcal{S}(S) + \sum_{i=2}^{d-1} 2\pi \max\{\text{diam}(S), 1\}^d \rho_H(A, B)^{i-1} \right) \\ &\cdot \rho_H(A, B) . \end{aligned} \tag{3.3.15}$$

Before we complete the proofs of these inequalities we note Cauchy's surface area formula and a further geometric lemma. Recall that if ν_d is the volume of the unit ball in d-dimensions, then Cauchy's surface area formula [Gru07, p. 106] states that for $A \in \mathfrak{C}^d$,

$$\mathcal{S}(A) = \frac{1}{\nu_{d-1}} \int_{\mathbb{S}^{d-1}} \mu_{d-1}(A|\mathbf{u}^{\perp})\mathrm{d}\mathbf{u},$$

where $A|\mathbf{u}^{\perp}$ denotes the projection of A onto the $(d-1)$-dimensional subspace of $\mathbb{R}^d$ perpendicular to $\mathbf{u}$.

Remark 3.3.8. In $d = 2$ Cauchy's formula says $\mathcal{S}(A) = \mathcal{W}(A)$.

The geometric lemma is a bound on the Lebesgue measure of the difference in volume of two convex sets.

Lemma 3.3.9. *Consider two sets $S_1, S_2, \in \mathfrak{C}_0^d$ with $\rho_H(S_1, S_2) = r$, then*

$$\mu_d(S_1 \backslash S_2) \leq \frac{2\pi^{d/2}(\mathrm{diam}(S_2) + r)^{d-1}}{\Gamma(\frac{d}{2})} \cdot r.$$

Proof. First we recall (3.3.12) and note that $Q_0(S) = \mu_d(S)$; for a comprehensive discussion on quermassintegrals see [Gru07, Ch. 6]. We also note one further result of Steiner, see [Gru07, Theorem 6.14] which states, for $S \in \mathfrak{C}^d$,

$$\mathcal{S}(S^{\lambda}) = d \sum_{i=0}^{d-1} \binom{d-1}{i} Q_{i+1}(S)\lambda^i = \sum_{i=1}^{d} i \binom{d}{i} Q_i(S)\lambda^{i-1}.$$

It is a simple exercise by comparison of terms in the summations and use of the fact $Q_0(S) = \mu_d(S)$ to see

$$\mu_d(S^{\lambda}) - \mu_d(S) = \lambda \sum_{i=1}^{d} \binom{d}{i} Q_i(S)\lambda^{i-1} \leq \lambda \mathcal{S}(S^{\lambda}). \tag{3.3.16}$$

Now, if $\rho_H(S_1, S_2) = r$, for any $s > r$, $S_1 \subseteq S_2^s$, so $S_1 \setminus S_2 \subseteq S_2^s \setminus S_2$. It follows from (3.3.16),

$$\mu_d(S_1 \setminus S_2) \leq \mu_d(S_2^s \setminus S_2) = \mu_d(S_2^s) - \mu_d(S_2) \leq s\mathcal{S}(S_2^s). \tag{3.3.17}$$

Now, recall $\mathbb{B}^d$ is the d-dimensional unit ball. Then notice that it follows from Cauchy's formula that for convex sets A and B such that $A \subseteq B$, $\mathcal{S}(A) \leq \mathcal{S}(B)$, so, because $S_2^s \subseteq (\mathrm{diam}(S_2) + s)\mathbb{B}^d$, we have

$$s\mathcal{S}(S_2^s) \leq s\mathcal{S}((\mathrm{diam}(S_2) + s)\mathbb{B}^d). \tag{3.3.18}$$

Since $s > r$ was arbitrary, the statement of the lemma follows from (3.3.17), (3.3.18) and the surface area formula for $\mathbb{B}^d$, see for example [Som58, p. 136]. □

Now we turn to the proof of Lemma 3.3.7.

Proof of Lemma 3.3.7. We first prove (3.3.13). By Cauchy's formula and the triangle inequality,

$$\begin{aligned}|\mathcal{W}(A) - \mathcal{W}(B)| &= \left|\int_{\mathbb{S}^{d-1}} (h_A(\mathbf{e}) - h_B(\mathbf{e}))\mathrm{d}\mathbf{e}\right| \\ &\leq \int_{\mathbb{S}^{d-1}} \sup_{\mathbf{e}\in\mathbb{S}^{d-1}} |h_A(\mathbf{e}) - h_B(\mathbf{e})|\mathrm{d}\mathbf{e} \\ &= 2 \sup_{\mathbf{e}\in\mathbb{S}^{d-1}} |h_A(\mathbf{e}) - h_B(\mathbf{e})|.\end{aligned}$$

This equation with (3.3.2) gives (3.3.13).

Next we consider (3.3.14). Suppose, without loss of generality, $\mathcal{S}(A) \geq \mathcal{S}(B)$. In this next calculation, we use Cauchy's surface area formula and crudely replace the difference in integrands, which should be a difference is $(d-1)$-dimensional measures of sets, by the $(d-1)$-dimensional measure of the points in the larger but not smaller set (ignoring the offset of the measure of the points in the smaller but not larger set). This calculation, using the volume of $\mathbb{B}^d$ formula, see [Som58, p. 136],

$$\begin{aligned}\rho_E(\mathcal{S}(A), \mathcal{S}(B)) &= \frac{\Gamma(\frac{d+1}{2})}{\pi^{(d-1)/2}} \int_{\mathbb{S}^{d-1}} \left(\mu_{d-1}(A|\mathbf{u}^\perp) - \mu_{d-1}(B|\mathbf{u}^\perp)\right) \mathrm{d}\mathbf{u} \\ &\leq \frac{\Gamma(\frac{d+1}{2})}{\pi^{(d-1)/2}} \int_{\mathbb{S}^{d-1}} \mu_{d-1}(A|\mathbf{u}^\perp \setminus B|\mathbf{u}^\perp)\mathrm{d}\mathbf{u}.\end{aligned}$$

Now, noting that for a set B, $\mathrm{diam}(B|\mathbf{u}^\perp) \leq \mathrm{diam}(B)$, and that $\rho_H(A|\mathbf{u}^\perp, B|\mathbf{u}^\perp) \leq \rho_H(A, B) = r$ we can apply Lemma 3.3.9 to get,

$$\begin{aligned}\rho_E(\mathcal{S}(A), \mathcal{S}(B)) &\leq \frac{\Gamma(\frac{d+1}{2})}{\pi^{(d-1)/2}} \int_{\mathbb{S}^{d-1}} \frac{2\pi^{(d-1)/2}(\mathrm{diam}(B) + r)^{d-2} \cdot r}{\Gamma(\frac{d-1}{2})}\mathrm{d}\mathbf{u} \\ &= \frac{\Gamma(\frac{d+1}{2})}{\pi^{(d-1)/2}} \cdot 2 \left(\frac{2\pi^{(d-1)/2}(\mathrm{diam}(B) + r)^{d-2} \cdot r}{\Gamma(\frac{d-1}{2})}\right) \\ &= 2(d-1)(\mathrm{diam}(B) + r)^{d-2} \cdot r.\end{aligned}$$

and the result follows.

And finally, we consider (3.3.15). Set $r = \rho_H(A, B)$. Then, by (1.3.2), $A \subseteq B^s$ for any $s > r$. Also note, if $Q_i(B)$ are the quermassintegrals of B then $Q_0(B) = \mathcal{V}(B)$,

$Q_1(B) = \mathcal{S}(B)$ and $Q_d(B) = \mathcal{V}(\mathbb{B}^d) \leq 2\pi^{d-1}$. Hence,

$$\begin{aligned}\mathcal{V}(A) &\leq \mathcal{V}(B^s) \\ &\leq \mathcal{V}(B) + \mathcal{S}(B)s + 2\pi^{d-1}s^d + \sum_{i=2}^{d-1} Q_i(B)s^i \ ,\end{aligned}$$

by the Steiner formula (3.3.12). However, as discussed at [Gru07, p. 109] the quermass-integrals can be expressed as the mean of the $(d-i)$-dimensional volumes of the projections of the set B into $(d-i)$-dimensional subspaces. Thus using the very loose bound that the d-dimensional volume of the sphere with radius r is $\mathcal{V}(r\mathbb{B}^d) \leq 2\pi^{d-1}r^d$, we can establish the crude bound $Q_i \leq 2\pi^{d-1}(\max\{\operatorname{diam} B, 1\})^d$ for all $i \in \{2, \ldots, d-1\}$ and so each Q_i is finite because B is compact (assume d fixed). By symmetry we can get a similar inequality starting from $\mathcal{V}(B)$ and since $s > r$ was arbitrary, (3.3.15) follows. □

So now we have the weak convergence result, continuity of the relevant functionals and the mapping theorem, we can return to the weak convergence of the functionals. The 2-dimensional statements for the surface area and volume were previously studied in [WX15b].

Theorem 3.3.10. *Suppose we have the walk defined at* $(\mathbf{W}_\mu)$ *with* $\mu = \mathbf{0}$, $(\mathbf{V})$ *and,* $\mathcal{W}_n$, $\mathcal{S}_n$ *and* $\mathcal{V}_n$ *are the mean width, surface area and volume respectively of the hull of the d-dimensional random walk. Then, as* $n \to \infty$,

$$\begin{aligned} n^{-1/2}\mathcal{W}_n &\xrightarrow{\mathrm{d}} \mathcal{W}\left(\Sigma^{1/2}h_d\right) \\ n^{-(d-1)/2}\mathcal{S}_n &\xrightarrow{\mathrm{d}} \mathcal{S}\left(\Sigma^{1/2}h_d\right) \\ n^{-d/2}\mathcal{V}_n &\xrightarrow{\mathrm{d}} \mathcal{V}\left(\Sigma^{1/2}h_1^d\right) = v_d\sqrt{\det(\Sigma)} \end{aligned}$$

where v_d *is the volume of* h_d.

Proof. Notice that Theorem 3.3.4 gives

$$n^{-1/2}\mathrm{hull}\{S_0, S_1, \ldots, S_n\} \Rightarrow \Sigma^{1/2}h_d, \text{ on } (\mathfrak{C}_0^d, \rho_H),$$

where h_d is the hull of the d-dimensional Brownian motion starting at $b_d(0) = \mathbf{0}$. Using this fact and Lemma 3.3.7, it only remains to observe that the rescaling of the walk by $n^{-1/2}$ in all directions rescales $\mathcal{W}$ by $n^{-1/2}$, $\mathcal{S}$ by $n^{-(d-1)/2}$ and $\mathcal{V}$ by $n^{-d/2}$ which are continuous functions and therefore the mapping from the original walk to that of Brownian motion is also continuous. The result with the given limits follows, with the additional equality for the volume functional following from the Jacobian of the transformation $\mathbf{x} \mapsto \Sigma^{1/2}\mathbf{x}$ being $\sqrt{\det \Sigma}$. □

In the special case $d = 2$, $L_n := \mathcal{S}_n$ is the perimeter length of hull$\{S_0, \ldots, S_n\}$; Cauchy's formula also confirms that L_n is equal to $\mathcal{W}_n$ in this case, see Remark 3.3.8.

Theorem 3.3.11. *Let $d = 2$. Suppose that we have a random walk as defined at* $(\mathbf{W}_\mu)$. *Then*

$$n^{-1} L_n \xrightarrow{\text{a.s.}} 2\|\mu\|.$$

Remark 3.3.12. This result was proven in [MW18] 'directly' from the strong law of large numbers and Cauchy's surface area formula. Snyder and Steele [SS93] had previously obtained the result under the stronger condition $\mathbb{E}(\|Z\|^2) < \infty$ as a consequence of an upper bound on $\mathbb{V}\text{ar}\, \mathcal{L}_n$ deduced from Steele's version of the Efron–Stein inequality. In fact, Snyder and Steele state the result only for the case $\mu \neq \mathbf{0}$, but their proof works equally well when $\mu = \mathbf{0}$.

Proof. Using $\mathcal{L}_n = \mathcal{W}_n$ in the case $d = 2$, the almost-sure convergence of Theorem 3.3.3, the continuity of $\mathcal{W}_n$ from Lemma 3.3.7, and the continuous mapping theorem from Theorem 3.1.4 to establish $n^{-1}\mathcal{L}_n \xrightarrow{\text{a.s.}} \mathcal{W}(I_\mu[0,1])$. Without loss of generality, we will assume $\mu = \|\mu\|\mathbf{e}_{\pi/2}$ in order to calculate the right hand side explicitly:

$$\begin{aligned}\mathcal{W}(I_\mu[0,1]) = \int_{\mathbb{S}} h_{I_\mu[0,1]}(e)\mathrm{d}e &= \int_0^\pi (0, \|\mu\|) \cdot (\cos\theta, \sin\theta)\mathrm{d}\theta + \int_\pi^{2\pi} (0,0) \cdot (\cos\theta, \sin\theta)\mathrm{d}\theta \\ &= -\|\mu\| \cos\pi + \|\mu\| \cos 0 = 2\|\mu\|.\end{aligned}$$ □

We finish this section with the weak convergence statement for the d-dimensional volume of the walk with drift. This was also studied in [WX15b] for the specific case $d = 2$.

Theorem 3.3.13. *Suppose we have the walk defined at* $(\boldsymbol{W}_\mu)$ *with* $\|\mu\| > 0$, $(\boldsymbol{V})$ *holds and $\mathcal{V}_n$ is the volume of the hull of the d-dimensional random walk. Then, as $n \to \infty$,*

$$n^{-(d+1)/2}\mathcal{V}_n \xrightarrow{\text{d}} \|\mu\|\sqrt{\det \Sigma_{\mu_\perp}}\,\tilde{v}_d,$$

where $\tilde{v}_d$ *is the volume of* $\tilde{h}_d := \operatorname{hull} \tilde{b}_d[0,1]$ *where* $\tilde{b}_d[0,1] = \left\{ \tilde{b}_d(t) : t \in [0,1] \right\}$ *with* $\tilde{b}_d(t)$ *described at* (3.3.5) *and* $\Sigma_{\mu_\perp}$ *as described at* (3.3.3).

Proof. Recall the definition of $\tilde{\Sigma}^{1/2}_{\mu_\perp}$ from (3.3.4). Then note that $\operatorname{hull} \tilde{\Sigma}^{1/2}_{\mu_\perp} \tilde{b}_d[0,1] = \tilde{\Sigma}^{1/2}_{\mu_\perp} \operatorname{hull} \tilde{b}_d[0,1]$ because left multiplication by $\tilde{\Sigma}^{1/2}_{\mu_\perp}$ is an affine transformation, and that $\mathcal{V}(\tilde{\Sigma}^{1/2}_{\mu_\perp} A) = \sqrt{\det \tilde{\Sigma}_{\mu_\perp}} \mathcal{V}(A) = \sqrt{\det \Sigma_{\mu_\perp}} \mathcal{V}(A)$ because $\sqrt{\det \tilde{\Sigma}_{\mu_\perp}}$ is the Jacobian of the transformation. It follows that,

$$\mathcal{V}(\psi_{n,\mu}(A)) = n^{-(d+1)/2} \left(\|\mu\| \sqrt{\det \Sigma_{\mu_\perp}} \right)^{-1} \mathcal{V}(A) \tag{3.3.19}$$

for $A \in \mathfrak{C}^d_0$. Then we use Lemma 3.3.6, the continuous mapping theorem, and the continuity of the functional, Lemma 3.3.7 in the usual way with (3.3.19) to complete the proof. □

3.4 Functionals of Brownian motion

With all the functional limit theorem results stated, it is natural to ask what we know about Brownian motion in order to understand the functional central limit theorem results better. Of course, there is already a great deal of literature on this subject, discussed in Section 1.1.3. However, we add to the literature with the following work on the expected diameter of planar Brownian motion. There remain many open questions relating to these functionals of Brownian motion which we discuss below too.

Recall $b_2 = (b_2(t), t \in \mathbb{R}_+)$ is standard planar Brownian motion, and consider the set $b_2[0,1] = \{b_2(t) : t \in [0,1]\}$. The Brownian convex hull $h_2 := \operatorname{hull} b_2[0,1]$ has been well-studied from Lévy [Lév48, §52.6, pp. 254–256] onwards; the expectations of the perimeter length $\ell_2 := \mathcal{W}(h_2) = \mathcal{S}(h_2)$ and area $a_2 := \mathcal{V}(h_2)$ are given by the exact formulae $\mathbb{E}\, \ell_2 = \sqrt{8\pi}$ (due to Letac and Tákacs [Let78; Tak80]) and $\mathbb{E}\, a_2 = \pi/2$ (due to El Bachir [EB83]).

Another characteristic is the *diameter*

$$d_2 := \operatorname{diam} h_2 = \operatorname{diam} b_2[0,1] = \sup_{\mathbf{x},\mathbf{y} \in b_2[0,1]} \|\mathbf{x} - \mathbf{y}\|,$$

for which, in contrast, no explicit formula is known. The exact formulae for $\mathbb{E}\, \ell_2$ and $\mathbb{E}\, a_2$ rest on geometric integral formulae of Cauchy; since no such formula is available for d_2, it may not be possible to obtain an explicit formula for $\mathbb{E}\, d_2$. However, one may get bounds.

By convexity, we have the almost-sure inequalities $2 \le \ell_2/d_2 \le \pi$, the extrema being the line segment and shapes of constant width (such as the disc). In other words,

$$\frac{\ell_2}{\pi} \le d_2 \le \frac{\ell_2}{2}.$$

The formula of Letac and Takács [Let78; Tak80] says that $\mathbb{E}\, \ell_2 = \sqrt{8\pi}$, so we get:

Proposition 3.4.1. $\sqrt{8/\pi} \le \mathbb{E}\, d_2 \le \sqrt{2\pi}$.

Note that $\sqrt{8/\pi} \approx 1.5958$ and $\sqrt{2\pi} \approx 2.5066$. In this section we improve both of these bounds.

For the lower bound, we note that $b_2[0,1]$ is compact and thus, as a corollary of Lemma 3.4.6 below, we have the formula

$$d_2 = \sup_{0 \le \theta \le \pi} r(\theta), \tag{3.4.1}$$

where r is the parametrized range function given by

$$r(\theta) = \sup_{0 \le s \le 1} (b_s \cdot \mathbf{e}_\theta) - \inf_{0 \le s \le 1} (b_s \cdot \mathbf{e}_\theta),$$

with $\mathbf{e}_\theta$ being the unit vector $(\cos\theta, \sin\theta)$. Feller [Fel51] established that

$$\mathbb{E}\, r(\theta) = \sqrt{8/\pi} \quad \text{and} \quad \mathbb{E}(r(\theta)^2) = 4\log 2, \tag{3.4.2}$$

and the density of $r(\theta)$ is given explicitly as

$$f(r) = \frac{8}{\sqrt{2\pi}} \sum_{k=1}^{\infty} (-1)^{k-1} k^2 \exp\{-k^2 r^2/2\}, \ (r \ge 0). \tag{3.4.3}$$

Combining (3.4.1) with (3.4.2) gives immediately $\mathbb{E}\, d_2 \ge \mathbb{E}\, r(0) = \sqrt{8/\pi}$, which is just the lower bound in Proposition 3.4.1. For a better result, a consequence of (3.4.1) is that $d_2 \ge \max\{r(0), r(\pi/2)\}$. Observing that $r(0)$ and $r(\pi/2)$ are independent, we get:

Lemma 3.4.2. $\mathbb{E}\, d_2 \ge \mathbb{E} \max\{X_1, X_2\}$, *where X_1 and X_2 are independent copies of $X := r(0)$.*

It seems hard to explicitly compute $\mathbb{E} \max\{X_1, X_2\}$ in Lemma 3.4.2, because although the density given at (3.4.3) is known explicitly, it is not very tractable. Instead we obtain a lower bound. Since

$$\max\{x, y\} = \frac{1}{2}(x + y + |x - y|)$$

we get

$$\mathbb{E}\max\{X_1, X_2\} = \mathbb{E}\,X + \frac{1}{2}\,\mathbb{E}\,|X_1 - X_2|. \tag{3.4.4}$$

Thus with Lemma 3.4.2, the lower bound in Proposition 3.4.1 is improved given any non-trivial lower bound for $\mathbb{E}\,|X_1 - X_2|$. Using the fact that for any $c \in \mathbb{R}$, if m is a median of X, $\mathbb{E}\,|X - c| \geq \mathbb{E}\,|X - m|$, we see that

$$\mathbb{E}\,|X_1 - X_2| \geq \mathbb{E}\,|X - m|.$$

Again, the intractability of the density at (3.4.3) makes it hard to exploit this. Instead, we provide the following as a crude lower bound on $\mathbb{E}\,|X_1 - X_2|$.

Lemma 3.4.3. *Taking X_1, X_2 to be two independent copies of the arbitrary random variable X, for any $a, h > 0$,*

$$\mathbb{E}\,|X_1 - X_2| \geq 2h\,\mathbb{P}(X \leq a)\,\mathbb{P}(X \geq a + h).$$

Proof. We have

$$\begin{aligned}
\mathbb{E}\,|X_1 - X_2| \geq \mathbb{E}\left[|X_1 - X_2|\mathbf{1}\{X_1 \leq a, X_2 \geq a + h\}\right] \\
+ \mathbb{E}\left[|X_1 - X_2|\mathbf{1}\{X_2 \leq a, X_1 \geq a + h\}\right] \\
&\geq h\,\mathbb{P}(X_1 \leq a)\,\mathbb{P}(X_2 \geq a + h) + h\,\mathbb{P}(X_2 \leq a)\,\mathbb{P}(X_1 \geq a + h) \\
&= 2h\,\mathbb{P}(X \leq a)\,\mathbb{P}(X \geq a + h),
\end{aligned}$$

which proves the statement. □

This lower bound yields the following result.

Proposition 3.4.4. *For $a, h > 0$ define*

$$g(a, h) := h\left(\frac{4}{\pi}\exp\left\{-\frac{\pi^2}{2a^2}\right\} - \frac{4}{3\pi}\exp\left\{-\frac{9\pi^2}{2a^2}\right\}\right)\left(1 - \frac{4}{\pi}\exp\left\{-\frac{\pi^2}{8(a+h)^2}\right\}\right).$$

Then $\mathbb{E}\,d_2 \geq \sqrt{8/\pi} + g(1.492, 0.337) \approx 1.6014$.

Proof. Consider

$$Z := \sup_{0 \le s \le 1} |b_s \cdot \mathbf{e}_0|.$$

Then it is known (see [JP75]) that for $x > 0$,

$$\frac{4}{\pi} \exp\left\{-\frac{\pi^2}{8x^2}\right\} - \frac{4}{3\pi} \exp\left\{-\frac{9\pi^2}{8x^2}\right\} \le \mathbb{P}(Z < x) \le \frac{4}{\pi} \exp\left\{-\frac{\pi^2}{8x^2}\right\}. \tag{3.4.5}$$

Moreover, with $X = r(0)$ as above, we have

$$Z \le X \le 2Z.$$

Since $X \le 2Z$, we have

$$\mathbb{P}(X \le a) \ge \mathbb{P}(Z \le a/2) \ge \frac{4}{\pi} \exp\left\{-\frac{\pi^2}{2a^2}\right\} - \frac{4}{3\pi} \exp\left\{-\frac{9\pi^2}{2a^2}\right\},$$

by the lower bound in (3.4.5). On the other hand,

$$\mathbb{P}(X \ge a + h) \ge \mathbb{P}(Z \ge a + h) \ge 1 - \frac{4}{\pi} \exp\left\{-\frac{\pi^2}{8(a+h)^2}\right\},$$

by the upper bound in (3.4.5). Combining these two bounds and applying Lemma 3.4.3 we get $\mathbb{E}\,|X_1 - X_2| \ge 2g(a, h)$. So from (3.4.4) and the fact that $\mathbb{E}\,X = \sqrt{8/\pi}$ by (3.4.2) we get $\mathbb{E}\,d_2 \ge \sqrt{8/\pi} + g(a, h)$. Numerical evaluation using MAPLE suggests that $(a, h) = (1.492, 0.337)$ is close to optimal, and this choice gives the statement in the proposition. □

We also improve the upper bound in Proposition 3.4.1.

Proposition 3.4.5. $\mathbb{E}\,d_2 \le \sqrt{8 \log 2} \approx 2.3548.$

Proof. First, we claim that

$$d_2^2 \le r(0)^2 + r(\pi/2)^2. \tag{3.4.6}$$

It follows from (3.4.6) and (3.4.2) that

$$\mathbb{E}(d_2^2) \le \mathbb{E}(X_1^2 + X_2^2) = 2\,\mathbb{E}(X^2) = 8 \log 2.$$

The result now follows by Jensen's inequality.

It remains to prove the claim (3.4.6). Note that the diameter is an increasing function, that is, if $A \subseteq B$ then $\operatorname{diam} A \le \operatorname{diam} B$. Note also, that by the definition of $r(\theta)$,

$b_2[0,1] \subseteq \mathbf{z} + [0, r(0)] \times [0, r(\pi/2)] =: R_{\mathbf{z}}$ for some $\mathbf{z} \in \mathbb{R}^2$. Since the diameter of the set $R_{\mathbf{z}}$ is attained at the diagonal,

$$\operatorname{diam} R_{\mathbf{z}} = \sqrt{r(0)^2 + r(\pi/2)^2},$$

for all $\mathbf{z} \in \mathbb{R}^2$, and we have $\operatorname{diam} b_2[0,1] \leq \operatorname{diam} R_{\mathbf{z}}$, the result follows. □

We make one further remark about second moments. In the proof of Proposition 3.4.5, we saw that $\mathbb{E}(d_2^2) \leq 8 \log 2 \approx 5.5452$. A bound in the other direction can be obtained from the fact that $d_2^2 \geq \ell_1^2/\pi^2$, and we have (see [WX15b, §4.1]) that

$$\mathbb{E}(\ell_2^2) = 4\pi \int_{-\pi/2}^{\pi/2} \mathrm{d}\theta \int_0^\infty \mathrm{d}u \cos\theta \frac{\cosh(u\theta)}{\sinh(u\pi/2)} \tanh\left(\frac{(2\theta+\pi)u}{4}\right) \approx 26.1677,$$

which gives $\mathbb{E}(d_2^2) \geq 2.651$.

Finally, for completeness, we state and prove the lemma which was used to obtain equation (3.4.1).

Lemma 3.4.6. *Let $A \subset \mathbb{R}^2$ be a nonempty compact set, and let $r_A(\theta) = \sup_{\mathbf{x}\in A}(\mathbf{x} \cdot \mathbf{e}_\theta) - \inf_{\mathbf{x}\in A}(\mathbf{x} \cdot \mathbf{e}_\theta)$. Then*

$$\operatorname{diam} A = \sup_{0\leq\theta\leq\pi} r_A(\theta).$$

Proof. Since A is compact, for each θ there exist $\mathbf{x}, \mathbf{y} \in A$ such that

$$\begin{aligned} r_A(\theta) &= \mathbf{x} \cdot \mathbf{e}_\theta - \mathbf{y} \cdot \mathbf{e}_\theta \\ &= (\mathbf{x} - \mathbf{y}) \cdot \mathbf{e}_\theta \leq \|\mathbf{x} - \mathbf{y}\|. \end{aligned}$$

So $\sup_{0\leq\theta\leq\pi} r_A(\theta) \leq \sup_{\mathbf{x},\mathbf{y}\in A} \|\mathbf{x} - \mathbf{y}\| = \operatorname{diam} A$.

It remains to show that $\sup_{0\leq\theta\leq\pi} r_A(\theta) \geq \operatorname{diam} A$. This is clearly true if A consists of a single point, so suppose that A contains at least two points. Suppose that the diameter of A is achieved by $\mathbf{x}, \mathbf{y} \in A$ and let $\mathbf{z} = \mathbf{y} - \mathbf{x}$ be such that $\hat{\mathbf{z}} := \mathbf{z}/\|\mathbf{z}\| = \mathbf{e}_{\theta_0}$ for $\theta_0 \in [0, \pi]$. Then

$$\begin{aligned} \sup_{0\leq\theta\leq\pi} r_A(\theta) \geq r_A(\theta_0) &\geq \mathbf{y} \cdot \mathbf{e}_{\theta_0} - \mathbf{x} \cdot \mathbf{e}_{\theta_0} \\ &= \mathbf{z} \cdot \hat{\mathbf{z}} = \|\mathbf{z}\| = \operatorname{diam} A, \end{aligned}$$

as required. □

3.5 Application of results to our examples

We start by demonstrating the functional law of large numbers, Theorem 3.1.2. For the zero drift case, we actually appeal to Theorem 3.3.3 because then we can show the convergence of the hulls to the point at the origin. Since the walk is bounded by the convex hull, the pictures don't need the walk to be printed out to demonstrate that they also are trivial on the law of large numbers scale. In Figure 3.2 and Figure 3.3 we show later hulls in a darker colour which certainly show the decreasing size of area covered by the scaled random walks.

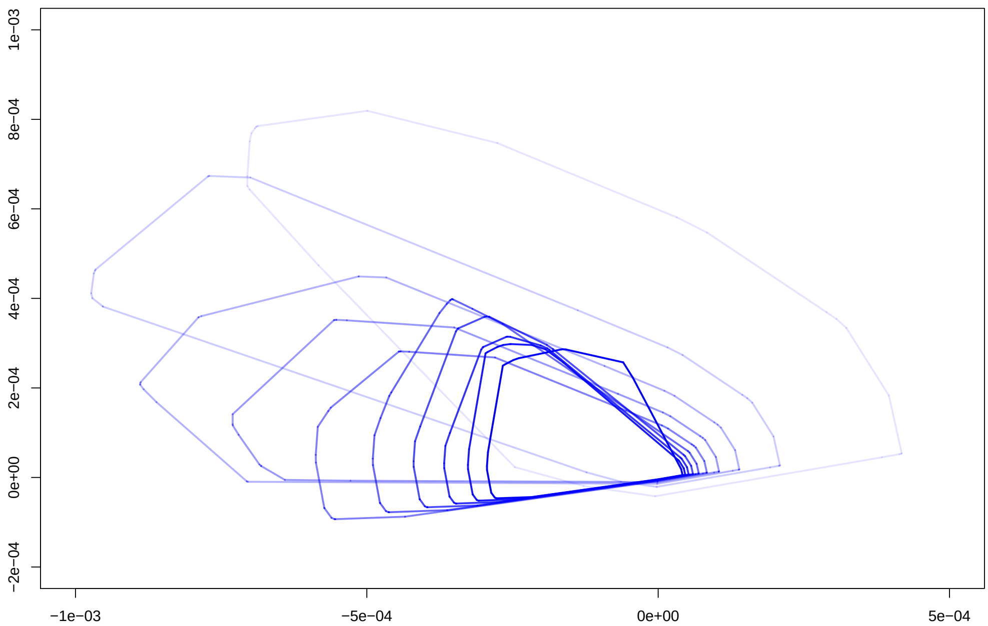


Figure 3.2: Convex hulls of one simple symmetric random walk scaled accordingly at 1,000,000 step intervals

We then turn to the case with drift. Here, after the law of large numbers scaling, we expect to see the trajectory converge to the linear vector from the origin to the mean. To demonstrate this, in Figures 3.4, 3.5 and 3.6, we plot our three walks with drift and at every $10,000$th point we draw a red cross. These crosses should be evenly spread out along the mean vector, which is what we see. Again, note that the vertical axis is on a smaller scale so small fluctuations in the vertical direction, as small as they are in the plots, should be even more microscopic. Note also, the walk with mean $(5,0)$ is converging to the vector from $(0,0)$ to $(5,0)$ as expected.

The next result to see in action is Theorem 3.1.10. It is not necessary to demonstrate this with all of our walks, so in one plot, Figure 3.7, we will use the simple symmetric random walk in black, the walk with fixed drift, unit mean, in blue and the walk with Normal drift and mean $(5,0)$ in red. For each we plot the running maximum scaled by the appropriate law of large numbers factor at each timestep on the vertical axis and the number of steps on the horizontal axis. It is clear that each walk converges to $\|\mu\|$ and quite quickly. It is hard to see any significant deviations from this value.

Our second example was the arcsine law. For this result we take our two zero drift examples and show the proportion of time spent on the positive side of the vertical axis as a process in itself, which is equivalent to taking A to be the upper half plane and plotting $\pi_n(A)$ against n. The plots show the simple symmetric random walk happened to spend most of the time on the positive side, in Figure 3.8, which would mean it spends almost no time on the negative side. Noting the obvious fact that the simple symmetric random walk is in fact symmetric and hence we were equally likely to observe a walk spending almost no time on the positive side, we see this picture backs up Corollary 3.1.16. The simulation of our standard Normal random walk, Figure 3.9, also shows how this proportion of time in a given set can vary more than was seen in our simple symmetric random walk simulation.

We also provide bar charts representing the empirical distributions for $\pi_n(A)$ where A is the upper half plane and also where A is the positive quadrant for each of the simple symmetric random walk, Figure 3.10 and Figure 3.12, and the standard Normal random walk, Figure 3.11 and Figure 3.13. These empirical distributions were established by running $10,000$ walks of each type and taking the proportion of time each had spent in the relevant sets in the first $10,000$ steps.

The first two plots are consistent with the conclusions of Lévy [Lév40b], that the upper half plane results display an empirical distribution that has greatest mass near 0 and 1 and the smallest mass around the centre. In order to add quantitative evidence that these empirical distributions are close to the arcsine law Lévy described, we calculate a simplified version of the Kolmogorov-Smirnov distance. Instead of taking $\rho_\infty(E_n, F)$ where $E_n(x) := \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\{\pi_{n,i}(A) \leq x\}$ is the empirical distribution function with $\pi_{n,i}$

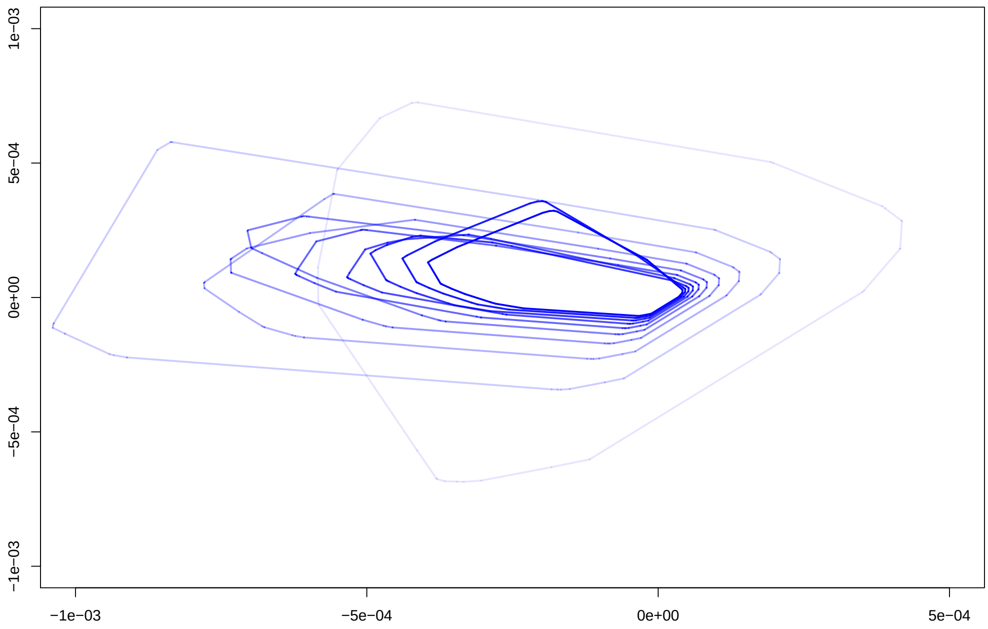


Figure 3.3: Convex hulls of one standard Normal random walk scaled accordingly at 1,000,000 step intervals

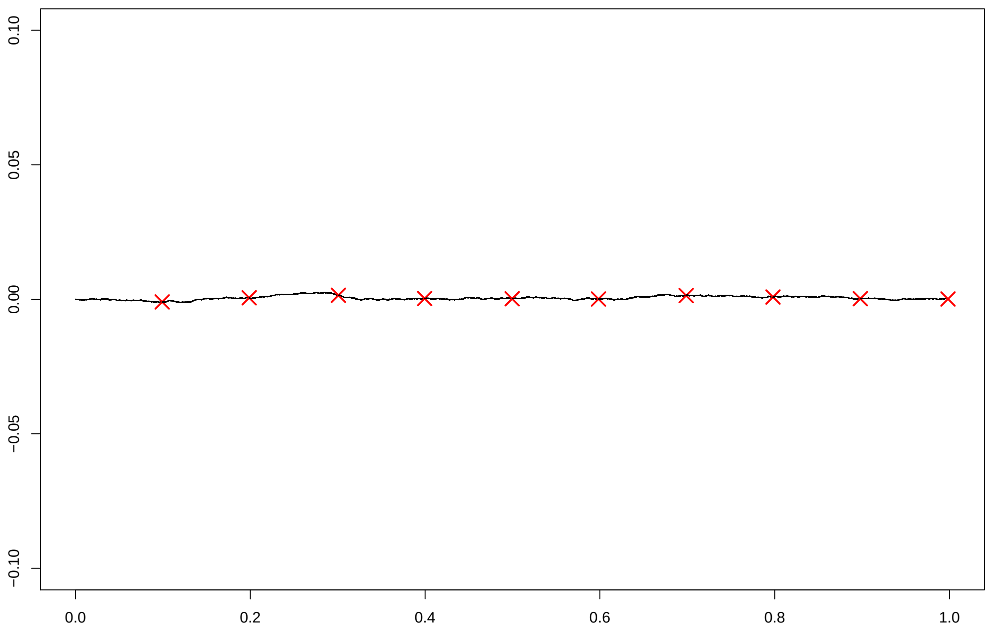


Figure 3.4: A simulation of one random walk with drift and all coordinates Normally distributed, unit mean to the right. Every 10,000th point of the walk plotted with a red cross.

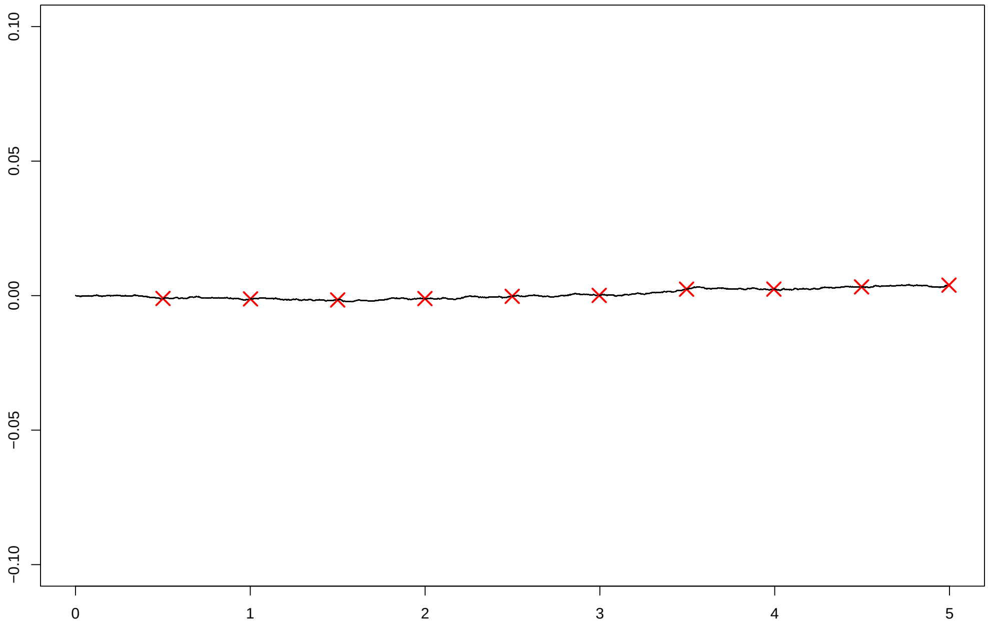


Figure 3.5: A simulation of one random walk with drift and all coordinates Normally distributed, mean of length 5 to the right. Every 10,000th point of the walk plotted with a red cross.

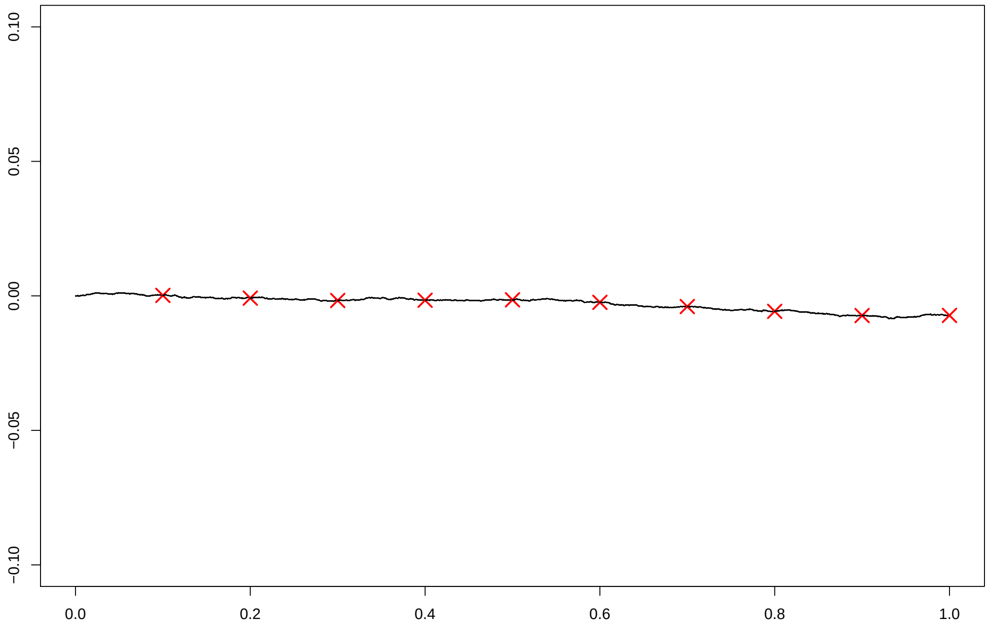


Figure 3.6: A simulation of one random walk with drift and no variance in the first coordinate, unit mean to the right. Every 10,000th point of the walk plotted with a red cross.

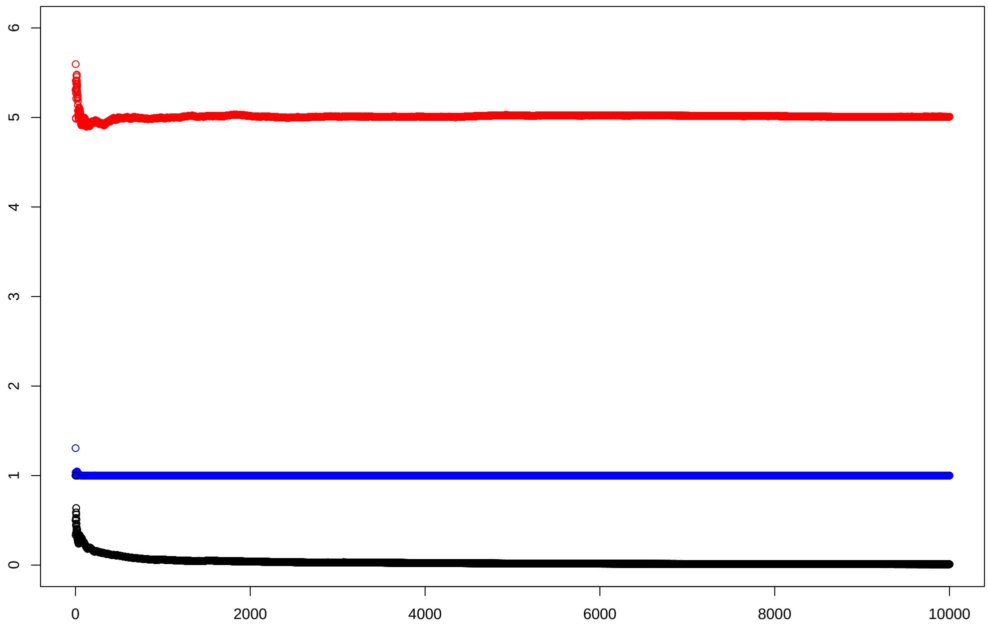


Figure 3.7: Scaled running maximum of three random walks plotted for the first 10^4 steps. Simple symmetric random walk in black, a random walk with drift and no variance in the first coordinate, unit mean to the right in blue, and a walk with Normal drift and mean $(5, 0)$ in red.

independent copies of $\pi_n(A)$, and $F(x) := 2/\pi \arcsin(\sqrt{x})$ is the arcsine law's cumulative distribution function, we just use a discrete approximation over the range $[0, 1]$ which we will call

$$\rho^k_{K-S}(F_{test}, F) = \sup_{0 \le i \le k} |E_n(i/k) - F(i/k)|. \tag{3.5.1}$$

We took $\rho^{40}_{K-S}(F_{test}, F)$ which coincides with the binning in the bar charts. For the simple symmetric random walk we got a a value of $\rho^{40}_{K-S}(F_{test}, F) = 0.099$ and for the standard Normal random walk a value of $\rho^{40}_{K-S}(F_{test}, F) = 0.094$. Both of these results are not too close to 0 but are not too far away either so could be considered as weak evidence to support the arcsine law, at least in comparison with the quadrant results below.

Meanwhile, the positive quadrant case backs up the conclusions of Bingham and Doney [BD88] that the distribution is not the arcsine law. These plots look starkly different with very little mass near 1 and most of the mass near 0. Again, calculating the simplified Kolmogorov-Smirnov distance, we get a value of $\rho^{40}_{K-S}(F_{test}, F) = 0.325$ for the simple symmetric random walk and $\rho^{40}_{K-S}(F_{test}, F) = 0.314$ for the standard Normal random walk. These clearly indicate that the arcsine law is not the limiting distribution for the case where A is a quadrant in the plane.

Beyond the preliminary examples, we next turned to results on the diameter. We have seen simulations demonstrating the law of large numbers result in the previous chapter, but now we can explore Theorem 3.2.6(b), the weak convergence statement for the diameter. For this, we provide the empirical distribution for $n^{-1/2} D_n$ for the simple symmetric random walk in 2-dimensions, Figure 3.14. The distribution seems slightly skewed to the right, with a mode in the region of 1.25 and a mean approximately 1.40. Multiplying these values by $\sqrt{2}$, which accounts for $\Sigma^{1/2}$ in Theorem 3.2.6(b), we get estimates of 1.77 and 1.98 for the mode and mean, respectively, of the distribution of d_2. This is in agreement with a larger simulation of diameters, this time following the Normal zero drift increment distribution. We carried out $1,000$ simulations of the random walk with $10,000$ steps, calculated the mean diameter of these simulations and then repeated this process $1,000$ times to get a vector of $1,000$ estimates for $\mathbb{E}\, d_2$. These values had mean 1.976 and variance only 0.0002 suggesting $\mathbb{E}\, d_2 \approx 1.98$. We note

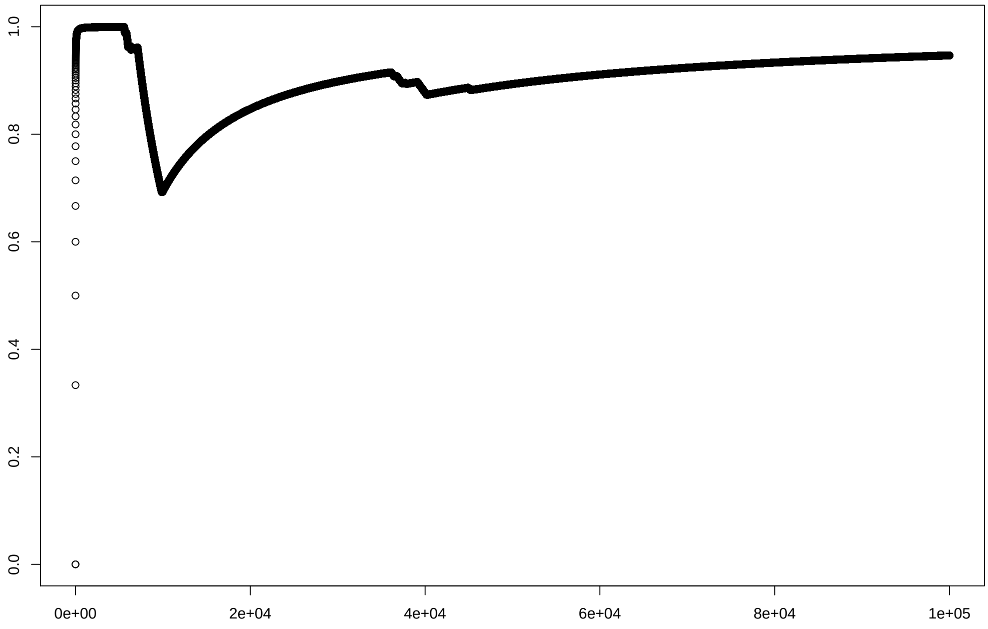


Figure 3.8: Proportion of time spent on the upper half plane is plotted against the number of steps taken, for 100, 000 steps of our simple symmetric random walk.

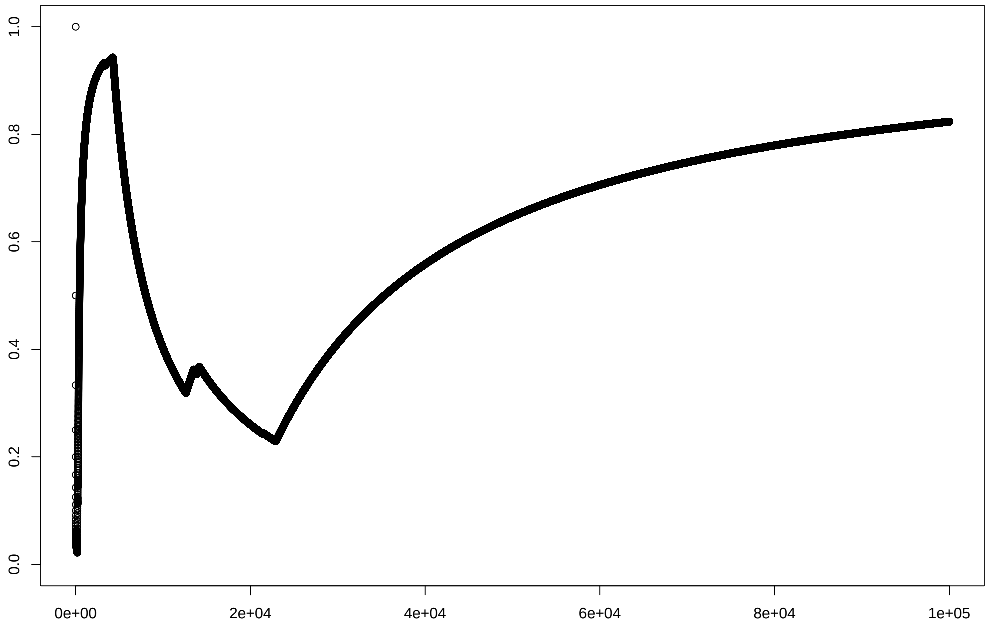


Figure 3.9: Proportion of time spent on the upper half plane is plotted against the number of steps taken, for 100, 000 steps of our standard Normal random walk.

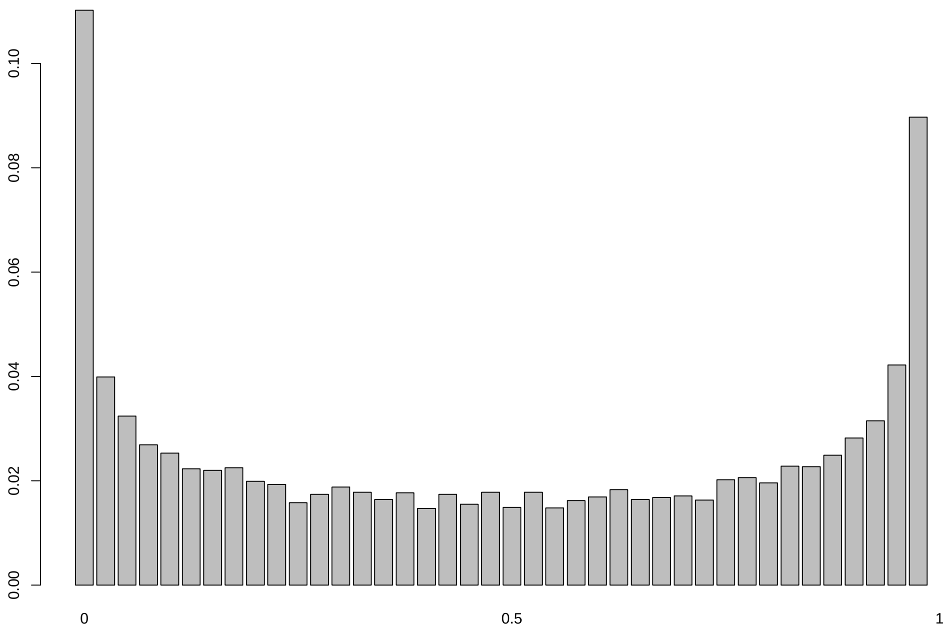


Figure 3.10: Empirical distribution of the proportion of time spent on the upper half plane for our simple symmetric random walk.

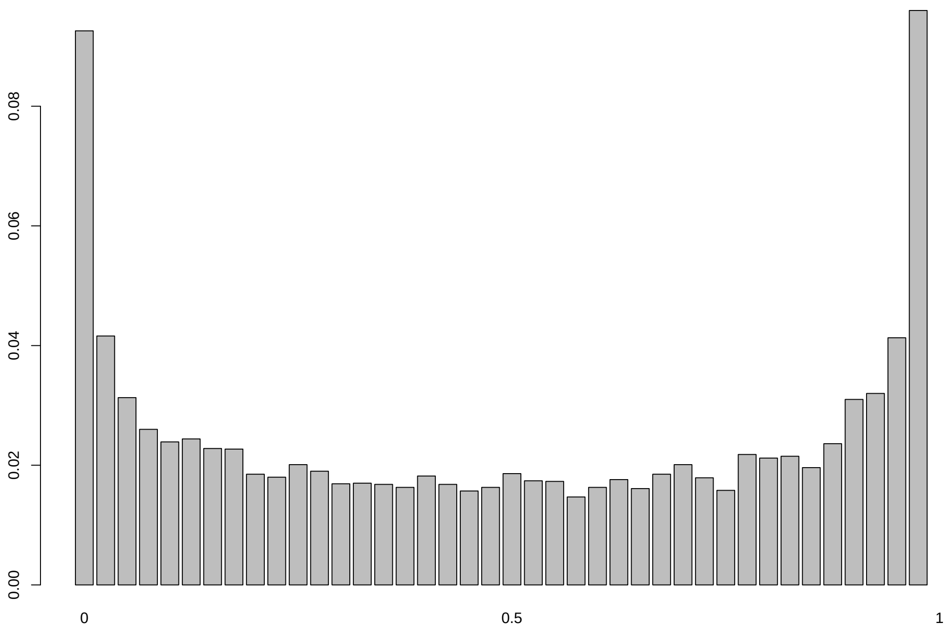


Figure 3.11: Empirical distribution of the proportion of time spent on the upper half plane for our standard Normal random walk.

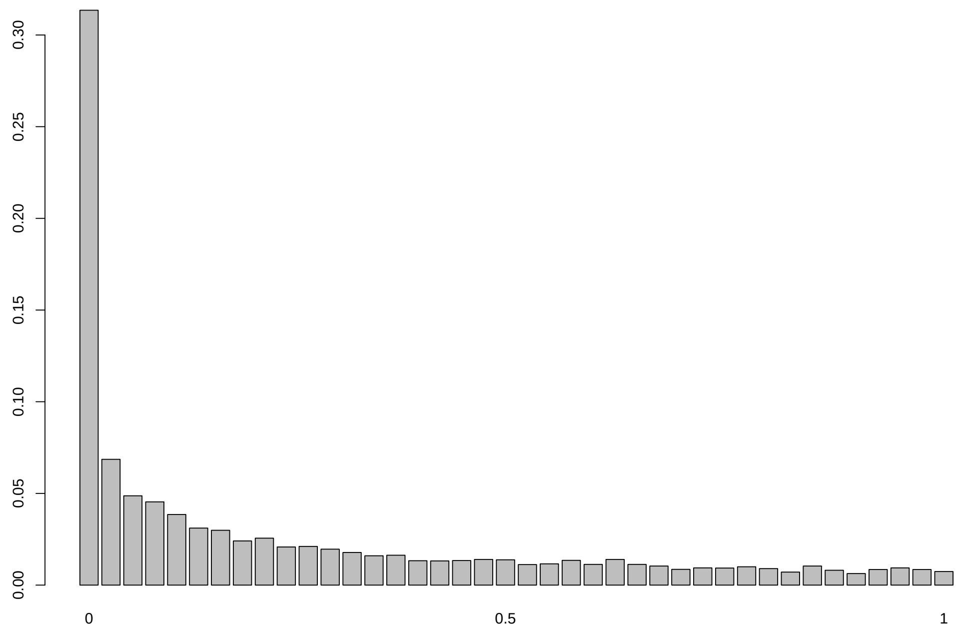


Figure 3.12: Empirical distribution of the proportion of time spent in the positive quadrant for our simple symmetric random walk.

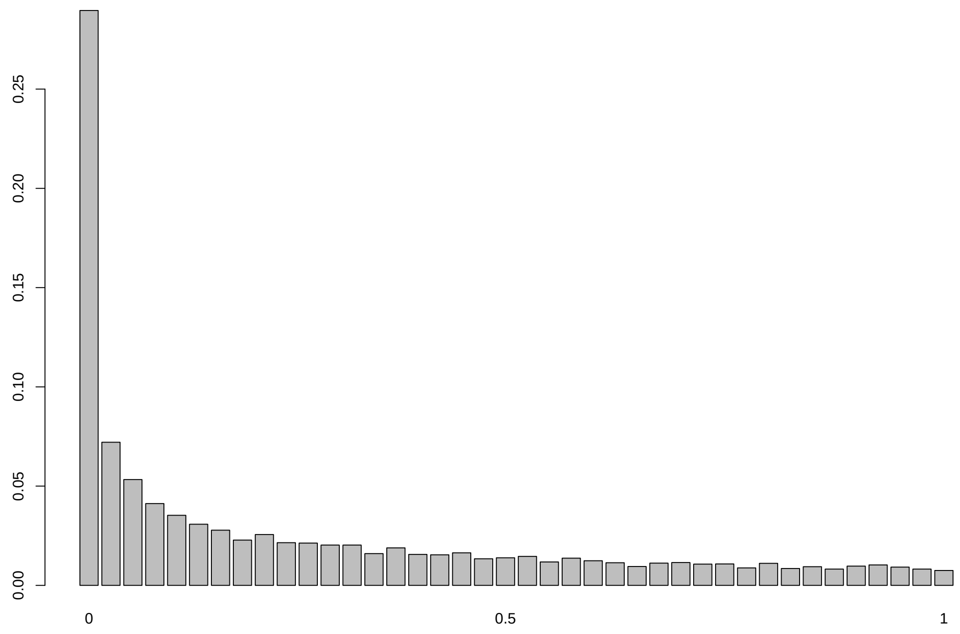


Figure 3.13: Empirical distribution of the proportion of time spent in the positive quadrant for our standard Normal random walk.

that it seems, running the simulations for different values of n it appears the diameter values increase as n increases. Although this looks like it is only a small bias in the simulations, it is not impossible that our simulations have enough error so that we are underestimating the diameter and in fact it could be $\mathbb{E}\, d_2 = 2$. All of which agrees with the bounds $1.6014 \le \mathbb{E}\, d_2 \le 2.3548$ established in Section 3.4.

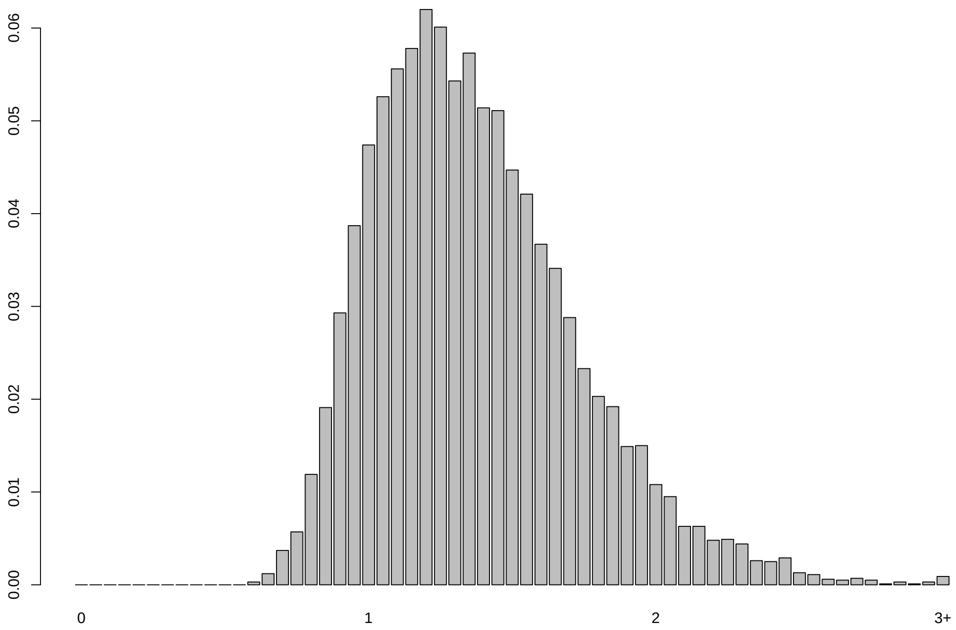


Figure 3.14: Empirical distribution of the rescaled diameter for our simple symmetric random walk.

These plots conclude the simulations for this section. We omit empirical distribution plots for the functionals in Theorem 3.3.10 and Theorem 3.3.13 because the functionals become increasingly complex to calculate for even $d = 3$. Note, however, that the law of large numbers for the perimeter length in $d = 2$ was examined in the previous chapter and some information on moments and the limit distribution in $d = 2$ was established in [WX15b].

Chapter 4

A zero-one law and shape result

In Chapter 2 a law of large numbers result for the ratio L_n/D_n was established for the 2-dimensional case with drift, Corollary 2.1.4. A natural question arises about the behaviour of this ratio for the 2-dimensional zero drift case.

Some insight can be found by the observation in Chapter 3 that we have a non-degenerate scaling limit for the zero drift case, namely planar Brownian motion. So the question could be solved by considering the ratio $\mathcal{L}(\Sigma^{1/2}b_2)/\mathcal{D}(\Sigma^{1/2}b_2)$. However, as discussed previously, this ratio gives crude information about the shape of the convex hull of the random walk, so we may be tempted to go further and ask about the possible limiting shapes of the convex hull of planar Brownian motion.

This is exactly what we do in this chapter[1], showing that the appropriately rescaled convex hull of the zero drift random walk infinitely often approximates any convex set with unit diameter. Hence, the convex hull infinitely often becomes arbitrarily close to a shape with $L_n/D_n = 2$ and infinitely often becomes arbitrarily close to a shape with $L_n/D_n = \pi$, and hence the ratio has no limit.

4.1 Shape results

Recall, $\mathcal{H}_n := \text{hull}(S_0, S_1, \dots, S_n)$, and with the extra condition $\mathbb{E}(\|Z\|^2) < \infty$, we let $\Sigma := \mathbb{E}(ZZ^\top)$, viewing Z as a column vector. Also, recall ρ_H is the Hausdorff distance

[1]Based on work published in [MW18] which was joint work between the authors.

between non-empty compact sets; see (1.3.2) above for a definition. Our result will be stated for elements of the set of compact convex sets in $\mathbb{R}^2$ containing the origin, which we recall is denoted $\mathfrak{C}_0^2$.

Then the formal statement of the limit, or lack thereof, of the shape of the convex hull is the following theorem.

Theorem 4.1.1. *Suppose we have the walk defined at* ($\mathbf{W}_\mu$) *with* $\mu = \mathbf{0}$ *and satisfying* ($\mathbf{V}$) *such that* Σ *is positive definite. Then, for any compact convex set* $K \in \mathfrak{C}_0^2$ *with* $\operatorname{diam} K = 1$,

$$\liminf_{n\to\infty} \rho_H(D_n^{-1}\mathcal{H}_n, K) = 0, \ \textit{a.s.}$$

Remark 4.1.2.

- Note that under the hypotheses of Theorem 4.1.1, $\mathbb{P}(Z = \mathbf{0}) < 1$, so that $D_n > 0$ for all but finitely many n, a.s.
- The non-random scaling of $n^{-1/2}$ might seem more natural, and we posit that a similar result would hold where K would be any compact convex set (not necessarily with $\operatorname{diam} K = 1$).

A consequence of Theorem 4.1.1 is the following result, which should be contrasted with Corollary 2.1.4.

Corollary 4.1.3. *Suppose we have the walk defined at* ($\mathbf{W}_\mu$) *with* $\mu = \mathbf{0}$ *and satisfying* ($\mathbf{V}$) *such that* Σ *is positive definite. Then,*

$$2 = \liminf_{n\to\infty} \frac{L_n}{D_n} < \limsup_{n\to\infty} \frac{L_n}{D_n} = \pi, \ \textit{a.s.}$$

4.2 A zero-one law for convex hulls

A key ingredient in the proof of Theorem 4.1.1 is a *zero-one law* (Theorem 4.2.1 below). Before we state the result, we need some extra notation. Define σ-algebras $\mathcal{F}_0 := \{\emptyset, \Omega\}$ and $\mathcal{F}_n := \sigma(Z_1, \ldots, Z_n)$ for $n \geq 1$; also set $\mathcal{F}_\infty := \sigma(\cup_{n\geq 0}\mathcal{F}_n)$. Also, recall the notation $\mathcal{B}(\mathfrak{C}_0^2)$ is used for the Borel σ-algebra in this case generated by the metric ρ_H.

Since the function $(\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_n) \mapsto \operatorname{hull}\{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_n\}$ (with $\mathbf{x}_0 := \mathbf{0}$) is continuous from $(\mathbb{R}^{2(n+1)}, \rho_E)$ to $(\mathcal{K}, \rho_H)$ where ρ_E is here the Euclidean metric in $\mathbb{R}^{2(n+1)}$, it is measurable from $(\mathbb{R}^{2(n+1)}, \mathcal{B}(\mathbb{R}^{2(n+1)}))$ to $(\mathfrak{C}_0^2, \mathcal{B}(\mathfrak{C}_0^2))$; thus $\mathcal{H}_n$ is a $\mathfrak{C}_0^2$-valued random variable, and $\mathcal{H}_n$ is $\mathcal{F}_n$-measurable.

For $n \geq 0$, set $\mathcal{T}_n := \sigma(\mathcal{H}_n, \mathcal{H}_{n+1}, \ldots)$ and define $\mathcal{T} := \cap_{n\geq 0}\mathcal{T}_n$. Also, for $n \geq 0$ define

$$r_n := \inf\{\|\mathbf{x}\| : \mathbf{x} \in \mathbb{R}^2 \setminus \mathcal{H}_n\}.$$

Note that r_n is non-decreasing. Here is the zero-one law.

Theorem 4.2.1. *Suppose that $r_n \to \infty$ a.s. Then if $A \in \mathcal{T}$, $\mathbb{P}(A) \in \{0, 1\}$.*

Next we give a sufficient condition for $r_n \to \infty$. Recall [Dur10, p. 190] that S_n is *recurrent* if there is a non-empty set $\mathcal{R}$ of points $\mathbf{x} \in \mathbb{R}^2$ (the recurrent values) such that, for any $\varepsilon > 0$, $\|S_n - \mathbf{x}\| < \varepsilon$ i.o., a.s.

Proposition 4.2.2. *If S_n is genuinely 2-dimensional and recurrent, then $r_n \to \infty$ a.s.*

Remark 4.2.3. One may also have $r_n \to \infty$ a.s. in the case of a transient walk, provided it visits all angles. However, $\lim_{n\to\infty} r_n < \infty$ a.s. may occur if the walk has a limiting direction, such as if there is a finite non-zero drift.

Let $B(\mathbf{x}; r)$ denote the closed Euclidean ball centred at $\mathbf{x} \in \mathbb{R}^2$ with radius r.

Proof of Proposition 4.2.2. Since S_n is recurrent, the set $\mathcal{R}$ of recurrent values is a closed subgroup of $\mathbb{R}^2$ and coincides with the set of *possible values* for the walk: see [Dur10, p. 190]. Since S_n is genuinely 2-dimensional, it follows from e.g. Theorem 21.2 of [BR10, p. 225] that $\mathcal{R}$ contains a further closed subgroup $\mathcal{R}'$ of the form $H\mathbb{Z}^2$ where H is a non-singular 2 by 2 matrix. Hence there exists $h > 0$ such that for every $\mathbf{x} \in \mathbb{R}^2$ there exists $\mathbf{y} \in \mathcal{R}'$ with $\|\mathbf{x} - \mathbf{y}\| < h/2$. In particular, for any $\mathbf{x} \in \mathbb{R}^2$, $\mathbb{P}(S_n \in B(\mathbf{x}; h) \text{ i.o.}) = 1$.

Fix $r > h$, and consider 4 discs, D_1, D_2, D_3, D_4, each of radius h, centred at $(\pm 2r, \pm 2r)$. Define T_r to be the first time at which the walk has visited all 4 discs, i.e.,

$$T_r := \min\{n \geq 0 : \exists\ i_1, i_2, i_3, i_4 \in [0, n] \text{ with } S_{i_j} \in D_j \text{ for } j = 1, 2, 3, 4\}.$$

The first paragraph of this proof shows that $T_r < \infty$ a.s. By construction, for $n \geq T_r$ we have that $\mathcal{H}_n$ contains the square $[-r, r]^2$, and so $n \geq T_r$ implies $r_n \geq r$. Hence,

$$\mathbb{P}\left(\liminf_{m\to\infty} r_m \geq r\right) \geq \mathbb{P}(T_r \leq n) \to 1,$$

as $n \to \infty$, and so $\liminf_{n\to\infty} r_n \geq r$, a.s. Since $r > h$ was arbitrary, the result follows. □

The first step in the proof of Theorem 4.2.1 is the following result, which uses the fact that $r_n \to \infty$ to show that any initial segment of the trajectory is eventually contained in the interior of the convex hull, uniformly over permutations of the initial increments.

Lemma 4.2.4. *Suppose that $r_n \to \infty$ a.s. Let $k \in \mathbb{N}$. Then there exists a random variable N_k with $\mathbb{P}(k < N_k < \infty) = 1$ such that (i) N_k is invariant under permutations of $Z_1, \ldots, Z_k$, and (ii) $\mathcal{H}_n = \text{hull}\{S_{k+1}, \ldots, S_n\}$ for all $n \geq N_k$ and all permutations of $Z_1, \ldots, Z_k$.*

Proof. Fix $k \in \mathbb{N}$. Let $R_k := \sum_{i=1}^{k} \|Z_i\|$ and define $N_k := \min\{n > k : r_n > R_k\}$. Note that since r_n is non-decreasing, $n \geq N_k$ implies $r_n > R_k$. Since $R_k < \infty$ a.s. and $r_n \to \infty$ a.s., we have $N_k < \infty$ a.s. Observe that if $r_n > R_k$ for $n > k$, then $S_0, S_1, \ldots, S_k$ are all contained in the interior of $\mathcal{H}_n$, for all permutations of $Z_1, \ldots, Z_k$, so that $\mathcal{H}_n = \mathcal{H}_{n,k} := \text{hull}\{S_{k+1}, \ldots, S_n\}$. So statement (ii) holds. Moreover, if $r_{n,k} := \inf\{\|\mathbf{x}\| : \mathbf{x} \in \mathbb{R}^2 \setminus \mathcal{H}_{n,k}\}$ we have that $\{r_n > R_k\} = \{r_{n,k} > R_k\}$. But the events $\{r_{n,k} > R_k\}$, $n > k$, which determine N_k, depend only on R_k and $S_{k+1}, S_{k+2}, \ldots$, and so statement (i) holds. □

Heuristically, it seems clear that Theorem 4.2.1 is true, since any $A \in \mathcal{T}$ is determined by $\mathcal{H}_{N_k}, \mathcal{H}_{N_k+1}, \ldots$, and Lemma 4.2.4 shows that this sequence in invariant under permutations of $Z_1, \ldots, Z_k$, as required for the Hewitt–Savage zero-one law. The formal proof is as follows.

Proof of Theorem 4.2.1. We adapt one of the standard proofs of the Hewitt–Savage zero-one law; see e.g. [Dur10, pp. 180–181]. Let $A \in \mathcal{T}$ and fix $\varepsilon > 0$. Recall a fact from measure theory: if $\mathcal{A}$ is an algebra and $A \in \sigma(\mathcal{A})$, then we can find $A' \in \mathcal{A}$ such that $\mathbb{P}(A \triangle A') < \varepsilon$ (see Theorem 1.3.1 or e.g. [Bil12, p. 179]). Applied to the algebra $\cup_{n\geq 0}\mathcal{F}_n$ which generates $\mathcal{F}_\infty \supseteq \mathcal{T}$, this result implies that we can find $k \geq 0$ and $A_k \in \mathcal{F}_k$ such that $\mathbb{P}(A \triangle A_k) < \varepsilon$. Fix this k, and fix n such that $\mathbb{P}(N_{2k} > n) < \varepsilon$, where N_{2k} is as given in Lemma 4.2.4. Applied to the algebra $\mathcal{A}_n := \cup_{m\geq 0}\sigma(\mathcal{H}_n, \mathcal{H}_{n+1}, \ldots, \mathcal{H}_{n+m})$, which has $\sigma(\mathcal{A}_n) \supseteq \mathcal{T}_n \supseteq \mathcal{T}$, the same measure-theoretic result shows that we can find $E_n \in \mathcal{A}_n$ such that $\mathbb{P}(A \triangle E_n) < \varepsilon$.

Now $A_k \in \mathcal{F}_k$ can be expressed as $A_k = \{Z_1 \in C_{k,1}, \dots, Z_k \in C_{k,k}\}$ for Borel sets $C_{k,1}, \dots, C_{k,k}$. Set $A_k' := \{Z_{k+1} \in C_{k,1}, \dots, Z_{2k} \in C_{k,k}\}$; since the Z_i are i.i.d., $\mathbb{P}(A_k') = \mathbb{P}(A_k)$, and A_k and A_k' are independent. We claim that

$$\mathbb{P}((A_k' \triangle E_n) \cap \{N_{2k} \le n\}) = \mathbb{P}((A_k \triangle E_n) \cap \{N_{2k} \le n\}) \le 2\varepsilon. \tag{4.2.1}$$

To see the equality in (4.2.1), observe that Lemma 4.2.4 shows that $E_n \cap \{N_{2k} \le n\}$ is invariant under permutations of $Z_1, \dots, Z_{2k}$, where the Z_i are i.i.d. For the inequality in (4.2.1), we use the fact that $\mathbb{P}(A \triangle B) \le \mathbb{P}(A \triangle C) + \mathbb{P}(B \triangle C)$ to get

$$\begin{aligned}\mathbb{P}((A_k \triangle E_n) \cap \{N_{2k} \le n\}) &\le \mathbb{P}(A_k \triangle E_n) \\ &\le \mathbb{P}(A_k \triangle A) + \mathbb{P}(E_n \triangle A) \le 2\varepsilon.\end{aligned}$$

Hence the claim (4.2.1) is verified. Since $\mathbb{P}((A \triangle B) \cap D) \le \mathbb{P}((A \triangle C) \cap D) + \mathbb{P}(B \triangle C)$, we also get that

$$\mathbb{P}((A \triangle A_k') \cap \{N_{2k} \le n\}) \le \mathbb{P}((A_k' \triangle E_n) \cap \{N_{2k} \le n\}) + \mathbb{P}(A \triangle E_n) \le 3\varepsilon,$$

by (4.2.1). Hence

$$\mathbb{P}(A \triangle A_k') \le \mathbb{P}(N_{2k} > n) + \mathbb{P}((A \triangle A_k') \cap \{N_{2k} \le n\}) \le 4\varepsilon.$$

The final sequence of the proof is a variation of the standard argument. First note that

$$|\mathbb{P}(A)^2 - \mathbb{P}(A)| \le |\mathbb{P}(A)^2 - \mathbb{P}(A_k \cap A_k')| + |\mathbb{P}(A_k \cap A_k') - \mathbb{P}(A)|. \tag{4.2.2}$$

For the first term on the right-hand side of (4.2.2), we use the fact that A_k and A_k' are independent with $\mathbb{P}(A_k) = \mathbb{P}(A_k')$, along with the property of the symmetric difference operator that $|\mathbb{P}(A) - \mathbb{P}(B)| \le \mathbb{P}(A \triangle B)$, to get

$$\begin{aligned}|\mathbb{P}(A)^2 - \mathbb{P}(A_k \cap A_k')| &= |\mathbb{P}(A)^2 - \mathbb{P}(A_k)^2| \\ &= |\mathbb{P}(A) + \mathbb{P}(A_k)||\mathbb{P}(A) - \mathbb{P}(A_k)| \\ &\le 2\mathbb{P}(A \triangle A_k) \le 2\varepsilon.\end{aligned}$$

Now considering the second term on the right-hand side of (4.2.2) and using the fact that $\mathbb{P}(A \triangle (B \cap C)) \leq \mathbb{P}(A \triangle B) + \mathbb{P}(A \triangle C)$, we have

$$\begin{aligned}|\mathbb{P}(A_k \cap A'_k) - \mathbb{P}(A)| &\leq \mathbb{P}(A \triangle (A_k \cap A'_k)) \\ &\leq \mathbb{P}(A \triangle A_k) + \mathbb{P}(A \triangle A'_k) \leq 5\varepsilon.\end{aligned}$$

Combining these two bounds, we obtain from (4.2.2) that $|\mathbb{P}(A)^2 - \mathbb{P}(A)| \leq 7\varepsilon$. Since $\varepsilon > 0$ was arbitrary, we get the result. □

The strategy of the proof of Theorem 4.1.1, carried out in the remainder of this section, is as follows. We use Donsker's theorem and the mapping theorem to show that $D_n^{-1}\mathcal{H}_n$ converges weakly to the convex hull of an appropriate Brownian motion, scaled to have unit diameter (Lemma 4.2.7). This limiting set has positive probability of being an arbitrarily good approximation to any given unit-diameter convex compact set K. An application of the zero-one law (Theorem 4.2.1) then completes the proof.

For $K \in \mathfrak{C}_0^2$ let $\mathcal{D}(K) := \operatorname{diam} K$. The next result shows that the map $K \mapsto \mathcal{D}(K)$ is continuous from $(\mathfrak{C}_0^2, \rho_H)$ to $(\mathbb{R}_+, \rho_E)$.

Lemma 4.2.5. *For $K_1, K_2 \in \mathfrak{C}_0^2$, $|\mathcal{D}(K_1) - \mathcal{D}(K_2)| \leq 2\rho_H(K_1, K_2)$.*

Proof. Let $\rho_H(K_1, K_2) = r$. From (1.3.1) we have that for any $\mathbf{x}_1, \mathbf{x}_2 \in K_1$ and any $s > r$, there exist $\mathbf{y}_1, \mathbf{y}_2 \in K_2$ such that $\|\mathbf{x}_i - \mathbf{y}_i\| \leq s$. Then,

$$\|\mathbf{x}_1 - \mathbf{x}_2\| \leq \|\mathbf{x}_1 - \mathbf{y}_1\| + \|\mathbf{y}_1 - \mathbf{y}_2\| + \|\mathbf{y}_2 - \mathbf{x}_2\| \leq 2s + \mathcal{D}(K_2).$$

Hence $\mathcal{D}(K_1) \leq 2s + \mathcal{D}(K_2)$, and since $s > r$ was arbitrary we get $\mathcal{D}(K_1) - \mathcal{D}(K_2) \leq 2r$. A symmetric argument gives $\mathcal{D}(K_2) - \mathcal{D}(K_1) \leq 2r$. □

For $K \in \mathfrak{C}_0^2$ and $\mathbf{x} \in \mathbb{S} := \{\mathbf{y} \in \mathbb{R}^2 : \|\mathbf{y}\| = 1\}$, define $h_K(\mathbf{x}) := \sup_{\mathbf{y} \in K}(\mathbf{y} \cdot \mathbf{x})$. Equivalent to (1.3.2) for $K_1, K_2 \in \mathfrak{C}_0^2$ is the formula [Gru07, p. 84]

$$\rho_H(K_1, K_2) = \sup_{\mathbf{x} \in \mathbb{S}} |h_{K_1}(\mathbf{x}) - h_{K_2}(\mathbf{x})|. \tag{4.2.3}$$

Let $\mathfrak{C}^\star := \{K \in \mathfrak{C}_0^2 : \mathcal{D}(K) > 0\} = \mathfrak{C}_0^2 \setminus \{\{\mathbf{0}\}\}$.

Lemma 4.2.6. *Suppose that $K_1, K_2 \in \mathfrak{C}^\star$. Then*

$$\rho_H(K_1/\mathcal{D}(K_1), K_2/\mathcal{D}(K_2)) \le \frac{3\rho_H(K_1, K_2)}{\mathcal{D}(K_1)}. \tag{4.2.4}$$

In particular, the map $K \mapsto K/\mathcal{D}(K)$ is continuous from $(\mathfrak{C}^\star, \rho_H)$ to $(\mathfrak{C}^\star, \rho_H)$.

Proof. We first claim that for $K_1, K_2 \in \mathfrak{C}_0^2$ and $\alpha_1, \alpha_2 > 0$,

$$\rho_H(\alpha_1 K_1, \alpha_2 K_2) \le \alpha_1 \rho_H(K_1, K_2) + |\alpha_1 - \alpha_2| \mathcal{D}(K_2). \tag{4.2.5}$$

Suppose that $K_1, K_2 \in \mathfrak{C}^\star$. Applying (4.2.5) with $\alpha_i = 1/\mathcal{D}(K_i)$, we get

$$\rho_H(K_1/\mathcal{D}(K_1), K_2/\mathcal{D}(K_2)) \le \frac{\rho_H(K_1, K_2)}{\mathcal{D}(K_1)} + \frac{|\mathcal{D}(K_1) - \mathcal{D}(K_2)|}{\mathcal{D}(K_1)},$$

from which (4.2.4) follows by Lemma 4.2.5. This gives the desired continuity.

It remains to verify the claim (4.2.5). From (4.2.3), with the observation that, for $\alpha > 0$, $h_{\alpha K}(\mathbf{x}) = \alpha h_K(\mathbf{x})$, it follows that

$$\begin{aligned} \rho_H(\alpha_1 K_1, \alpha_2 K_2) &= \sup_{\mathbf{x} \in \mathbb{S}} |\alpha_1 h_{K_1}(\mathbf{x}) - \alpha_1 h_{K_2}(\mathbf{x}) + (\alpha_1 - \alpha_2) h_{K_2}(\mathbf{x})| \\ &\le \alpha_1 \sup_{\mathbf{x} \in \mathbb{S}} |h_{K_1}(\mathbf{x}) - h_{K_2}(\mathbf{x})| + |\alpha_1 - \alpha_2| \sup_{\mathbf{x} \in \mathbb{S}} h_{K_2}(\mathbf{x}), \end{aligned}$$

from which the claim (4.2.5) follows. □

Suppose that $\Sigma := \mathbb{E}(ZZ^\top)$ is positive definite and recall $\Sigma^{1/2}$ denotes the (unique) positive-definite symmetric matrix such that $\Sigma^{1/2}\Sigma^{1/2} = \Sigma$. Further recall $(b_2(t), t \ge 0)$ is standard Brownian motion in $\mathbb{R}^2$ and $h_2 := \operatorname{hull} b_2[0,1]$, the convex hull of Brownian motion run for unit time. The map $\mathbf{x} \mapsto \Sigma^{1/2}\mathbf{x}$ is an affine transformation of $\mathbb{R}^2$, such that $\Sigma^{1/2} b_2$ is Brownian motion with covariance matrix Σ, and $\Sigma^{1/2} h_2 = \operatorname{hull} \Sigma^{1/2} b_2[0,1]$ is the corresponding convex hull.

Lemma 4.2.7. *Suppose that $\mathbb{E}(\|Z\|^2) < \infty$, $\mu = \mathbf{0}$, and Σ is positive definite. Then*

$$D_n^{-1} \mathcal{H}_n \Rightarrow \frac{\Sigma^{1/2} h_2}{\mathcal{D}(\Sigma^{1/2} h_2)},$$

in the sense of weak convergence on $(\mathfrak{C}_0^2, \rho_H)$.

Proof. The convergence $n^{-1/2}\mathcal{H}_n \Rightarrow \Sigma^{1/2} h_2$ is given in Theorem 3.3.4, see also Theorem 2.5 of [WX15b]. Since (by Lemma 4.2.6) $K \mapsto K/\mathcal{D}(K)$ is continuous on $\mathfrak{C}^\star$, and

$\mathbb{P}(\Sigma^{1/2}h_2 \in \mathfrak{C}^\star) = 1$, we may apply the mapping theorem [Bil99, p. 21] to deduce the result. □

Proof of Theorem 4.1.1. Fix $K \in \mathfrak{C}_0^2$ with $\mathcal{D}(K) = 1$. We claim that, for any $\varepsilon > 0$,

$$\mathbb{P}\left(\liminf_{n\to\infty} \rho_H\left(D_n^{-1}\mathcal{H}_n, K\right) \le \varepsilon\right) > 0. \tag{4.2.6}$$

Under the conditions of the theorem, S_n is genuinely 2-dimensional and recurrent [Dur10, p. 195], and so, by Proposition 4.2.2, $r_n \to \infty$ a.s. Since the event in (4.2.6) is in $\mathcal{T}$, the zero-one law (Theorem 4.2.1) shows that the probability in (4.2.6) must be equal to 1. Since $\varepsilon > 0$ was arbitrary, the statement of the theorem follows.

Thus it remains to prove the claim (4.2.6). To this end, observe that, for any $\varepsilon > 0$,

$$\begin{aligned}
\mathbb{P}\left(\liminf_{n\to\infty} \rho_H\left(D_n^{-1}\mathcal{H}_n, K\right) \le \varepsilon\right) &\ge \mathbb{P}\left(\rho_H\left(D_n^{-1}\mathcal{H}_n, K\right) < \varepsilon \text{ i.o.}\right) \\
&= \mathbb{P}\left(\bigcap_{n=1}^{\infty} \bigcup_{m\ge n} \left\{\rho_H(D_m^{-1}\mathcal{H}_m, K) < \varepsilon\right\}\right) \\
&= \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{m\ge n} \left\{\rho_H(D_m^{-1}\mathcal{H}_m, K) < \varepsilon\right\}\right) \\
&\ge \lim_{n\to\infty} \mathbb{P}\left(\rho_H(D_n^{-1}\mathcal{H}_n, K) < \varepsilon\right).
\end{aligned}$$

By the triangle inequality, $|\rho_H(K, K_1) - \rho_H(K, K_2)| \le \rho_H(K_1, K_2)$, i.e., for fixed K, the function $K_1 \mapsto \rho_H(K, K_1)$ is continuous. Thus by Lemma 4.2.7 and the continuous mapping theorem

$$\lim_{n\to\infty} \mathbb{P}\left(\rho_H(D_n^{-1}\mathcal{H}_n, K) < \varepsilon\right) = \mathbb{P}\left(\rho_H\left(\frac{\Sigma^{1/2}h_2}{\mathcal{D}(\Sigma^{1/2}h_2)}, K\right) < \varepsilon\right). \tag{4.2.7}$$

Let $\delta \in (0, \varepsilon/6)$. For convenience, set $A = \Sigma^{1/2}h_2$, note that A is not the normalised hull so we do not yet assume that $\mathcal{D}(A) = 1$. First suppose that $\mathbf{0}$ is in the interior of K. Then, it is not hard to see that $K \subseteq A \subseteq (1+\delta)K$ occurs with positive probability (one can force the Brownian motion to make a 'loop' in $((1+\delta)K) \setminus K$). On this event, we have $h_K(\mathbf{x}) \le h_A(\mathbf{x}) \le (1+\delta)h_K(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{S}$, so that, by (4.2.3),

$$\rho_H(A, K) = \sup_{\mathbf{x}\in\mathbb{S}} |h_A(\mathbf{x}) - h_K(\mathbf{x})| \le \delta \sup_{\mathbf{x}\in\mathbb{S}} h_K(\mathbf{x}) \le \delta\mathcal{D}(K) = \delta.$$

It follows from taking $K_1 = K$ and $K_2 = A$ in (4.2.4) that

$$\rho_H(A/\mathcal{D}(A), K) \le 3\rho_H(A, K) \le 3\delta < \varepsilon/2.$$

If $\mathbf{0}$ is not in the interior of K, then we can find $K' \in \mathfrak{C}_0^2$ with $K \subset K'$ such that $\mathbf{0}$ is in the interior of K' and $\rho_H(K, K') < \varepsilon/2$. Then

$$\rho_H(A/\mathcal{D}(A), K) \le \rho_H(A/\mathcal{D}(A), K') + \rho_H(K, K') < \varepsilon,$$

on the event $K' \subseteq A \subseteq (1+\delta)K'$, which has positive probability. Hence, in either case, the probability on the right-hand side of (4.2.7) is strictly positive, establishing (4.2.6).

□

Proof of Corollary 4.1.3. For $K \in \mathfrak{C}_0^2$, recall $\mathcal{L}(K)$ denotes the perimeter length of K; then, Lemma 2.4 of [WX15b] shows that

$$|\mathcal{L}(K_1) - \mathcal{L}(K_2)| \le 2\pi\rho_H(K_1, K_2), \text{ for any } K_1, K_2 \in \mathfrak{C}_0^2. \tag{4.2.8}$$

First, take K to be a unit-length line segment in $\mathbb{R}^2$ containing $\mathbf{0}$. Theorem 4.1.1 shows that, for any $\varepsilon > 0$, $\rho_H(D_n^{-1}\mathcal{H}_n, K) < \varepsilon$ i.o., a.s. Hence, by (4.2.8),

$$L_n/D_n = \mathcal{L}(D_n^{-1}\mathcal{H}_n) \le \mathcal{L}(K) + 2\pi\varepsilon, \text{ i.o.,}$$

and $\mathcal{L}(K) = 2$. Since $\varepsilon > 0$ was arbitrary, we get $\liminf_{n\to\infty} L_n/D_n \le 2$, and the first inequality in (2.1.4) shows that this latter inequality is in fact an equality.

Now take K to be a unit-diameter disc in $\mathbb{R}^2$ containing $\mathbf{0}$. Again, Theorem 4.1.1 shows that, for any $\varepsilon > 0$, $\rho_H(D_n^{-1}\mathcal{H}_n, K) < \varepsilon$ i.o., a.s. Hence, by (4.2.8),

$$L_n/D_n = \mathcal{L}(D_n^{-1}\mathcal{H}_n) \ge \mathcal{L}(K) - 2\pi\varepsilon, \text{ i.o.,}$$

and since now $\mathcal{L}(K) = \pi$ we get $\limsup_{n\to\infty} L_n/D_n \ge \pi$, which combined with the second inequality in (2.1.4) completes the proof. □

4.3 Application of results to our examples

The shape result is difficult to show in any static pictures from a simulation, however Corollary 4.1.3 can be shown by plotting the ratio L_n/D_n against n for both our simple symmetric random walk and the standard Normal random walk. These plots are below and, although they do not demonstrate many values near the extremities, they do show the ratio varying in time without the appearance of converging to a limit.

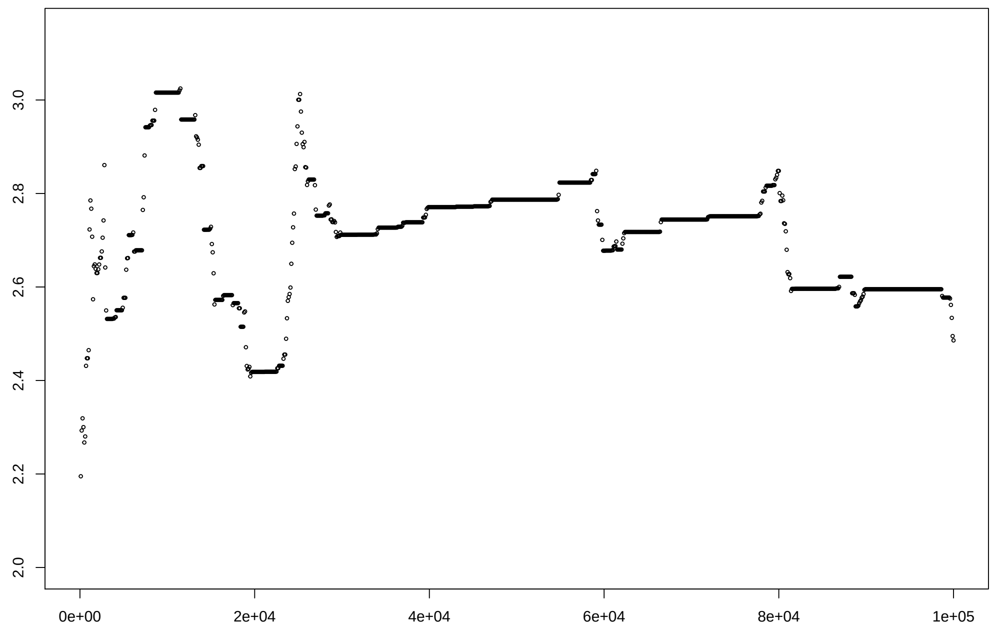


Figure 4.1: The ratio L_n/D_n plotted against n for $n = 1, \ldots, 10,000$ for our simple symmetric random walk.

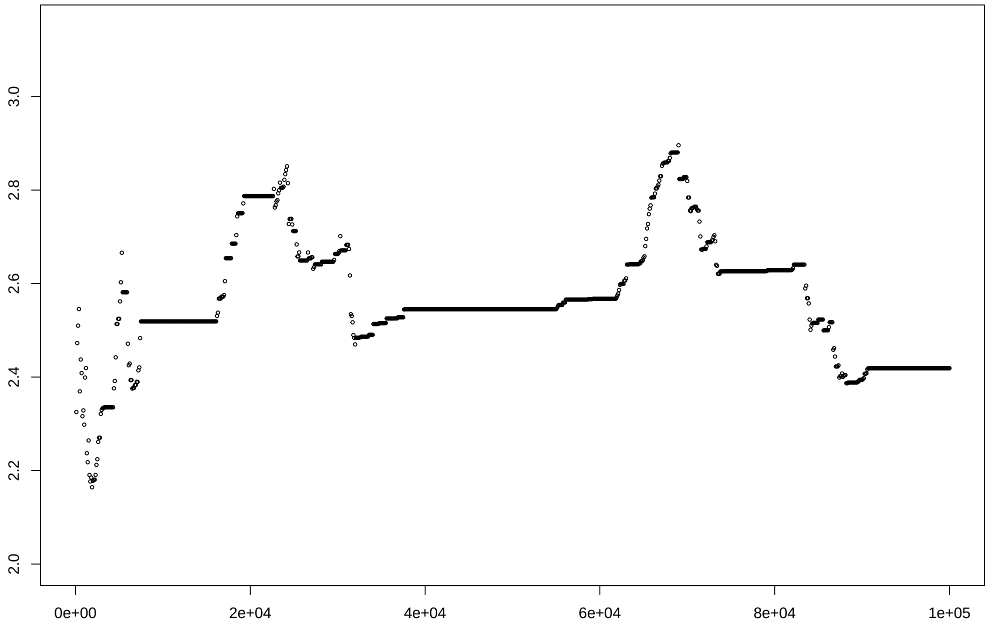


Figure 4.2: The ratio L_n/D_n plotted against n for $n = 1, \ldots, 10,000$ for our walk with Normal increments.

Chapter 5

Martingale difference method for diameter

So far, we have established both first and second order behaviour of functionals in the case with zero drift, but only have established the law of large numbers behaviour in the case with drift. Second order results for the perimeter length were established in [WX15a]. We will use the same method but adapt it for the diameter[1]. The method in question is the martingale difference method.

This method uses a sequence where the expected change is no change at all which gives useful formulae for the expectation and variance of related random variables, see Lemma 5.1.2 below or for example [Gut05, pp. 467–553] for an exposition of martingales and martingale difference sequences. As for their use in geometric probability theory, Steele [Ste90, p. 754] attributes the first use to Rhee and Talagrand [RT87] who apply the method to the travelling salesman problem. In turn, Rhee and Talagrand point towards some earlier uses in Banach space theory in particular referring to Milman and Schechtman [MS86]. Both Steele's and Rhee and Talagrand's use of the method uses a martingale difference sequence created by taking expectations having removed an element of the sequence of random variables, whilst we will use a slight modification whereby we resample (or replace after the removal) each random variable – this is the same as the method from [WX15a].

[1] Based on work published in [MW18], which was joint work between the authors.

We establish first Theorem 5.0.1, the analogue of Theorem 2.2.1, which states that the diameter is not far from the distance the walk travels in the direction of the mean.

Theorem 5.0.1. *Suppose that* $\mathbb{E}(\|Z\|^2) < \infty$ *and* $\mu \neq \mathbf{0}$*. Then, as* $n \to \infty$*,*

$$n^{-1/2}|D_n - S_n \cdot \hat{\mu}| \to 0, \text{ in } L^2. \tag{5.0.1}$$

Theorem 5.0.1 yields variance asymptotics and a central limit theorem when $\sigma_\mu^2 > 0$, as follows.

Corollary 5.0.2. *Suppose that* $\mathbb{E}(\|Z\|^2) < \infty$ *and* $\mu \neq \mathbf{0}$*. Then* $\lim_{n\to\infty} n^{-1} \mathbb{V}\mathrm{ar}\, D_n = \sigma_\mu^2$*. Moreover, if* $\sigma_\mu^2 > 0$*, for* $\zeta \sim \mathcal{N}(0,1)$*, as* $n \to \infty$*,*

$$\frac{D_n - \mathbb{E}\, D_n}{\sqrt{\mathbb{V}\mathrm{ar}\, D_n}} \xrightarrow{\mathrm{d}} \zeta, \text{ and } \frac{D_n - n\|\mu\|}{\sqrt{n\sigma_\mu^2}} \xrightarrow{\mathrm{d}} \zeta.$$

5.1 Diameter in the case with drift

The main aim of this section is to establish the following result, from which we will deduce Theorem 5.0.1.

Theorem 5.1.1. *Suppose that* $\mathbb{E}(\|Z\|^2) < \infty$ *and* $\mu \neq \mathbf{0}$*. Then, as* $n \to \infty$*,*

$$n^{-1/2}\,|D_n - \mathbb{E}\, D_n - (S_n - \mathbb{E}\, S_n) \cdot \hat{\mu}| \to 0, \text{ in } L^2. \tag{5.1.1}$$

Theorem 5.1.1 is the analogue for D_n of the result (1.1.4) for L_n, established in Theorem 1.3 of [WX15a]. Our approach to proving Theorem 5.1.1 is similar in outline to that in [WX15a], where a martingale difference idea (which we explain below in the present context) was combined with Cauchy's formula for the perimeter length. Here, the place of Cauchy's formula is taken by the formula

$$\operatorname{diam} A = \sup_{0\le\theta\le\pi} \rho_A(\theta), \tag{5.1.2}$$

where $A \subset \mathbb{R}^d$ is a non-empty compact set, and $\rho_A(\theta) := \sup_{\mathbf{x}\in A}(\mathbf{x}\cdot\mathbf{e}_\theta) - \inf_{\mathbf{x}\in A}(\mathbf{x}\cdot\mathbf{e}_\theta)$; see Lemma 3.4.6 for a derivation of (5.1.2).

Now we describe the martingale difference construction. Recall that $\mathcal{F}_0 := \{\emptyset, \Omega\}$ and $\mathcal{F}_n := \sigma(Z_1, \ldots, Z_n)$ for $n \ge 1$. Let $Z_1', Z_2', \ldots$ be an independent copy of the sequence

$Z_1, Z_2, \ldots$. Fix $n \in \mathbb{N}$. For $i \in \{1, \ldots, n\}$, define

$$S_j^{(i)} := \begin{cases} S_j & \text{if } j < i, \\ S_j - Z_i + Z_i' & \text{if } j \geq i; \end{cases}$$

then $(S_j^{(i)}; 0 \leq j \leq n)$ is the random walk $(S_j; 0 \leq j \leq n)$ but with Z_i 'resampled' and replaced by Z_i'. For $i \in \{1, \ldots, n\}$, define

$$D_n^{(i)} := \operatorname{diam}\{S_0^{(i)}, \ldots, S_n^{(i)}\}, \text{ and } \Delta_{n,i} := \mathbb{E}(D_n - D_n^{(i)} \mid \mathcal{F}_i). \tag{5.1.3}$$

Observe that we also have the representation $\Delta_{n,i} = \mathbb{E}(D_n \mid \mathcal{F}_i) - \mathbb{E}(D_n \mid \mathcal{F}_{i-1})$ and hence $\Delta_{n,i}$ is a martingale difference sequence, i.e., $\Delta_{n,i}$ is $\mathcal{F}_i$-measurable with $\mathbb{E}(\Delta_{n,i} \mid \mathcal{F}_{i-1}) = 0$. The utility of this construction is the following result (see e.g. Lemma 2.1 of [WX15a]).

Lemma 5.1.2. *Let $n \in \mathbb{N}$. Then $D_n - \mathbb{E}\, D_n = \sum_{i=1}^n \Delta_{n,i}$, and $\mathbb{V}\mathrm{ar}\, D_n = \sum_{i=1}^n \mathbb{E}(\Delta_{n,i}^2)$.*

Recall that $\mathbf{e}_\theta$ denotes the unit vector in direction θ. For $\theta \in [0, \pi]$, define

$$M_n(\theta) := \max_{0 \leq j \leq n} (S_j \cdot \mathbf{e}_\theta), \text{ and } m_n(\theta) := \min_{0 \leq j \leq n} (S_j \cdot \mathbf{e}_\theta),$$

and define $R_n(\theta) := M_n(\theta) - m_n(\theta)$. Note that since $S_0 = \mathbf{0}$, we have $M_n(\theta) \geq 0$ and $m_n(\theta) \leq 0$, a.s. It follows from (5.1.2) that $D_n = \sup_{0 \leq \theta \leq \pi} R_n(\theta)$.

Similarly, when the ith increment is resampled, $D_n^{(i)} = \sup_{0 \leq \theta \leq \pi} R_n^{(i)}(\theta)$, where $R_n^{(i)}(\theta) := M_n^{(i)}(\theta) - m_n^{(i)}(\theta)$, with

$$M_n^{(i)}(\theta) := \max_{0 \leq j \leq n} (S_j^{(i)} \cdot \mathbf{e}_\theta), \text{ and } m_n^{(i)}(\theta) := \min_{0 \leq j \leq n} (S_j^{(i)} \cdot \mathbf{e}_\theta).$$

Thus to study $\Delta_{n,i}$ as defined at (5.1.3), we are interested in

$$D_n - D_n^{(i)} = \sup_{0 \leq \theta \leq \pi} R_n(\theta) - \sup_{0 \leq \theta \leq \pi} R_n^{(i)}(\theta). \tag{5.1.4}$$

For the remainder of this section we suppose, without loss of generality, that $\mu = \|\mu\| \mathbf{e}_{\pi/2}$ with $\|\mu\| \in (0, \infty)$. An important observation is that the diameter does not deviate far from the direction of the drift. For $\delta \in (0, \pi/2)$ and $i \in \{1, \ldots, n\}$, define the event

$$A_{n,i}(\delta) := \left\{ \left| \frac{\pi}{2} - \arg\max_{0\leq\theta\leq\pi} R_n(\theta) \right| < \delta \right\} \cap \left\{ \left| \frac{\pi}{2} - \arg\max_{0\leq\theta\leq\pi} R_n^{(i)}(\theta) \right| < \delta \right\}.$$

Lemma 5.1.3. *Suppose that* $\mathbb{E}\,\|Z\| < \infty$ *and* $\mu = \|\mu\|\mathbf{e}_{\pi/2} \neq \mathbf{0}$. *Then for any* $\delta \in (0, \pi/2)$, $\lim_{n\to\infty} \min_{1\leq i\leq n} \mathbb{P}(A_{n,i}(\delta)) = 1$.

Proof. Fix $\delta \in (0, \pi/2)$. Note that $S_j \cdot \mathbf{e}_0$ is a random walk on $\mathbb{R}$ with mean increment $\mathbb{E}(Z \cdot \mathbf{e}_0) = \mu \cdot \mathbf{e}_0 = 0$. Hence the strong law of large numbers implies that for any $\varepsilon > 0$,

$$\max_{0\leq j\leq n} |S_j \cdot \mathbf{e}_0| \leq \varepsilon n,$$

for all $n \geq N_\varepsilon$ with $\mathbb{P}(N_\varepsilon < \infty) = 1$. Similarly, since $S_j \cdot \mathbf{e}_{\pi/2}$ is a random walk on $\mathbb{R}$ with mean increment $\|\mu\| > 0$, there exists N' with $\mathbb{P}(N' < \infty) = 1$ such that

$$S_j \cdot \mathbf{e}_{\pi/2} \geq \tfrac{1}{2}\|\mu\| j, \text{ for all } j \geq N'.$$

Let $A'_n(\varepsilon)$ denote the event

$$\left\{ \max_{0\leq j\leq n} |S_j \cdot \mathbf{e}_0| \leq \varepsilon n \right\} \cap \left\{ S_n \cdot \mathbf{e}_{\pi/2} \geq \tfrac{1}{2}\|\mu\| n \right\}.$$

Then if $A'_n(\varepsilon)$ occurs, any line segment that achieves the diameter has length at least $\frac{1}{2}\|\mu\|n$ and horizontal component at most $2\varepsilon n$. Thus if $\theta_n = \arg\max_{0\leq\theta\leq\pi} R_n(\theta)$ we have

$$|\cos\theta_n| \leq \frac{4\varepsilon}{\|\mu\|}, \text{ on } A'_n(\varepsilon).$$

Thus for ε sufficiently small we have that $A'_n(\varepsilon)$ implies $|\theta_n - \pi/2| < \delta$. Hence

$$\mathbb{P}(|\theta_n - \pi/2| < \delta) \geq \mathbb{P}(A'_n(\varepsilon)) \geq \mathbb{P}\left(n \geq \max\{N_\varepsilon, N'\}\right) \to \mathbb{P}\left(\max\{N_\varepsilon, N'\} < \infty\right) = 1.$$

But $\theta_n^{(i)} = \arg\max_{0\leq\theta\leq\pi} R_n^{(i)}(\theta)$ has the same distribution as θ_n, so

$$\min_{1\leq i\leq n} \mathbb{P}(\{|\theta_n - \pi/2| < \delta\} \cap \{|\theta_n^{(i)} - \pi/2| < \delta\}) \geq 1 - 2\mathbb{P}(|\theta_n - \pi/2| \geq \delta),$$

and the result follows. □

Lemma 5.1.3 tells us that the key to understanding (5.1.4) is to understand what is happening with $R_n(\theta)$ and $R_n^{(i)}(\theta)$ for $\theta \approx \pi/2$. The next important observation is that for $\theta \in (0, \pi)$, the one-dimensional random walk $S_j \cdot \mathbf{e}_\theta$ has drift $\mu \cdot \mathbf{e}_\theta = \mu \sin\theta > 0$, so, with very high probability $M_n(\theta)$ is attained somewhere near the end of the walk, and $m_n(\theta)$ somewhere near the start.

To formalize this statement, and its consequence for $R_n(\theta) - R_n^{(i)}(\theta)$, define

$$\bar{J}_n(\theta) := \arg\max_{0\le j\le n}(S_j \cdot \mathbf{e}_\theta), \text{ and } \underline{J}_n(\theta) := \arg\min_{0\le j\le n}(S_j \cdot \mathbf{e}_\theta);$$
$$\bar{J}_n^{(i)}(\theta) := \arg\max_{0\le j\le n}(S_j^{(i)} \cdot \mathbf{e}_\theta), \text{ and } \underline{J}_n^{(i)}(\theta) := \arg\min_{0\le j\le n}(S_j^{(i)} \cdot \mathbf{e}_\theta).$$

For $\gamma \in (0, 1/2)$ (a constant that will be chosen to be suitably small later in our argument), we denote by $E_{n,i}(\gamma)$ the event that the following occur:

- for all $\theta \in [\pi/4, 3\pi/4]$, $\underline{J}_n(\theta) < \gamma n$ and $\bar{J}_n(\theta) > (1-\gamma)n$;
- for all $\theta \in [\pi/4, 3\pi/4]$, $\underline{J}_n^{(i)}(\theta) < \gamma n$ and $\bar{J}_n^{(i)}(\theta) > (1-\gamma)n$;

note that the choice of interval $[\pi/4, 3\pi/4]$ could be replaced by any other interval containing $\pi/2$ and bounded away from 0 and π. Define $I_{n,\gamma} := \{1, \ldots, n\} \cap [\gamma n, (1-\gamma)n]$. The next result is contained in Lemma 4.1 of [WX15a].

Lemma 5.1.4. *For any $\gamma \in (0, 1/2)$ the following hold.*

(i) If $i \in I_{n,\gamma}$, then, on the event $E_{n,i}(\gamma)$,

$$R_n(\theta) - R_n^{(i)}(\theta) = (Z_i - Z_i') \cdot \mathbf{e}_\theta, \text{ for any } \theta \in [\pi/4, 3\pi/4]. \tag{5.1.5}$$

(ii) If $\mathbb{E}\,\|Z\| < \infty$ and $\mu \neq \mathbf{0}$ then $\lim_{n\to\infty} \min_{1\le i\le n} \mathbb{P}(E_{n,i}(\gamma)) = 1$.

In light of Lemma 5.1.3, the key to estimating (5.1.4) is provided by the following.

Lemma 5.1.5. *Let $\gamma \in (0, 1/2)$. Then for any $\delta \in (0, \pi/4)$ and any $i \in I_{n,\gamma}$, on $E_{n,i}(\gamma)$,*

$$\left| \sup_{|\theta-\pi/2|\le\delta} R_n(\theta) - \sup_{|\theta-\pi/2|\le\delta} R_n^{(i)}(\theta) - (Z_i - Z_i') \cdot \mathbf{e}_{\pi/2} \right| \le \delta \|Z_i - Z_i'\|.$$

Before proving Lemma 5.1.5, we need a simple geometrical lemma.

Lemma 5.1.6. *For any* $\mathbf{x} \in \mathbb{R}^2$ *and* $\theta_1, \theta_2 \in \mathbb{R}$,

$$|\mathbf{x} \cdot \mathbf{e}_{\theta_1} - \mathbf{x} \cdot \mathbf{e}_{\theta_2}| \leq \|\mathbf{x}\||\theta_1 - \theta_2|.$$

Proof. We have

$$\begin{aligned}\mathbf{e}_{\theta_1} - \mathbf{e}_{\theta_2} &= (\cos\theta_1 - \cos\theta_2, \sin\theta_1 - \sin\theta_2) \\ &= \left(-2\sin\left(\frac{\theta_1 - \theta_2}{2}\right)\sin\left(\frac{\theta_1 + \theta_2}{2}\right), 2\sin\left(\frac{\theta_1 - \theta_2}{2}\right)\cos\left(\frac{\theta_1 + \theta_2}{2}\right)\right),\end{aligned}$$

so that $\|\mathbf{e}_{\theta_1} - \mathbf{e}_{\theta_2}\|^2 = 4\sin^2\left(\frac{\theta_1 - \theta_2}{2}\right)$, and hence $\|\mathbf{e}_{\theta_1} - \mathbf{e}_{\theta_2}\| = 2\left|\sin\left(\frac{\theta_1 - \theta_2}{2}\right)\right|$. Now use the inequality $|\sin x| \leq |x|$ (valid for all $x \in \mathbb{R}$) to get

$$\|\mathbf{e}_{\theta_1} - \mathbf{e}_{\theta_2}\| \leq |\theta_1 - \theta_2|,$$

and the result follows. □

Proof of Lemma 5.1.5. We claim that with $i \in I_{n,\gamma}$, for any $\theta_1, \theta_2 \in [\pi/4, 3\pi/4]$, on the event $E_{n,i}(\gamma)$, it holds that

$$\inf_{\theta_1 \leq \theta \leq \theta_2} (Z_i - Z_i') \cdot \mathbf{e}_\theta \leq \sup_{\theta_1 \leq \theta \leq \theta_2} R_n(\theta) - \sup_{\theta_1 \leq \theta \leq \theta_2} R_n^{(i)}(\theta) \leq \sup_{\theta_1 \leq \theta \leq \theta_2} (Z_i - Z_i') \cdot \mathbf{e}_\theta. \quad (5.1.6)$$

Given the claim (5.1.6), and that, as follows from Lemma 5.1.6,

$$\begin{aligned}&\sup_{|\theta - \pi/2| \leq \delta} (Z_i - Z_i') \cdot \mathbf{e}_\theta \leq (Z_i - Z_i') \cdot \mathbf{e}_{\pi/2} + \delta\|Z_i - Z_i'\|, \text{ and} \\ &\inf_{|\theta - \pi/2| \leq \delta} (Z_i - Z_i') \cdot \mathbf{e}_\theta \geq (Z_i - Z_i') \cdot \mathbf{e}_{\pi/2} - \delta\|Z_i - Z_i'\|,\end{aligned}$$

the statement in the lemma follows on taking $\theta_1 = \pi/2 - \delta$ and $\theta_2 = \pi/2 + \delta$.

It remains to establish the claim (5.1.6). First we note that for $f, g : \mathbb{R} \to \mathbb{R}$ and I an interval in $(-\pi, \pi]$, with $\sup_{\theta \in I} |f(\theta)| < \infty$ and $\sup_{\theta \in I} |g(\theta)| < \infty$,

$$\inf_{\theta \in I}(f(\theta) - g(\theta)) \leq \sup_{\theta \in I} f(\theta) - \sup_{\theta \in I} g(\theta) \leq \sup_{\theta \in I}(f(\theta) - g(\theta)). \quad (5.1.7)$$

In particular, taking $I = [\theta_1, \theta_2]$, with $\theta_1, \theta_2 \in [\pi/3, 3\pi/4]$, we have

$$\inf_{\theta_1 \leq \theta \leq \theta_2} \left(R_n(\theta) - R_n^{(i)}(\theta)\right) \leq \sup_{\theta_1 \leq \theta \leq \theta_2} R_n(\theta) - \sup_{\theta_1 \leq \theta \leq \theta_2} R_n^{(i)}(\theta) \leq \sup_{\theta_1 \leq \theta \leq \theta_2} \left(R_n(\theta) - R_n^{(i)}(\theta)\right),$$

and, on the event $E_{n,i}(\gamma)$, we have from (5.1.5) that

$$R_n(\theta) - R_n^{(i)}(\theta) = (Z_i - Z_i') \cdot \mathbf{e}_\theta, \text{ for all } \theta \in [\theta_1, \theta_2],$$

which establishes the claim (5.1.6). □

To obtain rough estimates when the events $A_{n,i}(\delta)$ and $E_{n,i}(\gamma)$ do not occur, we need the following bound.

Lemma 5.1.7. *For any $i \in \{1, 2, \dots, n\}$, a.s.,*

$$|D_n^{(i)} - D_n| \leq 2\|Z_i\| + 2\|Z_i'\|.$$

Proof. Lemma 3.1 from [WX15a] states that, for any $i \in \{1, 2, \dots, n\}$, a.s.,

$$\sup_{0 \leq \theta \leq \pi} \left| R_n(\theta) - R_n^{(i)}(\theta) \right| \leq 2\|Z_i\| + 2\|Z_i'\|.$$

Now from (5.1.4) and (5.1.7) we obtain the result. □

Now define the event $B_{n,i}(\gamma, \delta) := E_{n,i}(\gamma) \cap A_{n,i}(\delta)$. Let $B_{n,i}^{\mathrm{c}}(\gamma, \delta)$ denote the complementary event. The preceding results in this section can now be combined to obtain the following approximation lemma for $\Delta_{n,i}$ as given by (5.1.3).

Lemma 5.1.8. *Suppose that $\mathbb{E}\|Z\| < \infty$ and $\mu \neq \mathbf{0}$. For any $\gamma \in (0, 1/2)$, $\delta \in (0, \pi/4)$, and $i \in I_{n,\gamma}$, we have, a.s.,*

$$|\Delta_{n,i} - (Z_i - \mu) \cdot \hat{\mu}| \leq 3\|Z_i\| \mathbb{P}(B_{n,i}^{\mathrm{c}}(\gamma, \delta) \mid \mathcal{F}_i) + 3\,\mathbb{E}[\|Z_i'\| \mathbf{1}(B_{n,i}^{\mathrm{c}}(\gamma, \delta)) \mid \mathcal{F}_i] \\ + \delta \left(\|Z_i\| + \mathbb{E}\|Z\| \right).$$

Proof. First observe that, since Z_i is $\mathcal{F}_i$-measurable and Z_i' is independent of $\mathcal{F}_i$,

$$\Delta_{n,i} - (Z_i - \mu) \cdot \hat{\mu} = \mathbb{E}[D_n - D_n^{(i)} - (Z_i - Z_i') \cdot \hat{\mu} \mid \mathcal{F}_i].$$

Hence, by the triangle inequality,

$$|\Delta_{n,i} - (Z_i - \mu) \cdot \hat{\mu}| \leq \mathbb{E}\left[\left| D_n - D_n^{(i)} - (Z_i - Z_i') \cdot \hat{\mu} \right| \mathbf{1}(B_{n,i}(\gamma, \delta)) \mid \mathcal{F}_i \right] \\ + \mathbb{E}\left[\left| D_n - D_n^{(i)} - (Z_i - Z_i') \cdot \hat{\mu} \right| \mathbf{1}(B_{n,i}^{\mathrm{c}}(\gamma, \delta)) \mid \mathcal{F}_i \right].$$

Here, by Lemma 5.1.7, we have that

$$\mathbb{E}\left[\left|D_n - D_n^{(i)} - (Z_i - Z_i') \cdot \hat{\mu}\right| \mathbf{1}(B_{n,i}^{\mathrm{c}}(\gamma,\delta)) \mid \mathcal{F}_i\right] \\ \leq 3\,\mathbb{E}\left[(\|Z_i\| + \|Z_i'\|)\mathbf{1}(B_{n,i}^{\mathrm{c}}(\gamma,\delta)) \mid \mathcal{F}_i\right].$$

Now, on $A_{n,i}(\delta)$ we have that

$$D_n = \sup_{|\theta-\pi/2|\leq\delta} R_n(\theta), \text{ and } D_n^{(i)} = \sup_{|\theta-\pi/2|\leq\delta} R_n^{(i)}(\theta),$$

and hence, by Lemma 5.1.5, on $A_{n,i}(\delta) \cap E_{n,i}(\gamma)$,

$$|D_n - D_n^{(i)} - (Z_i - Z_i') \cdot \hat{\mu}| \leq \delta \|Z_i - Z_i'\|.$$

Hence

$$\mathbb{E}\left[\left|D_n - D_n^{(i)} - (Z_i - Z_i') \cdot \hat{\mu}\right| \mathbf{1}(B_{n,i}(\gamma,\delta)) \mid \mathcal{F}_i\right] \leq \delta\,\mathbb{E}[\|Z_i\| + \|Z_i'\| \mid \mathcal{F}_i].$$

Combining these bounds, and using the fact that Z_i is $\mathcal{F}_i$-measurable and Z_i' is independent of $\mathcal{F}_i$, we obtain the result. □

We are now almost ready to complete the proof of Theorem 5.1.1. To do so, we present an analogue of Lemma 6.1 from [WX15a]; we set $V_i := (Z_i - \mu) \cdot \hat{\mu}$, and $W_{n,i} := \Delta_{n,i} - V_i$.

Lemma 5.1.9. *Suppose that* $\mathbb{E}(\|Z\|^2) < \infty$ *and* $\mu \neq \mathbf{0}$*. Then*

$$\lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} \mathbb{E}(W_{n,i}^2) = 0.$$

Proof. The proof is similar to that of Lemma 6.1 of [WX15a]. Fix $\varepsilon \in (0,1)$. Take $\gamma \in (0,1/2)$ and $\delta \in (0,\pi/4)$, to be specified later. Note that from Lemma 5.1.7 we have $|W_{n,i}| \leq 3(\|Z_i\| + \mathbb{E}\,\|Z\|)$, so that, provided $\mathbb{E}(\|Z\|^2) < \infty$, we have $\mathbb{E}(W_{n,i}^2) \leq C_0$ for all n and all i, for some constant $C_0 < \infty$, depending only on the distribution of Z. Hence

$$\frac{1}{n} \sum_{i \notin I_{n,\gamma}} \mathbb{E}(W_{n,i}^2) \leq 2\gamma C_0.$$

From now on choose and fix $\gamma > 0$ small enough so that $2\gamma C_0 < \varepsilon$.

Now consider $i \in I_{n,\gamma}$. For such i, Lemma 5.1.8 yields an upper bound for $|W_{n,i}|$. Note that, for any $C_1 < \infty$, since Z_i' is independent of $\mathcal{F}_i$,

$$\mathbb{E}[\|Z_i'\|\mathbf{1}(B_{n,i}^{\mathrm{c}}(\gamma,\delta)) \mid \mathcal{F}_i] \leq \mathbb{E}[\|Z\|\mathbf{1}\{\|Z\| \geq C_1\}] + C_1\mathbb{P}[B_{n,i}^{\mathrm{c}}(\gamma,\delta) \mid \mathcal{F}_i].$$

Given $\varepsilon \in (0,1)$ we can take $C_1 = C_1(\varepsilon)$ large enough such that $\mathbb{E}[\|Z\|\mathbf{1}\{\|Z\| \geq C_1\}] \leq \varepsilon$, by dominated convergence; for convenience we take $C_1 > 1$ and $C_1 > \mathbb{E}\,\|Z\|$. Hence from Lemma 5.1.8 we obtain

$$|W_{n,i}| \leq 3(\|Z_i\| + C_1)\mathbb{P}[B_{n,i}^{\mathrm{c}}(\gamma,\delta) \mid \mathcal{F}_i] + 3\varepsilon + \delta\left(\|Z_i\| + \mathbb{E}\,\|Z\|\right).$$

Using the fact that $\mathbb{P}[B_{n,i}^{\mathrm{c}}(\gamma,\delta) \mid \mathcal{F}_i] \leq 1$, $\varepsilon \leq 1$, $\delta \leq 1$, and $C_1 > 1$, $C_1 > \mathbb{E}\,\|Z\|$, we can square both sides of the last display and collect terms to obtain

$$W_{n,i}^2 \leq 27C_1^2(1 + \|Z_i\|)^2\mathbb{P}[B_{n,i}^{\mathrm{c}}(\gamma,\delta) \mid \mathcal{F}_i] + 9\varepsilon + 13C_1^2\delta\left(1 + \|Z_i\|\right)^2.$$

Since $\mathbb{E}(\|Z\|^2) < \infty$, it follows that, given ε and hence C_1, we can choose $\delta \in (0,\pi/4)$ sufficiently small so that $13C_1^2\delta\,\mathbb{E}[(1+\|Z_i\|)^2] < \varepsilon$; fix such a δ from now on. Then

$$\mathbb{E}(W_{n,i}^2) \leq 27C_1^2\,\mathbb{E}[(1 + \|Z_i\|)^2\mathbb{P}[B_{n,i}^{\mathrm{c}}(\gamma,\delta) \mid \mathcal{F}_i]] + 10\varepsilon.$$

Here we have that, for any $C_2 > 0$,

$$\mathbb{E}[(1 + \|Z_i\|)^2\mathbb{P}[B_{n,i}^{\mathrm{c}}(\gamma,\delta) \mid \mathcal{F}_i]] \leq (1 + C_2)^2\mathbb{P}(B_{n,i}^{\mathrm{c}}(\gamma,\delta)) + \mathbb{E}[(1 + \|Z\|)^2\mathbf{1}\{\|Z\| \geq C_2\}],$$

where dominated convergence shows that we may choose C_2 large enough so that the last term is less than ε/C_1^2, say. Then,

$$\mathbb{E}(W_{n,i}^2) \leq 37\varepsilon + 27C_1^2(1 + C_2)^2\mathbb{P}(B_{n,i}^{\mathrm{c}}(\gamma,\delta)).$$

Finally, we see from Lemmas 5.1.3 and 5.1.4 than $\max_{1\leq i\leq n}\mathbb{P}(B_{n,i}^{\mathrm{c}}(\gamma,\delta)) \to 0$, so that, for given $\varepsilon > 0$ (and hence C_1 and C_2) we may choose $n \geq n_0$ sufficiently large so that $\max_{i\in I_{n,\gamma}}\mathbb{E}(W_{n,i}^2) \leq 38\varepsilon$. Hence

$$\frac{1}{n}\sum_{i\in I_{n,\gamma}}\mathbb{E}(W_{n,i}^2) \leq 38\varepsilon,$$

for all $n \geq n_0$. Combining this result with the estimate for $i \notin I_{n,\gamma}$, we see that

$$\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(W_{n,i}^2) \leq 39\varepsilon,$$

for all $n \geq n_0$. Since $\varepsilon > 0$ was arbitrary, the result follows. □

Proof of Theorem 5.1.1. First note that $W_{n,i}$ is $\mathcal{F}_i$-measurable with $\mathbb{E}(W_{n,i} \mid \mathcal{F}_{i-1}) = \mathbb{E}(\Delta_{n,i} \mid \mathcal{F}_{i-1}) - \mathbb{E} V_i = 0$, so that $W_{n,i}$ is a martingale difference sequence. Therefore by orthogonality, $n^{-1} \mathbb{E}[(\sum_{i=1}^n W_{n,i})^2] = n^{-1} \sum_{i=1}^n \mathbb{E}(W_{n,i}^2) \to 0$ as $n \to \infty$, by Lemma 5.1.9. In other words, $n^{-1/2} \sum_{i=1}^n W_{n,i} \to 0$ in L^2. But, by Lemma 5.1.2,

$$\sum_{i=1}^n W_{n,i} = \sum_{i=1}^n \Delta_{n,i} - \sum_{i=1}^n (Z_i - \mu) \cdot \hat{\mu} = D_n - \mathbb{E} D_n - (S_n - \mathbb{E} S_n) \cdot \hat{\mu}.$$

This yields the statement in the theorem. □

Finally we can give the proof of Theorem 5.0.1.

Proof of Theorem 5.0.1. Lemma 2.2.3 shows that

$$n^{-1/2} |\mathbb{E} D_n - \mathbb{E} S_n \cdot \hat{\mu}| \to 0. \tag{5.1.8}$$

Then by the triangle inequality

$$n^{-1/2}|D_n - S_n \cdot \hat{\mu}| \leq n^{-1/2}|D_n - \mathbb{E} D_n - (S_n - \mathbb{E} S_n) \cdot \hat{\mu}| + n^{-1/2}|\mathbb{E} D_n - \mathbb{E} S_n \cdot \hat{\mu}|,$$

which tends to 0 in L^2 by (5.1.1) and (5.1.8). □

Proof of Corollary 5.0.2. Corollary 5.0.2 is deduced from Theorem 5.0.1 in a very similar manner to how Theorems 1.1 and 1.2 in [WX15a] were deduced from Theorem 1.3 there, so we omit the details. □

5.2 Application of results to our examples

The first result and idea we would like to see in action is from Theorem 5.0.1. We hope to see that the difference between the diameter and distance in the direction of the mean does not grow as fast as $n^{1/2}$. To demonstrate this, Figures 5.1 and 5.2 are plots of the left hand side of (5.0.1) against n for our random walk with drift and all coordinates Normally distributed and our random walk with drift and no variance in the first coordinate, both with unit mean. Note, of course, $D_n \geq S_n \cdot \hat{\mu}$ so all of these values must be bounded below by 0.

The case with the Normal increments does seem to show a decreasing amount of points far away from 0 but the process doesn't look like a smooth path unlike the increments where the jump in the direction of the mean is fixed. This is not so surprising because, at least heuristically, the walk with Normal increments is more likely to attain the diameter by a distance induced by two points other than the first at $\mathbf{0}$ and S_n. Since the Normal increments allow for jumps back towards the origin which could be into the interior of the hull, when these types of jump occur, the difference plotted $D_n - S_n \cdot \hat{\mu}$ will be non-zero. Whereas the fixed increments can't move towards the origin horizontally, so any differences between D_n and $S_n \cdot \hat{\mu}$ must be created by movement in the vertical direction which is a relatively smooth process.

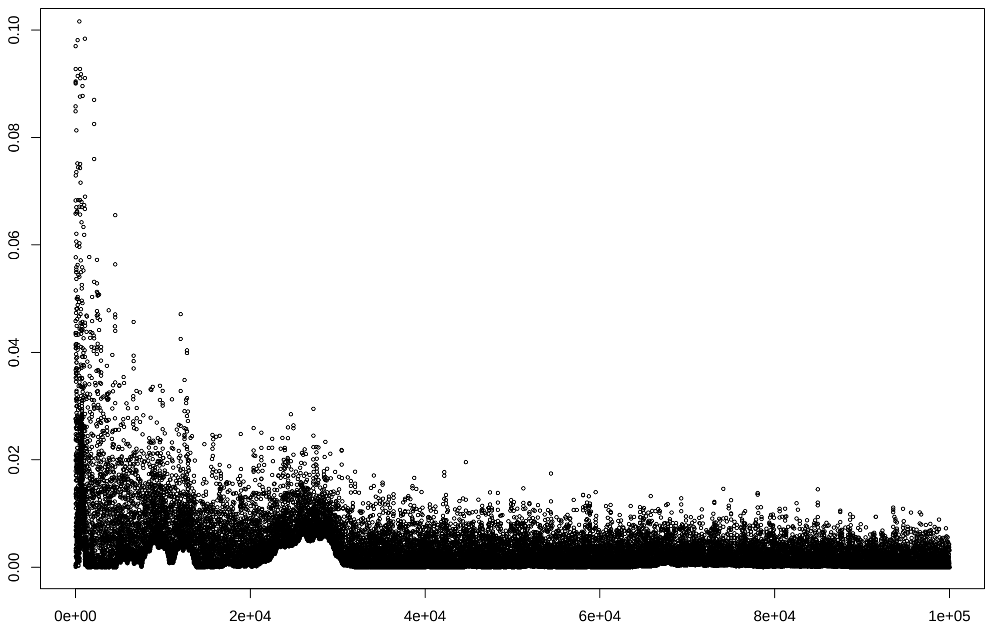


Figure 5.1: Difference between the diameter and length of the vector created by projecting the endpoint of the walk onto the direction of the mean $(D_n - S_n \cdot \hat{\mu})$ for our random walk with drift and all coordinates Normally distributed, unit mean, plotted for the first 10^5 steps.

For Corollary 5.0.2, we simulated $10,000$ independent copies of the random walks with Normal drift, unit mean and have plotted the empirical distribution of $D_n - n\|\mu\|/\sqrt{n\sigma_\mu^2}$ by taking the simulated values of this quantity when each walk reached $n = 100,000$. This is Figure 5.3. Note, we only show the range $[-2, 2]$. By visual inspection alone the plot looks like it is following the standard Normal distribution. Further, 448 of the

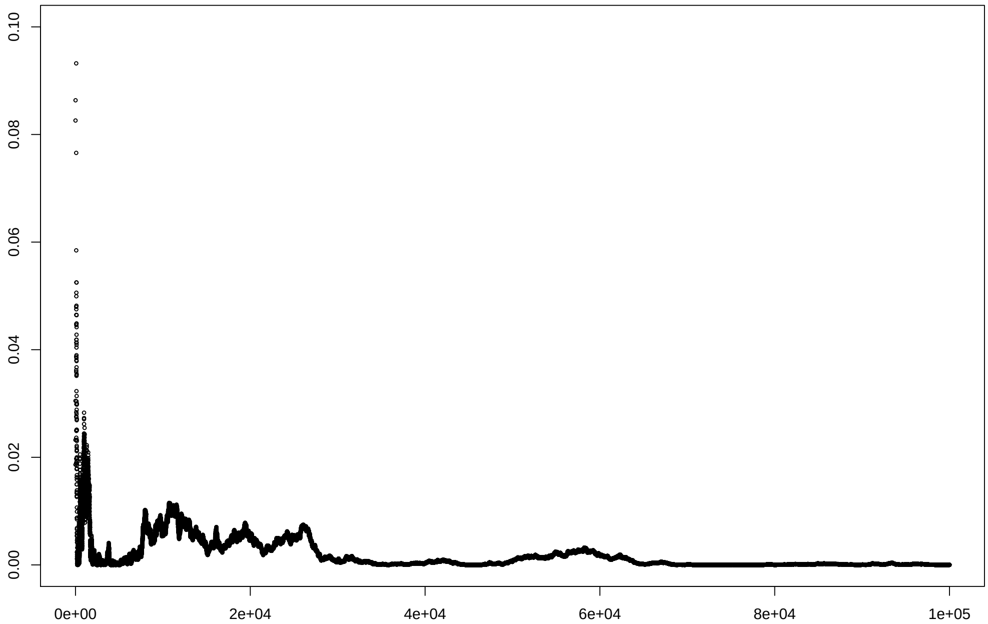


Figure 5.2: Difference between the diameter and length of the vector created by projecting the endpoint of the walk onto the direction of the mean $(D_n - S_n \cdot \hat{\mu})$ for our random walk with drift and no variance in the first coordinate, unit mean, plotted for the first 10^5 steps.

walks attained values outside the plotted range which is consistent with the well-known fact that 95% of the probability mass of a standard Normal distribution is between $[-1.96, 1.96]$. Again we can verify this using a simplified version of the Kolmogorov-Smirnov distance as described at (3.5.1). Here though we will take $F(x) = \Phi(x)$ where $\Phi(x)$ is the cumulative distribution function of the standard Normal distribution, and we also slightly modify our definition of the test to run over the range $[-2, 2]$ as follows

$$\rho^k_{K-S}(F_{test}, F) = \sup_{0 \le i \le k} |\mathbb{P}(F_{test} \le -2 + 4i/k) - F(-2 + 4i/k)|.$$

Using this definition, we will take $k = 80$ so that we are again in line with the binning shown in our empirical distribution bar chart, Figure 5.3, and using the notation D_{emp} to represent the empirical distribution, we find $\rho^{80}_{K-S}(D_{emp}, F) = 0.023$. This certainly supports the central limit theorem of Corollary 5.0.2.

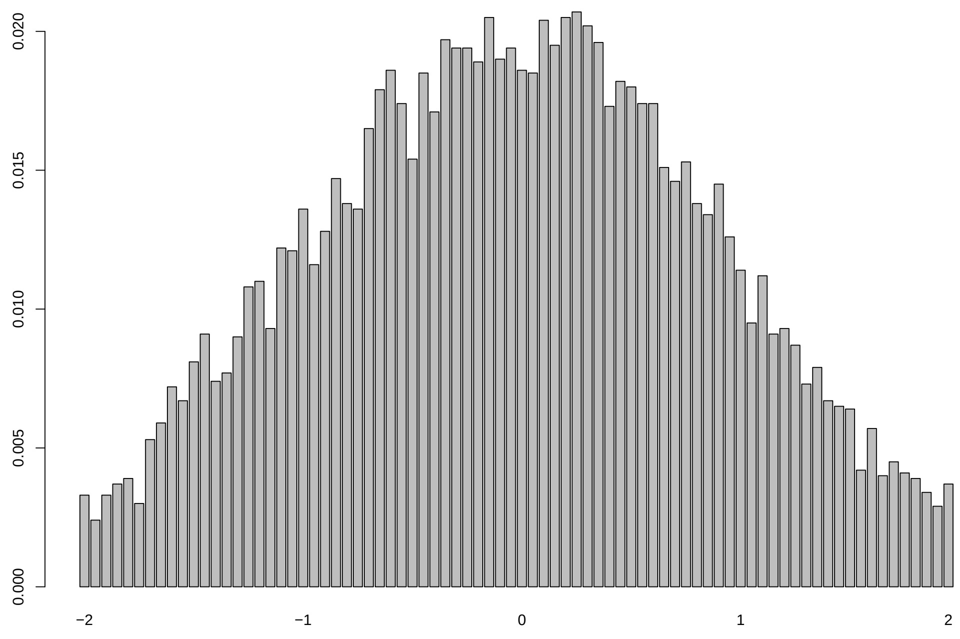


Figure 5.3: Empirical distribution of the centralised and normalised diameter for the random walk with drift and all coordinates Normally distributed, unit mean.

Chapter 6

Time-space processes

The final chapter of results fills in some of the details of the degenerate case, where there is no variation in the direction of the mean, $\sigma_\mu^2 = 0$, which corresponds to the case where $Z \cdot \hat{\mu} = \|\mu\|$ a.s., and is of its own interest. It includes, for example, the case where $Z = (1, 1)$ or $(1, -1)$, each with probability $1/2$, in which the two-dimensional walk S_n corresponds to the space-time diagram of a one-dimensional simple symmetric random walk. In fact, all of the processes in this degenerate case can be considered as time-space processes.

Regarding the diameter, in the case $\sigma_\mu^2 = 0$ Corollary 5.0.2 says only that $\mathbb{V}\mathrm{ar}\, D_n = o(n)$. We will show that the variance converges to a constant and determine the limiting distribution of the centred diameter too[1]. These results require some additional conditions. For the perimeter length we do not obtain the limiting distribution but do show that $\mathbb{V}\mathrm{ar}\, L_n = o(n^\varepsilon)$ for any $\varepsilon > 0$ and conjecture that $\mathbb{V}\mathrm{ar}\, L_n = O(\log n)$. In studying the perimeter length, the heuristic was motivated by consideration of the number of faces, F_n, of the convex hull, and we use results regarding the expected number of faces in our proof. Whilst considering this heuristic, we establish a partial classification of when $\liminf_{n\to\infty} F_n = 1$ and when $\lim_{n\to\infty} F_n = \infty$ which is to be contrasted with the results of Qiao and Steele, which stated that there exists a random walk for which $F_n = 1$ finitely often. In the case where $\lim_{n\to\infty} F_n = \infty$ we go further by showing that $F_n = O(\log n)$ with bounds on the possible constant in the growth

[1]Based on work published in [MW18], which was joint work between the authors.

rate, if such a constant exists.

6.1 Diameter limit distribution

In this section we will prove the following theorem.

Theorem 6.1.1. *Suppose that* $\mathbb{E}(\|Z\|^p) < \infty$ *for some* $p > 2$, $\mu \neq \mathbf{0}$, *and* $\sigma_\mu^2 = 0$. *Then,*

$$D_n - \|\mu\| n \xrightarrow{\text{d}} \frac{\sigma_{\mu_\perp}^2 \zeta^2}{2\|\mu\|}, \tag{6.1.1}$$

where $\zeta \sim \mathcal{N}(0, 1)$. *Further, if, in addition,* $\mathbb{E}(\|Z\|^p) < \infty$ *for some* $p > 4$, *then*

$$\lim_{n \to \infty} \mathbb{V}\text{ar}\, D_n = \frac{\sigma_{\mu_\perp}^4}{2\|\mu\|^2}. \tag{6.1.2}$$

Remark 6.1.2.

(i) The higher moments conditions required in Theorem 6.1.1 are necessary for the proofs that we employ, however we propose that $2 + \varepsilon$ moment should suffice; see also Remark 6.1.7 below.

(ii) The statement (6.1.1) may be written as

$$D_n - S_n \cdot \hat{\mu} \xrightarrow{\text{d}} \frac{\sigma_{\mu_\perp}^2 \zeta^2}{2\|\mu\|}. \tag{6.1.3}$$

It is natural to ask whether (6.1.3) also holds in the case where $\sigma_\mu^2 > 0$; if it did, then it would provide an alternative proof of the central limit theorem in Corollary 5.0.2. Simulations suggest that when $\sigma_\mu^2 > 0$, equation (6.1.3) holds in some, but not all cases, see Section 6.3.

The aim of this section is to prove Theorem 6.1.1; thus we assume $\mu \neq \mathbf{0}$. An important result for the proof is the following lemma which is interesting in its own right in that it not only confirms the intuition that the diameter is close to $\|S_n\|$ but seems particularly strong in that it states the difference between these values converges to 0. We present the proof of this lemma later in the chapter.

Lemma 6.1.3. *Suppose that* $\mathbb{E}(\|Z\|^p) < \infty$ *for some* $p > 2$, $\mu \neq \mathbf{0}$, *and* $\sigma_\mu^2 = 0$. *Then, as* $n \to \infty$, $D_n - \|S_n\| \to 0$, *a.s.*

The first step in the proof is to state another result that will enable us to obtain the second statement in Theorem 6.1.1 from the first.

Lemma 6.1.4. *Suppose that* $\mathbb{E}(\|Z\|^p) < \infty$ *for some* $p > 4$, $\mu \neq \mathbf{0}$, *and* $\sigma_\mu^2 = 0$. *Then* $(D_n - \|\mu\|n)^2$ *is uniformly integrable.*

Again, we write $X_n := S_n \cdot \hat{\mu}$ and $Y_n := S_n \cdot \hat{\mu}_\perp$, where $\hat{\mu}_\perp$ is any fixed unit vector orthogonal to μ. Note that if $\sigma_\mu^2 = 0$, then $X_n = n\|\mu\|$ is deterministic.

Proof of Lemma 6.1.4. For $i \leq j$, we have $\|S_j - S_i\|^2 = (Y_j - Y_i)^2 + (X_j - X_i)^2$, so that

$$\begin{aligned}(D_n - \|\mu\|n)^2 &= \left(\max_{0\leq i\leq j\leq n}\left((Y_j - Y_i)^2 + \|\mu\|^2(j-i)^2\right)^{1/2} - \|\mu\|n\right)^2 \\ &\leq \left(\|\mu\|n \max_{0\leq i\leq j\leq n}\left(1 + \frac{(Y_j - Y_i)^2}{\|\mu\|^2 n^2}\right)^{1/2} - \|\mu\|n\right)^2.\end{aligned}$$

Since $(1+y)^{1/2} \leq 1 + (y/2)$ for $y \geq 0$, and $(a-b)^2 \leq 2(a^2+b^2)$ for $a, b \in \mathbb{R}$, we obtain

$$(D_n - \|\mu\|n)^2 \leq \left(\|\mu\|n \max_{0\leq i\leq j\leq n}\frac{(Y_j - Y_i)^2}{2\|\mu\|^2 n^2}\right)^2 \leq \frac{4}{\|\mu\|^2}\max_{1\leq i\leq n}\frac{Y_i^4}{n^2}.$$

Now, $|Y_n|$ is a non-negative submartingale, so Doob's L^p inequality [Gut05, p. 505] yields

$$\mathbb{E}\left[\left(\max_{1\leq i\leq n}\frac{Y_i^4}{n^2}\right)^{p/4}\right] = n^{-p/2}\,\mathbb{E}\left(\max_{1\leq i\leq n}|Y_i|^p\right) \leq C_p n^{-p/2}\,\mathbb{E}(|Y_n|^p),$$

for any $p > 1$ and some constant $C_p < \infty$. Under the assumption that $\mathbb{E}(\|Z\|^p) < \infty$ for $p > 4$, Y_n is a random walk on $\mathbb{R}$ whose increments have zero mean and finite pth moments, so, by the Marcinkiewicz–Zygmund inequality, see Theorem 1.3.18 or e.g. [Gut05, p. 151], $\mathbb{E}(|Y_n|^p) \leq Cn^{p/2}$. Hence

$$\sup_{n\geq 0}\mathbb{E}\left[\left((D_n - \|\mu\|n)^2\right)^{p/4}\right] < \infty,$$

which, since $p/4 > 1$, establishes uniform integrability. □

Next we show that, under the conditions of Theorem 6.1.1, the diameter must be attained by a point 'close to' the start and one 'close to' the end of the walk.

Lemma 6.1.5. *Suppose that* $\mathbb{E}(\|Z\|^2) < \infty$, $\mu \neq \mathbf{0}$, *and* $\sigma_\mu^2 = 0$. *Let* $\beta \in (0,1)$. *Then, a.s., for all but finitely many* n,

$$D_n = \max_{\substack{0 \le i \le n^\beta \\ n-n^\beta \le j \le n}} \|S_j - S_i\|.$$

Proof. Fix $\beta \in (0,1)$. Since $D_n = \max_{0 \le i,j \le n} \|S_j - S_i\|$, we have

$$D_n = \max \left\{ \max_{\substack{0 \le i \le n^\beta \\ n-n^\beta \le j \le n}} \|S_j - S_i\|, \max_{\substack{0 \le i \le n^\beta \\ 0 \le j \le n-n^\beta}} \|S_j - S_i\|, \max_{n^\beta \le i,j \le n} \|S_j - S_i\| \right\}. \tag{6.1.4}$$

It is clear that

$$\max_{\substack{0 \le i \le n^\beta \\ n-n^\beta \le j \le n}} \|S_j - S_i\| \ge \|S_n\| \ge |X_n| = \|\mu\| n.$$

We aim to show that the other two terms on the right-hand side of (6.1.4) are strictly less than $\|\mu\|n$ for all but finitely many n.

A consequence of the law of the iterated logarithm, see Theorem 1.3.16, is that, for any $\varepsilon > 0$, a.s., for all but finitely many n, $\max_{0 \le i \le n} Y_i^2 \le n^{1+\varepsilon}$; see e.g. [Gut05, p. 384]. Take $\varepsilon \in (0, \beta)$. Then,

$$\begin{aligned}
\max_{\substack{0 \le i \le n^\beta \\ 0 \le j \le n-n^\beta}} \|S_j - S_i\|^2 &\le \max_{\substack{0 \le i \le n^\beta \\ 0 \le j \le n-n^\beta}} |X_j - X_i|^2 + \max_{\substack{0 \le i \le n^\beta \\ 0 \le j \le n-n^\beta}} |Y_j - Y_i|^2 \\
&\le \|\mu\|^2 (n - n^\beta)^2 + \max_{0 \le j \le n-n^\beta} Y_j^2 + \max_{0 \le i \le n^\beta} Y_i^2 + 2 \max_{\substack{0 \le i \le n^\beta \\ 0 \le j \le n-n^\beta}} |Y_j||Y_i| \\
&\le \|\mu\|^2 n^2 - 2\|\mu\|^2 n^{1+\beta} + \|\mu\|^2 n^{2\beta} + n^{1+\varepsilon},
\end{aligned}$$

for all but finitely many n. Since $\varepsilon < \beta < 1$, this last expression is strictly less than $\|\mu\|^2 n^2$ for all n sufficiently large. Similarly,

$$\begin{aligned}
\max_{n^\beta \le i,j \le n} \|S_j - S_i\|^2 &\le \|\mu\|^2 (n - n^\beta)^2 + \max_{n^\beta \le j \le n} Y_j^2 + \max_{n^\beta \le i \le n} Y_i^2 + 2 \max_{n^\beta \le i,j \le n} |Y_j||Y_i| \\
&\le \|\mu\|^2 n^2 - 2\|\mu\|^2 n^{1+\beta} + \|\mu\|^2 n^{2\beta} + n^{1+\varepsilon},
\end{aligned}$$

for all but finitely many n, and, as before, this is strictly less than $\|\mu\|^2 n^2$ for all n sufficiently large. Then (6.1.4) yields the result. □

The next result is required to control the fluctuations in the last part of the walk and can be considered as a technical result which will help us prove the more intuitive result, Lemma 6.1.3 below.

Lemma 6.1.6. *Let* $\xi, \xi_1, \xi_2, \dots$ *be i.i.d. random variables with* $\mathbb{E}(|\xi|^p) < \infty$ *for some* $p > 2$*, and* $\mathbb{E}\,\xi = 0$*. For* $0 \le j \le n$*, let* $T_{n,j} := \sum_{i=n-j}^{n} \xi_i$*. Then there exist* $\beta_0 \in (0, 1/2)$ *and* $\varepsilon_0 \in (0, 1/2)$ *such that for any* $\beta \in (0, \beta_0)$ *and any* $\varepsilon \in (0, \varepsilon_0)$*,*

$$\lim_{n\to\infty} \max_{0\le j\le n^\beta} \frac{|T_{n,j}|}{n^{(1/2)-\varepsilon}} = 0, \ \ a.s.$$

Remark 6.1.7. On first sight, by the fact that there are $O(n^\beta)$ terms in the sum $T_{n,j}$, one's intuition may be misled to conclude that $T_{n,j}$ should be only of size about $n^{\beta/2}$. However, note that assuming only $\mathbb{E}(\xi^2) < \infty$, $\max_{0\le i\le n} \xi_i$ can be almost as big as $n^{1/2}$, and with probability at least $1/n$ this maximal value is a member of $T_{n,j}$, and so it seems reasonable to expect that $T_{n,j}$ should be almost as big as $n^{1/2}$ infinitely often. Thus our $p > 2$ moments condition seems to be necessary.

Proof of Lemma 6.1.6. Let $\xi_i' = \xi_i \mathbf{1}\{|\xi_i| \le i^{1/2-\delta}\}$ and $\xi_i'' = \xi_i \mathbf{1}\{|\xi_i| > i^{1/2-\delta}\}$ for some $\delta \in (0, 1/2)$ to be chosen later. Then we use the subadditivity of the supremum, the triangle inequality, and the condition $\varepsilon \in (0, \varepsilon_0)$ to get

$$\max_{0\le j\le n^\beta} \frac{|T_{n,j}|}{n^{1/2-\varepsilon}} \le \max_{0\le j\le n^\beta} \frac{|\sum_{i=n-j}^{n}(\xi_i' - \mathbb{E}\,\xi_i')|}{n^{1/2-\varepsilon}} + \frac{\sum_{i=n-n^\beta}^{n} |\mathbb{E}\,\xi_i'|}{n^{1/2-\varepsilon_0}} + \frac{\sum_{i=n-n^\beta}^{n} |\xi_i''|}{n^{1/2-\varepsilon_0}}, \quad (6.1.5)$$

where, and for the rest of this proof, if n^β appears in the index of a sum, we understand it to be shorthand for $\lfloor n^\beta \rfloor$. By Markov's inequality, since $\mathbb{E}(|\xi|^p) < \infty$ for $p > 2$ we have

$$\mathbb{P}\left(|\xi_i| > i^{1/2-\delta}\right) \le \frac{\mathbb{E}(|\xi|^p)}{i^{(1/2-\delta)p}} = O(i^{\delta p - p/2}).$$

Suppose that $\delta \in (0, (p-2)/2p)$, so that $\delta p - p/2 < -1$, and thus the Borel–Cantelli lemma implies that $\xi_i'' = 0$ for all but finitely many i. Thus, for any $\beta, \varepsilon_0 \in (0, 1/2)$,

$$\lim_{n\to\infty} \frac{\sum_{i=n-n^\beta}^{n} |\xi_i''|}{n^{1/2-\varepsilon_0}} = 0, \ \text{a.s.}$$

For the second term on the right-hand side of (6.1.5), $\mathbb{E}\,\xi = 0$ implies $|\mathbb{E}\,\xi_i'| = |\mathbb{E}\,\xi_i''|$, so

$$\sum_{i=n-n^\beta}^{n} |\mathbb{E}\,\xi_i'| = \sum_{i=n-n^\beta}^{n} |\mathbb{E}\,\xi_i''| \le (n^\beta + 1)\,\mathbb{E}\left(|\xi| \mathbf{1}\{|\xi| > (n/2)^{1/2-\delta}\}\right),$$

for all n large enough so that $n - n^\beta > n/2$. Here

$$\mathbb{E}\left(|\xi|\mathbf{1}\{|\xi| > (n/2)^{1/2-\delta}\}\right) = \mathbb{E}\left(|\xi|^2|\xi|^{-1}\mathbf{1}\{|\xi| > (n/2)^{1/2-\delta}\}\right) \le Cn^{\delta-1/2},$$

for some constant C depending only on $\mathbb{E}(\xi^2)$. Suppose that $\delta \le 1/4$. Then we get $\sum_{i=n-n^\beta}^{n} |\mathbb{E}\,\xi_i'| = O(n^{\beta-1/4})$, so that, for any $\beta \in (0, 1/2)$ and $\varepsilon_0 \in (0, 1/4)$,

$$\lim_{n\to\infty} \frac{\sum_{i=n-n^\beta}^{n} |\mathbb{E}\,\xi_i'|}{n^{1/2-\varepsilon_0}} = 0, \text{ a.s.}$$

Finally, we consider the first term on the right-hand side of (6.1.5), with the truncated, centralised sum, which we denote as $T'_{n,j} := \sum_{i=n-j}^{n}(\xi_i' - \mathbb{E}\,\xi_i')$. The $\xi_i' - \mathbb{E}\,\xi_i'$ are independent, zero-mean random variables with $|\xi_i' - \mathbb{E}\,\xi_i'| \le 2n^{1/2-\delta}$ for $i \le n$, so we may apply the Azuma–Hoeffding inequality, see Theorem 1.3.17, to obtain, for any $t \ge 0$,

$$\mathbb{P}\left(|T'_{n,j}| \ge t\right) \le 2\exp\left(-\frac{t^2}{8(j+1)n^{1-2\delta}}\right).$$

In particular, taking $t = n^{1/2-\varepsilon_0}$ we obtain

$$\begin{aligned}\mathbb{P}\left(\max_{0\le j\le n^\beta} |T'_{n,j}| \ge n^{1/2-\varepsilon_0}\right) &\le (n^\beta + 1)\max_{0\le j\le n^\beta} \mathbb{P}\left(|T'_{n,j}| \ge n^{1/2-\varepsilon_0}\right) \\ &\le 2(n^\beta+1)\exp\left(-\frac{n^{1-2\varepsilon_0}}{16n^{1+\beta-2\delta}}\right),\end{aligned} \tag{6.1.6}$$

for all n sufficiently large. Now choose and fix $\delta = \delta(p) := \min\{1/4, (p-2)/4p\}$, so $\delta > 0$ satisfies the bounds earlier in this proof, and then choose $\beta < \beta_0 := \delta$ such that

$$\frac{n^{1-2\varepsilon_0}}{n^{1+\beta-2\delta}} = n^{2\delta-2\varepsilon_0-\beta} \ge n^{\delta-2\varepsilon_0}.$$

So choosing $\varepsilon_0 = \delta/4$ we have that the probability bound in (6.1.6) is summable. Thus by the Borel–Cantelli lemma, we have that $\max_{0\le j\le n^\beta} |T'_{n,j}| \le n^{1/2-\varepsilon_0}$ for all but finitely many n, a.s. It follows that, for any $\varepsilon \in (0, \varepsilon_0)$,

$$\lim_{n\to\infty} \frac{|\sum_{i=n-n^\beta}^{n}(\xi_i' - \mathbb{E}\,\xi_i')|}{n^{1/2-\varepsilon}} = 0, \text{ a.s.},$$

which completes the proof. □

The main remaining step in the proof of Theorem 6.1.1 is the proof of Lemma 6.1.3.

Proof of Lemma 6.1.3. Using the fact that $\|S_n\|^2 = \|\mu\|^2 n^2 + Y_n^2$, we have that, for

$j \leq n$,

$$\begin{aligned}\|S_j - S_i\|^2 &= \|\mu\|^2(j-i)^2 + (Y_j - Y_i)^2 \\ &= \|S_n\|^2 + \|\mu\|^2 i^2 + \|\mu\|^2 j^2 - 2\|\mu\|^2 ij - \|\mu\|^2 n^2 + Y_i^2 + Y_j^2 - 2Y_iY_j - Y_n^2 \\ &\leq \|S_n\|^2 + \|\mu\|^2 i^2 - (Y_n - Y_j)(Y_n + Y_j) + 2Y_i(Y_n - Y_j) - 2Y_iY_n + Y_i^2.\end{aligned}$$

Here we have that, for any $\varepsilon > 0$, $\max_{0\leq i \leq n^\beta} |Y_iY_n| \leq n^{\frac{1+\beta}{2}+\varepsilon}$ and $\max_{0 \leq i \leq n^\beta} Y_i^2 \leq n^{\beta+\varepsilon}$ almost surely for all but finitely many n. For the terms involving Y_j, Lemma 6.1.6 shows that we may choose $\beta \in (0, 1/2)$ such that, for any sufficiently small $\varepsilon > 0$,

$$\max_{n-n^\beta \leq j \leq n} |Y_n - Y_j| \leq n^{\frac{1}{2}-\varepsilon} \text{ a.s. and } \max_{n-n^\beta \leq j \leq n} |Y_n - Y_j||Y_n + Y_j| \leq n^{1-\varepsilon} \text{ a.s.}$$

for all but finitely many n. With this choice of β and sufficiently small ε, we combine these bounds to obtain

$$\max_{\substack{0 \leq i \leq n^\beta \\ n-n^\beta \leq j \leq n}} \|S_j - S_i\|^2 \leq \|S_n\|^2 + \|\mu\|^2 n^{2\beta} + n^{1-\varepsilon} + n^{\frac{1+\beta}{2}+\varepsilon} + n^{\beta+\varepsilon} \text{ a.s.}$$

for all but finitely many n. Since $\beta \in (0, 1/2)$, we may apply Lemma 6.1.5 and choose $\varepsilon > 0$ sufficiently small to see that $D_n^2 \leq \|S_n\|^2 + n^{1-\varepsilon}$, for all but finitely many n. Hence

$$D_n \leq \|S_n\| \left(1 + \|S_n\|^{-2} n^{1-\varepsilon}\right)^{1/2} \leq \|S_n\| \left(1 + \|\mu\|^{-2} n^{-1-\varepsilon}\right)^{1/2} \text{ a.s.}$$

since $\|S_n\| \geq n\|\mu\|$. Using the fact that $(1+x)^{1/2} \leq 1 + (x/2)$ for $x \geq 0$, we get

$$D_n \leq \|S_n\| \left(1 + \tfrac{1}{2}\|\mu\|^{-2} n^{-1-\varepsilon}\right) \leq \|S_n\| + \|\mu\|^{-1} n^{-\varepsilon} \text{ a.s.}$$

for all but finitely many n, since, by the strong law of large numbers, $\|S_n\| \leq 2\|\mu\|n$ a.s. all but finitely often. Combined with the bound $D_n \geq \|S_n\|$, this completes the proof. □

Proof of Theorem 6.1.1. Combining Lemmas 6.1.3 and 2.2.6 with Slutsky's theorem [Gut05, p. 249] and the fact that, in this case, $X_n = \|\mu\|n$, we obtain (6.1.1).

From Lemma 6.1.4 we have that, if $\mathbb{E}(\|Z\|^p) < \infty$ for $p > 4$, both $D_n - \|\mu\|n$ and

$(D_n - \|\mu\|n)^2$ are uniformly integrable. Thus from (6.1.1) we obtain

$$\lim_{n\to\infty} \mathbb{E}(D_n - \|\mu\|n) = \mathbb{E}\left[\frac{\sigma^2_{\mu_\perp}\zeta^2}{2\|\mu\|}\right] = \frac{\sigma^2_{\mu_\perp}}{2\|\mu\|}, \text{ and}$$
$$\lim_{n\to\infty} \mathbb{E}[(D_n - \|\mu\|n)^2] = \mathbb{E}\left[\frac{\sigma^4_{\mu_\perp}\zeta^4}{4\|\mu\|^2}\right] = \frac{3\sigma^4_{\mu_\perp}}{4\|\mu\|^2}.$$

Using the fact that

$$\mathbb{V}\text{ar}\, D_n = \mathbb{V}\text{ar}(D_n - \|\mu\|n) = \mathbb{E}[(D_n - \|\mu\|n)^2] - \mathbb{E}[D_n - \|\mu\|n]^2,$$

we obtain (6.1.2) on letting $n \to \infty$. □

6.2 Faces and Perimeter length

We continue to consider increments with $\mathbb{E}[(Z\cdot\hat{\mu}-\mu)^2] = 0$ and as before, without loss of generality, we assume $\mu = \|\mu\|\mathbf{e}_0$. We make use of some extra notation here, letting $Z^\perp = Z - (Z\cdot\hat{\mu})\hat{\mu}$ and $Z_i^\perp = Z_i - (Z_i\cdot\hat{\mu})\hat{\mu}$ so $S_n^\perp = \sum_{i=1}^n Z_i^\perp = S_n - (S_n\cdot\hat{\mu})\hat{\mu}$. Note that if $\mathbb{E}\,\|Z\| < \infty$ then $\mathbb{E}\,Z^\perp = 0$.

6.2.1 Faces of the convex hull

Qiao and Steele [QS05] summarise nicely what is known about the faces of such walks in the introduction of their paper.

If we let the number of records in an i.i.d. sequence $X_1, X_2, \ldots$ of random variables with a continuous distribution be denoted R_n, that is

$$R_n = \max\{k : X^\perp_{n_1} < X^\perp_{n_2} < \ldots < X^\perp_{n_k}, 1 \le n_1 < n_2 < \ldots < n_k \le n\},$$

then Rényi [Rén62] showed that this number has the same distribution as a sum of n independent Bernoulli random variables with success probability $1/k$.

Sparre Andersen had previously established the same result for the number of faces of the concave majorant, although he stated the result in terms of the number of vertices in the concave majorant, where the increments of the random walk have a common density, see Section 1.1.3 or [SA54, p. 217]. Goldie [Gol89] connects these results, stating that the number of faces of the concave majorant (in fact, Goldie discusses the convex minorant) of the time-space process, which we denote F_n^+, had the same distribution as the number of records, that is $F_n^+ \stackrel{d}{=} R_n$ and gives a clear explanation of how the Bernoulli sum representation can be established.

Then a Borel-Cantelli argument, using the Bernoulli sum representation of R_n and the monotonicity $R_{n+1} \geq R_n$ gives $\lim_{n\to\infty} R_n / \log n = 1$ almost surely. However, F_n^+ is not monotone so the same argument does not hold. In fact, the lack of monotonicity shows that the process $\{F_n : 1 \leq n < \infty\}$ must be different to the process $\{R_n : 1 \leq n < \infty\}$, whilst the latter is equivalent to the Bernoulli sum process. This is discussed in more detail at [Ste02, §8]. However, Qiao and Steele do note that the Borel-Cantelli lemma is enough to establish, if Z has a density then, a.s.,

$$1 \leq \limsup_{n\to\infty} \frac{F_n^+}{\log n}. \tag{6.2.1}$$

In particular

$$\limsup_{n\to\infty} F_n^+ = \infty, \text{ a.s.}$$

The main result of Qiao and Steele's paper asserts that there exists some distribution of Z such that $\liminf_{n\to\infty} F_n^+ = 1$, and further, almost surely we have $F_n^+ = m$ infinitely often for any $m \in \mathbb{N}$.

We now wish to show that the distribution of Z in their theorem is required to have the property $\mathbb{E}\,\|Z\| = \infty$, and thus it may in fact still be true that under the assumption of finite mean $\lim_{n\to\infty} F_n^+ / \log n = 1$ almost surely.

Lemma 6.2.1. *Let* $\limsup_{n\to\infty} S_n^\perp = +\infty$ *a.s.* $\lim_{n\to\infty} S_n^\perp / n = 0$ *a.s. and* $\sigma_\mu^2 = 0$ *then* $\mathbb{P}(F_n^+ = 1 \text{ i.o.}) = 0$.

Note, $\mathbb{E}\, Z^\perp = 0$ and $\mathbb{P}(Z^\perp = 0) < 1$ are sufficient conditions for the conditions of the lemma since then the strong law of large numbers, Theorem 1.3.14, implies the second condition and the first condition is implied by, for example, [Kal02, Prop 9.14].

Proof. Let n_1 be the first ascending ladder time, that is

$$n_1 := \min\{n > 0 : S_n^\perp > 0\}.$$

Then $\mathbb{P}(n_1 < \infty) = 1$ by the condition $\limsup_{n\to\infty} S_n^\perp = +\infty$ a.s. Denote the angle relative to $\mathbf{e}_0$ of the leftmost edge of the concave majorant at time n by α_n. Then at time n_1 we have $\alpha_{n_1} = \tan^{-1}(S_{n_1}^\perp/n_1)$, and note $0 < \alpha_n < \pi/2$ for all $n > n_1$. But $\lim_{n\to\infty} S_n^\perp/n = 0$ a.s., so there exists some almost surely finite time N such that $S_n^\perp < \tan(\alpha_{n_1})n$ for all $n > N$, note that by this definition of N we have $N \geq n_1$. Since the angle of the leftmost edge of the concave majorant is non-decreasing, at time N, $\alpha_N \geq \alpha_{n_1}$. However, to change the first edge, we require $S_n^\perp > \tan(\alpha_N)n > \tan(\alpha_{n_1})n$ for some $n > N$, of course this contradicts the definition of N so the first edge is fixed and $\mathbb{P}(F_n^+ \geq 2$ for all $n \geq N) = 1$ and the proof is complete. □

In fact, just as Qiao and Steele's result is actually stated in terms of returns of the process $\{F_n : 1 \leq n < \infty\}$ to any natural number m and not just returns to 1, we can extend our result to the following.

Theorem 6.2.2. *When $\sigma_\mu^2 = 0$,*

(i) If $\limsup_{n\to\infty} S_n^\perp = \infty$ a.s. but $\lim_{n\to\infty} S_n^\perp/n = 0$ a.s. then $\lim_{n\to\infty} F_n^+ = \infty$ a.s.

(ii) If $\limsup_{n\to\infty} S_n^\perp/n = \infty$ a.s. then $\liminf_{n\to\infty} F_n^+ = 1$ a.s.

Remark 6.2.3.

- If $\mathbb{E}\, Z^\perp = 0$ then the SLLN puts us in the case (i). It is not difficult to think of many classical examples that fall into this category, for example the time-space process of the simple symmetric random walk or the walk with $Z^\perp \sim \mathcal{N}(0,1)$, the standard Normal distribution.

- Otherwise, a result of Kesten [Kes70, Corollary 3] states that if $\mathbb{E}\,|Z^\perp| = \infty$, then as $n \to \infty$, $n^{-1}S_n^\perp$ either: (a) tends to $+\infty$ a.s.; (b) tends to $-\infty$ a.s.; or (c) oscillates:
$$-\infty = \liminf_{n\to\infty} n^{-1}S_n^\perp < \limsup_{n\to\infty} n^{-1}S_n^\perp = +\infty, a.s.$$
Erickson [Eri73] gives criteria for classifying such behaviour.

  Clearly (a) implies (ii), whilst (b) implies that $\liminf_{n\to\infty} F_n^- = 1$ a.s. by changing sign and (ii). Then (c) implies both $\liminf_{n\to\infty} F_n^+ = 1$ a.s. and $\liminf_{n\to\infty} F_n^- = 1$ a.s., although trivially, since $\mathbb{P}(Z^\perp = 0) < 1$, there exists some time n_0 such that, for all $n > n_0$ we will have a truly 2-dimensional convex hull (not just a line), and so the minimum number of faces of the hull at any time is 3. Thus, for all $n > n_0$, $F_n^+ = F_n^- = 1$ is not possible.

In the case (a) ((b)), a similar proof to that of part (i) of the theorem can be employed to show that $\lim_{n\to\infty} F_n^- = \infty$ ($\lim_n \to F_n^+ = \infty$). Instead of creating a set of times when a new face is created with positive angle, see below for details, all that is required is a new face in the convex minorant (concave majorant), and then the last time in which the walk goes below (above) the line created by extending this new face. By the assumptions in the theorem and the case we are considering, all of these times will be finite almost surely ensuring an increasing lower bound on the number of faces after each of these almost surely finite times.

For an example of a walk in category (c), consider $Z^\perp$ to have a Cauchy distribution. Then $\mathbb{E}\,|Z^\perp| = \infty$ and

$$\limsup_{n\to\infty} \frac{S_n^\perp}{n} = +\infty \text{ and } \liminf_{n\to\infty} \frac{S_n^\perp}{n} = -\infty,$$

so $F_n^+ = 1$ i.o. and $F_n^- = 1$ i.o. Another example would be the symmetric α-stable distribution with $\alpha \in (0,1)$, see Section 7.2 for further discussion and references.

- Case (ii) can be compared with Qiao and Steele's result. In fact, assuming that Z has a density, then in this case $F_n^+ = m$ i.o. a.s. for any natural number m. To see this, note that a corollary of (6.2.1) is that $\limsup_{n\to\infty} F_n^+ = \infty$, and if $F_n^+ > F_{n-1}^+$ then we have $F_n^+ - F_{n-1}^+ = 1$ so if F_n^+ increases it cannot 'jump over' any number, so the fact that $F_n^+ = 1$ i.o. a.s. means that $F_n^+ = m$ i.o. a.s. as claimed.

Proof of Theorem 6.2.2. We will consider an increasing sequence of times, alternating between the time when a new face is created with angle greater than 0, and the time when the trajectory last exits the cone, centred at the origin, with this angle. Specifically, we denote these times as $\{n_1, N_1, n_2, N_2, \ldots\}$ where $N_0 := 0$ and for $i \in \mathbb{N}$,

$$n_i := \min\{k > N_{i-1} : S_k^\perp - \max_{0\le j\le N_{i-1}} S_j^\perp > 0\},$$

and

$$N_i := \max\left\{k \ge n_i : \frac{S_k^\perp}{k} \ge \frac{S_{n_i}^\perp - \max_{0\le j\le N_{i-1}} S_j^\perp}{n_i - \arg\max_{0\le j\le N_{i-1}} S_j^\perp}\right\},$$

where we use the convention that $\arg\max_{0\le j\le N_{i-1}} S_j^\perp$ is the minimal value at which the maximum is attained. Note, n_1 is the same as n_1 in the previous proof. Also note that if N_{i-1} is almost surely finite then n_i is also almost surely finite by the fact $\limsup_{n\to\infty} S_n^\perp = \infty$ a.s., and then, if n_i and N_{i-1} are finite, so is N_i by the fact $\lim_{n\to\infty} S_n^\perp/n = 0$ a.s.

The important observation is that, after each of the times N_i, the previous faces with angle greater than the angle of the face created at time n_i cannot be altered, because, by definition the walk remains below the line parallel to the angle of the face and positioned below, or at the same height, as the face. By only considering n_{i+1} after N_i, we find a new face with positive angle after this time, to start the process again, but the face containing the n_ith increment cannot be changed so we must have at least one distinct face between each of the N_i.

To formalise this, note that at any time $n > N_i$, for all $1 \le j \le i$ there is at least one face whose endpoints have horizontal coordinates in the interval $[N_{j-1}, N_j]$, and whose relative angle is α_j, satisfying

$$0 < \tan^{-1}(S_{n_1}^\perp/n_1) \le \alpha_1 < \pi/2$$

and for any $j \ge 2$

$$0 < \alpha_j < \alpha_{j-1}.$$

Thus we have at least i distinct faces which are no longer able to be changed of the concave majorant, $F_n^+ \ge i$ for all $n \ge N_i$, and since N_i was almost surely finite, $\liminf_{n\to\infty} F_n^+ \ge i$ a.s., and since i was arbitrary we have proven (i).

Conversely, for (ii) consider α_j to be the angle of the first edge of the convex minorant for the trajectory at some fixed time j. Then, due to the fact this is the time-space process, we have $-\pi/2 < \alpha_j < \pi/2$ for all $j \ge 1$. However, the vector from $\mathbf{0}$ to $S_n^\perp$ has angle $\tan^{-1}(S_n^\perp/n)$, and since $\limsup_{n\to\infty} S_n^\perp/n = \infty$ a.s., we have $\limsup_{n\to\infty} \tan^{-1}(S_n^\perp/n) = \pi/2$. Thus, there is some time N_j with $\mathbb{P}(N_j < \infty) = 1$ such that

$$N_j := \min\{k \ge j : \tan^{-1}(S_k/k) > \alpha_j\},$$

at which time $F_n^+ = 1$. Since j was arbitrarily chosen, we have proven the result. □

Further to showing that the $\liminf_{n\to\infty} F_n^+ = \infty$ in the case of finite first moment, we also will use the Bernoulli sum representation of the number of faces along with Bennett's inequality to show that the growth rate is no faster than logarithmic.

Theorem 6.2.4. *Suppose that Z has a continuous probability distribution and $\sigma_\mu^2 = 0$. Then*

$$1 \le \limsup_{n\to\infty} \frac{F_n^+}{\log n} \le e \ a.s.$$

Proof. By Sparre Andersen and Rényi's work [SA54; Rén62], we have

$$\mathbb{P}(F_n^+ \le k) = \mathbb{P}\left(\sum_{i=1}^n Y_i \le k\right),$$

where $Y_i \sim Ber(1/i)$ that is, independent Bernoulli random variables with success probability $1/i$.

The relevant inequality for such sums of random variables is Bennett's inequality [Ben62]. For $X_1, X_2, \ldots, X_n$, independent random variables with mean zero and finite variance, if $X_i \le 1$ a.s., and $\sigma_n^2 = \sum_{i=1}^n \mathbb{V}\mathrm{ar}(X_i)$, then for any $t \ge 0$,

$$\mathbb{P}\left(\sum_{i=1}^n X_i \ge t\right) \le \exp\left(t - \sigma_n^2 \log\left(1 + \frac{t}{\sigma_n^2}\right) - t\log\left(1 + \frac{t}{\sigma_n^2}\right)\right).$$

In order to use this we will centralise the previously described Bernoulli random variables, $X_i := Y_i - 1/i$. Then we have $X_i \le Y_i \le 1$ a.s. so the assumptions of the inequality are met. This use of Bennett's inequality gives

$$\mathbb{P}\left(F_n^+ - \sum_{i=1}^n \frac{1}{i} \ge t\right) \le \exp\left(t - \sigma_n^2 \log\left(1 + \frac{t}{\sigma_n^2}\right) - t\log\left(1 + \frac{t}{\sigma_n^2}\right)\right),$$

where $\sigma_n^2 = \sum_{i=1}^n i^{-1}(1 - i^{-1})$. Using the fact $\sum_{i=1}^n i^{-2} \le \pi^2/6$ and that $\log(n) \le \sum_{i=1}^n i^{-1} \le 1 + \log(n)$ for n large enough, we get that $\log(n) - \pi^2/6 \le \sigma_n^2 \le 1 + \log n$ for n large enough. Put $t = (e - 1 + \varepsilon)\log(n)$ for some $\varepsilon > 0$, then

$$\begin{aligned}
\log\left(1 + \frac{t}{\sigma^2}\right) &\ge \log\left(1 + \frac{(e-1+\varepsilon)\log n}{1 + \log n}\right) \\
&= \log\left(e + \varepsilon - \frac{e - 1 + \varepsilon}{1 + \log n}\right) \\
&= 1 + \log\left(1 + \frac{\varepsilon}{e} + O((\log n)^{-1})\right) \\
&> 1 + \frac{\varepsilon}{2e},
\end{aligned}$$

for all n large enough. So putting this and $t = (e-1+\varepsilon)\log(n)$ into Bennet's inequality,

$$\begin{aligned}\mathbb{P}\left(F_n - \sum_{i=1}^{n}\frac{1}{i} \geq t\right) &\leq \exp\left((e-1+\varepsilon)\log(n) - \sigma_n^2\left(1+\frac{\varepsilon}{2e}\right)\right.\\ &\qquad \left. -(e-1+\varepsilon)\log(n)\left(1+\frac{\varepsilon}{2e}\right)\right)\\ &= \exp\left(-\frac{\varepsilon(e-1+\varepsilon)}{2e}\log(n) - \sigma_n^2\left(1+\frac{\varepsilon}{2e}\right)\right),\end{aligned}$$

and since $\sigma_n^2 > \log(n) - \pi^2/6$,

$$\begin{aligned}\mathbb{P}\left(F_n - \sum_{i=1}^{n}\frac{1}{i} \geq t\right) &\leq \exp\left(-\frac{\varepsilon(e-1+\varepsilon)}{2e}\log(n) - \left(\log(n) - \frac{\pi^2}{6}\right)\left(1+\frac{\varepsilon}{2e}\right)\right)\\ &\leq \exp\left(-\left(1+\frac{\varepsilon(e+\varepsilon)}{2e}\right)\log n + O(1)\right)\\ &\leq n^{-1-\gamma}\end{aligned}$$

for some $\gamma > 0$.

Applying the Borel-Cantelli lemma, we get that

$$F_n > (e-1+\varepsilon)\log(n) + \sum_{i=1}^{n}\frac{1}{i}$$

finitely often, and so

$$\limsup_{n\to\infty}\frac{F_n^+}{\log n} \leq \lim_{n\to\infty}\left((e-1+\varepsilon) + (\log n)^{-1}\sum_{i=1}^{n}\frac{1}{i}\right) = e+\varepsilon.$$

Since ε was arbitrary, we get the upper bound in the lemma.

The lower bound was discussed at (6.2.1), and was proven by Qiao and Steele [QS05].

□

6.2.2 Variance of the perimeter length of the convex hull

Just as with the diameter, when $\mathbb{E}(\|Z\|^2) < \infty$, $\mu \neq \mathbf{0}$, and $\sigma_\mu^2 = 0$, Theorem 2.2.1 (see also Theorem 1 in [WX15a]) only shows that $\mathbb{V}\mathrm{ar}\, L_n = o(n)$. It was conjectured in [WX15a] that $\mathbb{V}\mathrm{ar}\, L_n = O(\log n)$ in this case. We make the following stronger conjecture.

Conjecture 6.2.5. *Suppose that* $\mathbb{E}(\|Z\|^2) < \infty$, $\mu \neq \mathbf{0}$, $\sigma_\mu^2 = 0$, *and* $\sigma_{\mu_\perp}^2 > 0$. *Then*

$$\lim_{n\to\infty} \frac{\mathbb{V}\mathrm{ar}\, L_n}{\log n} \text{ exists in } (0, \infty).$$

This will remain as a conjecture but we will prove the following weaker but related statement.

Theorem 6.2.6. *Suppose that* Z *has a continuous probability distribution, that* $\mathbb{P}(\|Z\| < C) = 1$ *for some* $C < \infty$, $\mu \neq \mathbf{0}$, $\sigma_\mu^2 = 0$ *and* $\sigma_{\mu_\perp}^2 > 0$. *Then for any* $\varepsilon > 0$,

$$\lim_{n\to\infty} \frac{\mathbb{V}\mathrm{ar}\, L_n}{n^\varepsilon} = 0.$$

In order to get closer to the conjecture above, of course we wish to sharpen our upper bound to be of order $\log n$, but considering a lower bound would also be interesting. We propose the following conjecture which is not only necessary for Conjecture 6.2.5 to be true but would be in contrast to (6.1.2), which states that the variance of the diameter in the time-space degenerate case converges to a positive finite limit without any scaling.

Conjecture 6.2.7. *Suppose that* $\mathbb{E}(\|Z\|^2) < \infty$, $\mu \neq \mathbf{0}$, $\sigma_\mu^2 = 0$, *and* $\sigma_{\mu_\perp}^2 > 0$. *Then*

$$\lim_{n\to\infty} \mathbb{V}\mathrm{ar}\, L_n = \infty.$$

In order to prove Theorem 6.2.6, we start by just considering the concave majorant. To simplify the following notation and subsequent descriptions we, without loss of generality, consider the case $\mu = (1, 0)$ for the rest of this section.

Let the faces of the concave majorant be denoted $e_1^+, \ldots, e_{F_n^+}^+$ where F_n^+ is the number of faces. We say an increment Z_i *belongs to* the jth face of the concave majorant if $e_j^+ = S_{h_r} - S_{h_l} = (h_r - h_l, S_{h_r}^\perp - S_{h_r}^\perp)$ with $h_l < i \leq h_r$ and conversely we call e_j^+ the *parent face* of Z_i. Then we use $y_i := S_{h_r}^\perp - S_{h_l}^\perp$ to denote the vertical component of the parent face of the ith increment, and likewise we denote the horizontal component by $x_i := h_r - h_l$. We use l_i to denote the length of this parent face and α_i to denote its angle.

We denote the point on the concave majorant with horizontal coordinate j as b_j.

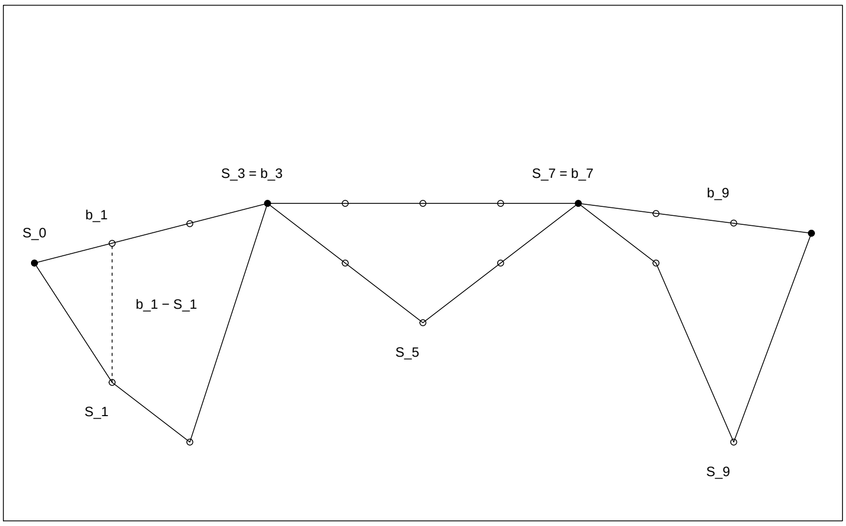


Figure 6.1: Picture to demonstrate the definition of b_i. The upper line shows the concave majorant with b_i points. The lower line is the random walk with points where the walk intersects the convex hull shown by filled in points.

Formally, this is the point interpolating between S_{h_l} and S_{h_r} as follows,

$$b_j = S_{h_l} + \frac{i - h_l}{h_r - h_l}(S_{h_r} - S_{h_l}).$$

Similar to the other notations, we will denote the point on the hull with horizontal coordinate j after resampling the ith increment as $b_j^{(i)}$.

When applying the Efron–Stein inequality, we will consider the change in perimeter length upon resampling Z_i. We also consider this as replacing Z_i with Z_i' and hence denote the perimeter length of the convex hull L_n and $L_n^{(i)}$ before and after the replacement, with M_n and $M_n^{(i)}$ denoting the length of the concave majorant before and after the replacement.

Before embarking on the proof of Theorem 6.2.6, we will give the heuristic behind the proof including a probability bound of an event which will be a particular use to us. The idea is to use Steele's version of the Efron–Stein jackknife inequality, see for example [BLM13, §3.1], which states

$$\mathbb{V}\text{ar}\, M_n \leq \frac{1}{2}\sum_{i=1}^{n} \mathbb{E}\left((M_n - M_n^{(i)})^2\right). \tag{6.2.2}$$

Given this sum, we will split indices into the (random) subsets of i where the parent faces are short or long. The short faces, by definition, will not correspond to more than $O(n^\varepsilon)$ of the i, and the long faces will be flat and thus the difference, $M_n - M_n^{(i)}$ will be controllable, see Lemma 6.2.10 below. In particular, for the long faces, we will need to consider the event

$$\begin{aligned} E_n = &\left\{\cup_{k=0}^{n} \cup_{m=\lfloor n^\varepsilon \rfloor}^{n} \{|S_{k+m}^\perp - S_k^\perp| \geq m^{1/2+\delta}\}\right\} \\ &\cup \left\{\cup_{k=0}^{n} \cup_{m=\lfloor n^\varepsilon \rfloor}^{n} \{|S_{k+m}^{(i)\perp} - S_k^{(i)\perp}| \geq m^{1/2+\delta}\}\right\}, \end{aligned} \tag{6.2.3}$$

with $\varepsilon > 0$, $\delta > 0$. This event describes the situation where a path of the walk with at least n^ε increments, let us say x increments, creates an angle of more than $x^{1/2+\delta}$ between the start and end point of the path. We now consider a bound on the probability of this event.

Lemma 6.2.8. *Suppose that $|Z \cdot \hat{\mu}_\perp| \leq c$ for some $c < \infty$ and $\sigma_\mu^2 = 0$. Then with E_n as defined at (6.2.3), for any $\varepsilon > 0$, there exists $\varepsilon' > 0$, $c_1 > 0$ such that $\mathbb{P}(E_n) \leq \exp(-c_1 n^{\varepsilon'})$.*

Proof. We apply the Azuma–Hoeffding inequality, see Theorem 1.3.17, which states that for any $k \in \{0, \ldots, n\}$ and $m \geq n^\varepsilon$,

$$\mathbb{P}\left(|\tilde{S}_{k+m} - \tilde{S}_k| \geq m^{1/2+\delta}\right) \leq 2\exp\left(\frac{-m^{1+2\delta}}{2mc^2}\right) \leq 2\exp(-c_0 m^{2\delta}). \tag{6.2.4}$$

where $\tilde{S}_j$ is used to represent either S_j or $S_j^{(i)}$, but the two $\tilde{S}_j$ must both be S_j or both be $S_j^{(i)}$, i.e. the bound holds both before and after the resampling. Then an application of the union bound with (6.2.4) gives

$$\begin{aligned} \mathbb{P}(E_n) &\leq 2(n+1)^2 \max_{0 \leq k \leq n} \max_{m \geq n^\varepsilon} \mathbb{P}\left(|\tilde{S}_{k+m} - \tilde{S}_k| \geq m^{1/2+\delta}\right) \\ &\leq 4(n+1)^2 \exp(-c_0 n^{2\delta\varepsilon}) \\ &\leq \exp(-c_1 n^{\varepsilon'}). \end{aligned}$$

□

Remark 6.2.9. Lemma 6.2.8 with the application of the Azuma-Hoeffding inequality is the only place where we really require bounded jumps. Use elsewhere is only for convenience.

The final component before we can prove Theorem 6.2.6 is to describe the control we have under this event when considering long faces.

Lemma 6.2.10. *Suppose* $\sigma_\mu^2 = 0$*. On the events* E_n^c*,* $\{x_i \geq n^\varepsilon\}$ *and* $\{|i - h_l| \wedge |h_r - i| > x_i^\varepsilon\}$*, we have*

$$|M_n - M_n^{(i)}| \leq (i - h_l)^{-1/2+2\delta} \tag{6.2.5}$$

for all $i \in \{1, \dots, n\}$ *and all* n *sufficiently large.*

Proof. We begin by considering a few different cases. First, we will consider the cases where the change upon the resampling in the vertical component of the ith increment is positive and where it is negative separately. Trivially, we do not need to consider the case where $Z_i \cdot (0,1)^\top = Z_i' \cdot (0,1)^\top$ because then the convex hull does not change, so $M_n = M_n^{(i)}$ and the bound in (6.2.5) holds. Within each of these cases we will also separate the situations where $M_n > M_n^{(i)}$ and where $M_n < M_n^{(i)}$ after the resampling.

Let us begin with the case $Z_i \cdot (0,1)^\top < Z_i' \cdot (0,1)^\top$ and $M_n < M_n^{(i)}$. Then we show in Figure 6.2, that, by convexity, $M_n^{(i)}$ is less than the length of the green path and M_n is greater than the length of the blue path, and this bound on the change in length is something we can calculate.

Note, using the notation from the figure, $\Delta = |(Z_i - Z_i') \cdot \hat{\mu}_\perp|$ and $x = i - h_l'$, and under the event E_n^c we know $y < (i - h_l')^{1/2+\delta}$. Then, since the lengths of the faces before vertex h_l' and after b_i remain unchanged, we only need to consider the change of the length of the path between $S_{h_l'}$ and S_i or $S_i^{(i)}$. Note that $\Delta > 0$, so we have

$$\begin{aligned}
M_n^{(i)} - M_n &\leq \sqrt{x^2 + (y+\Delta)^2} - \sqrt{x^2 + y^2} \\
&\leq \sqrt{x^2 + (|y|+\Delta)^2} - \sqrt{x^2 + y^2} \\
&= \frac{\left(\sqrt{x^2 + (|y|+\Delta)^2} - \sqrt{x^2+y^2}\right)\left(\sqrt{x^2 + (|y|+\Delta)^2} + \sqrt{x^2+y^2}\right)}{\left(\sqrt{x^2 + (|y|+\Delta)^2} + \sqrt{x^2+y^2}\right)} \\
&\leq \frac{2|y|\Delta + \Delta^2}{2\sqrt{x^2+y^2}} \\
&= \frac{|y|\Delta}{\sqrt{x^2+y^2}} + O\left(\frac{1}{\sqrt{x^2+y^2}}\right).
\end{aligned}$$

Then, applying the bounds on y and x gives that

$$M_n^{(i)} - M_n \leq \frac{|y|\Delta}{\sqrt{x^2+y^2}} + O\left(\frac{1}{\sqrt{x^2+y^2}}\right)$$

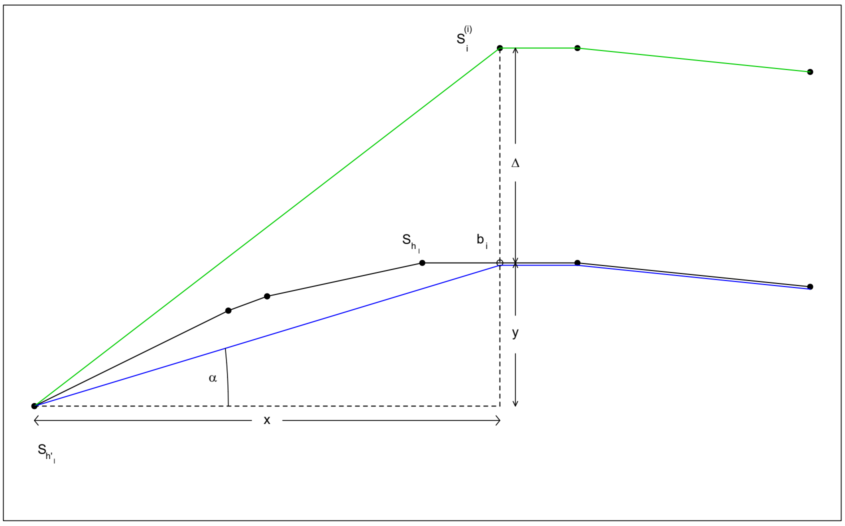


Figure 6.2: Picture to demonstrate the change in M_n possible upon resampling Z_i. The black line shows the concave majorant with b_i indicated. The lower blue line is the shortest path the concave majorant could be if it goes through h_l' and b_i, with everything fixed up to translations before and after these points respectively. The green upper line shows the longest path possible after the resampling, when b_i moves upwards.

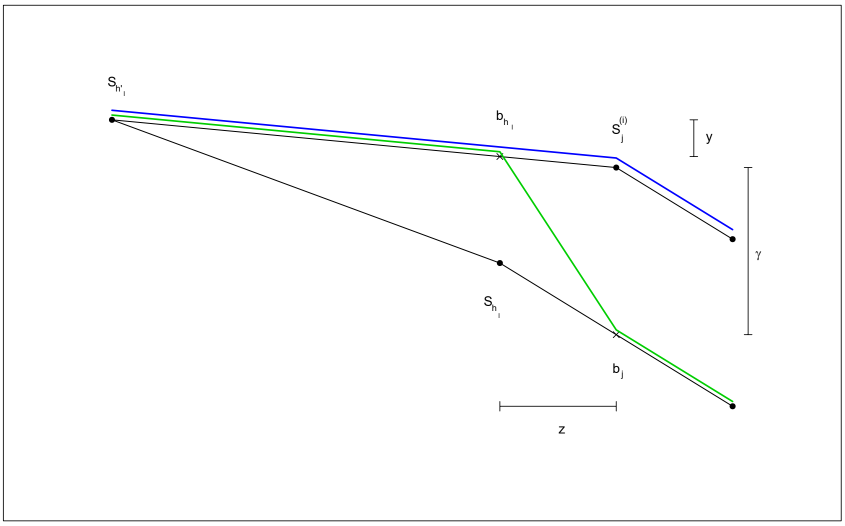


Figure 6.3: Picture to demonstrate the change in M_n possible upon resampling Z_i. The black lines shows the concave majorant before and after resampling. The upper blue line is in fact the concave majorant after resampling but shows the lower bound of $M_n^{(i)}$, with everything fixed up to translations before and after these points respectively. The green upper line shows the longest path possible after the resampling, when b_i moves upwards.

$$\begin{aligned} &\leq (i-h_l)^{-1/2+\delta} + O\left((i-h_l')^{-1}\right) \\ &\leq (i-h_l)^{-1/2+2\delta}, \end{aligned} \tag{6.2.6}$$

for all n large enough. Now let's consider the case $M_n > M_n^{(i)}$. Then any new convex hull corner points must still have the index in $[i, h_r]$. Let j be the smallest index in this range that refers to a convex corner hull point. Note, j must exist, because the point S_{h_r} must still be in the boundary of the convex hull. Then, given j we know the shortest possible path for $L_n^{(i)}$ contributing to $M_n^{(i)}$ is the path from $S_{h_l'}$ to $S_j^{(i)}$ and then to $S_{h_r}^{(i)}$, which is shown in blue in Figure 6.3. If $j = h_r$ then there is one fewer vertex in the hull than Figure 6.3 suggests, but the argument does not change. We then wish to find an upper bound for the length of the original hull. By convexity of the section of the concave majorant between h_l' and j, such an upper bound is the length of the path shown in green. The green path has the same length as the blue path except for the vector starting at $b_{h_l}^{(i)}$. The angle of the section from $b_{h_l}^{(i)}$ to $S_j^{(i)}$ is the same as that of $S_{h_l'}$ to $S_j^{(i)}$ and this angle is controlled by the event E_n^c so it has angle of size smaller than $(j-h_l)^{-1/2+\delta}$. The difference in vertical position between $S_j^{(i)}$ and b_j, γ in the figure, is less than that between $S_j^{(i)}$ and S_j, because $S_j^{(i)}$ is inside the hull, which is precisely Δ from before. Note that both $\gamma > 0$ and $\Delta > 0$. These facts mean we can compute a similar bound as we used in calculating the change in length in (6.2.6),

$$\begin{aligned} M_n - M_n^{(i)} &\leq \sqrt{z^2+(y+\gamma)^2} - \sqrt{z^2+y^2} \\ &\leq \sqrt{z^2+(|y|+\Delta)^2} - \sqrt{z^2+y^2} \\ &\leq \frac{|y|\Delta}{\sqrt{z^2+y^2}} + O\left(\frac{1}{\sqrt{z^2+y^2}}\right) \\ &\leq (i-h_l)^{-1/2+\delta} + O\left((i-h_l')^{-1}\right) \\ &\leq (i-h_l)^{-1/2+2\delta}, \end{aligned}$$

for all n large enough. Similar arguments hold for the two cases where $Z_i' \cdot (0,1) < Z_i \cdot (0,1)$. □

Proof of Theorem 6.2.6. Take $\varepsilon > 0$, then separating (6.2.2) into two parts, we get

$$\mathbb{V}\mathrm{ar}\, M_n \leq \frac{1}{2}\sum_{i=1}^n \mathbb{E}\left((M_n - M_n^{(i)})^2 \mathbf{1}\{x_i \leq n^\varepsilon\}\right) + \frac{1}{2}\sum_{i=1}^n \mathbb{E}\left((M_n - M_n^{(i)})^2 \mathbf{1}\{x_i > n^\varepsilon\}\right). \tag{6.2.7}$$

Further to this, we will need the Azuma–Hoeffding formula to apply within a long face, so we will split the sum of the elements in the long faces into those elements "in the middle" of the faces and those near the ends as follows.

$$\begin{aligned} \mathbb{V}\mathrm{ar}\, M_n \leq & \frac{1}{2}\sum_{i=1}^n \mathbb{E}\left((M_n - M_n^{(i)})^2 \mathbf{1}\{x_i \leq n^\varepsilon\}\right) \\ & + \frac{1}{2}\sum_{i=1}^n \mathbb{E}\left((M_n - M_n^{(i)})^2 \mathbf{1}\{x_i > n^\varepsilon\}\mathbf{1}\{|i-h_l| \wedge |h_r - i| \leq x_i^\varepsilon\}\right) \\ & + \frac{1}{2}\sum_{i=1}^n \mathbb{E}\left((M_n - M_n^{(i)})^2 \mathbf{1}\{x_i > n^\varepsilon\}\mathbf{1}\{|i-h_l| \wedge |h_r - i| > x_i^\varepsilon\}\right). \end{aligned} \tag{6.2.8}$$

For the next part, we consider just the first two terms. Note Lemma 3.1 of [WX15a] with Cauchy's formula gives that $(M_n - M_n^{(i)})^2 \leq 2\pi\|Z_i\| + 2\pi\|Z_i'\|$. If we then use the assumption $\mathbb{P}(\|Z\| < C) = 1$ then we choose C_0 such that $4\pi\|Z\| \leq C_0$ a.s., then

$$\frac{1}{2}\sum_{i=1}^n \mathbb{E}\left((M_n - M_n^{(i)})^2 \mathbf{1}\{x_i \leq n^\varepsilon\}\right) \leq \frac{1}{2}\mathbb{E}\sum_{i=1}^n \left(C_0 \mathbf{1}\{x_i \leq n^\varepsilon\}\right) \tag{6.2.9}$$

$$
\begin{aligned}
&= C_1\,\mathbb{E}\sum_{x=1}^{n^\varepsilon}\sum_{j=1}^{F_n^+}\mathbf{1}\{\|e_j^+\cdot\hat\mu\| = x\}x \\
&\le C_1 n^\varepsilon\,\mathbb{E}\sum_{x=1}^{n^\varepsilon}\sum_{j=1}^{F_n^+}\mathbf{1}\{\|e_j^+\cdot\hat\mu\| = x\} \\
&\le C_1 n^\varepsilon\,\mathbb{E}\,F_n^+ \\
&\le C_2 n^\varepsilon \log n \qquad (6.2.10)
\end{aligned}
$$

where the final inequality uses Sparre Andersen's [SA54] result $\mathbb{E}\,F_n^+ = 1 + \sum_{i=1}^{n-1}(i+1)^{-1}$ which we have bounded by $(C_2/C_1)\log n$ for some large enough C_2. Since ε was arbitrary we have, for any $\varepsilon > 0$

$$
\lim_{n\to\infty} n^{-\varepsilon}\left(\frac{1}{2}\sum_{i=1}^{n}\mathbb{E}\left((M_n - M_n^{(i)})^2\mathbf{1}\{x_i \le n^\varepsilon\}\right)\right) = 0. \qquad (6.2.11)
$$

We can use the same method of bounding the squared difference in perimeter length by a constant and reindexing for the second term in (6.2.8). This time we get

$$
\begin{aligned}
&\frac{1}{2}\sum_{i=1}^{n}\mathbb{E}\left((M_n - M_n^{(i)})^2\mathbf{1}\{x_i > n^\varepsilon\}\mathbf{1}\{|i-h_l|\wedge|h_r - i| \le x_i^\varepsilon\}\right) \\
&\qquad\le C_1\,\mathbb{E}\sum_{j=1}^{F_n^+}\sum_{x=n^\varepsilon}^{n}\left(2\sum_{i=1}^{x^\varepsilon}\mathbf{1}\{\|e_j^+\cdot\hat\mu\| = x\}\right) \\
&\qquad\le C_3 n^\varepsilon\,\mathbb{E}\,F_n^+ \\
&\qquad\le C_4 n^\varepsilon \log n, \qquad (6.2.12)
\end{aligned}
$$

and again, since ε was arbitrary we have, for any $\varepsilon > 0$

$$
\lim_{n\to\infty} n^{-\varepsilon}\left(\frac{1}{2}\sum_{i=1}^{n}\mathbb{E}\left((M_n - M_n^{(i)})^2\mathbf{1}\{x_i > n^\varepsilon\}\mathbf{1}\{|i-h_l|\wedge|h_r - i| \le x_i^\varepsilon\}\right)\right) = 0. \qquad (6.2.13)
$$

Now we consider the third term of (6.2.8). We begin by adding a further condition to this term, the event E_n described at (6.2.3),

$$
\begin{aligned}
&\frac{1}{2}\sum_{i=1}^{n}\mathbb{E}\left((M_n - M_n^{(i)})^2\mathbf{1}\{x_i > n^\varepsilon\}\mathbf{1}\{|i-h_l|\wedge|h_r - i| > x_i^\varepsilon\}\right) \\
&\quad\le \exp(-c_0 n^{\varepsilon'})\sum_{i=1}^{n}\mathbb{E}\left((M_n - M_n^{(i)})^2\mathbf{1}\{E_n\}\mathbf{1}\{x_i > n^\varepsilon\}\mathbf{1}\{|i-h_l|\wedge|h_r - i| > x_i^\varepsilon\}\right) \\
&\quad+ \sum_{i=1}^{n}\mathbb{E}\left((M_n - M_n^{(i)})^2\mathbf{1}\{E_n^c\}\mathbf{1}\{x_i > n^\varepsilon\}\mathbf{1}\{|i-h_l|\wedge|h_r - i| > x_i^\varepsilon\}\right).
\end{aligned}
$$

Again using the assumption $(M_n - M_n^{(i)})^2 \le C_0$ as at (6.2.9) we can simplify further to

$$\frac{1}{2}\sum_{i=1}^{n} \mathbb{E}\left((M_n - M_n^{(i)})^2 \mathbf{1}\{x_i > n^\varepsilon\}\mathbf{1}\{|i-h_l| \wedge |h_r - i| > x_i^\varepsilon\}\right)$$
$$\le C_0 n \exp(-c_0 n^{\varepsilon'})$$
$$+ \frac{1}{2}\sum_{i=1}^{n} \mathbb{E}\left((M_n - M_n^{(i)})^2 \mathbf{1}\{E_n^c\}\mathbf{1}\{x_i > n^\varepsilon\}\mathbf{1}\{|i-h_l| \wedge |h_r - i| > x_i^\varepsilon\}\right). \quad (6.2.14)$$

Then, focusing on the second term in (6.2.14), we can apply Lemma 6.2.10. Choosing $\delta \in (0, \varepsilon/4)$ we get

$$\frac{1}{2}\sum_{i=1}^{n} \mathbb{E}\left((M_n - M_n^{(i)})^2 \mathbf{1}\{E_n^c\}\mathbf{1}\{x_i > n^\varepsilon\}\mathbf{1}\{|i-h_l| \wedge |h_r - i| > x_i^\varepsilon\}\right)$$
$$\le \frac{1}{2}\mathbb{E}\sum_{x=n^\varepsilon}^{n}\sum_{j=1}^{F_n^+}\sum_{x^\varepsilon \le k \le x - x^\varepsilon} k^{-1+4\delta}\mathbf{1}\{\|e_j^+ \cdot \hat{\mu}\| = x\}$$
$$\le C\,\mathbb{E}\sum_{x=n^\varepsilon}^{n}\sum_{j=1}^{F_n^+} x^{4\delta}\mathbf{1}\{\|e_j^+ \cdot \hat{\mu}\| = x\} \quad (6.2.15)$$
$$\le C n^{4\delta}\,\mathbb{E}\,F_n^+ \le n^\varepsilon. \quad (6.2.16)$$

Using the bounds (6.2.10), (6.2.12), (6.2.14) and (6.2.16) we see that

$$\mathbb{V}\mathrm{ar}\, M_n \le C' n^\varepsilon,$$

for some large enough C'. By symmetry we also have $\mathbb{V}\mathrm{ar}\, M_n^- \le C' n^\varepsilon$ where M_n^- is the length of the convex minorant. Then by the Cauchy–Schwarz inequality

$$\mathbb{V}\mathrm{ar}\, L_n \le 2\,\mathbb{V}\mathrm{ar}\, M_n + 2\,\mathbb{V}\mathrm{ar}\, M_n^- \le 4C' n^\varepsilon,$$

and since ε was arbitrary, the result follows. □

6.3 Application of results to our examples

Our first simulation for this chapter demonstrates the convergence of the difference between the diameter and $\|\mu\| n$ to a squared Normal distribution as described in Theorem 6.1.1. Figure 6.4 shows the empirical distribution of the the difference between the diameter and $\|\mu\| n$ for the random walk with drift and no variance in the first coordinate, unit mean, in the darker bars, whilst the lighter bars demonstrate what the limiting distribution looks like – these bars are in fact the empirical distribution of $10,000$ simulated values of a Normal distribution transformed appropriately.

As with in previous chapters, we can compare these distributions with a simplified version of the Kolmogorov-Smirnov distance. Here, we are considering the range $[0,4]$ and so this time our measurement of difference between the distributions is

$$\rho^k_{K-S}(F_{test}, F) = \sup_{0\leq i\leq k} |\mathbb{P}(F_{test} \leq 4i/k) - F(4i/k)|.$$

Again we use $k = 80$ to match the binning in the bar charts, and will take $F = \sigma^2_{\mu_\perp}\zeta^2/2\|\mu\|$, the limiting distribution of Theorem 6.1.1. If we take D_{emp} as the empirical cumulative distribution of $D_n - \|\mu\|n$, then we get $\rho^{80}_{K-S}(D_{emp}, F) = 0.0085$ showing strong evidence of the convergence stated in the theorem.

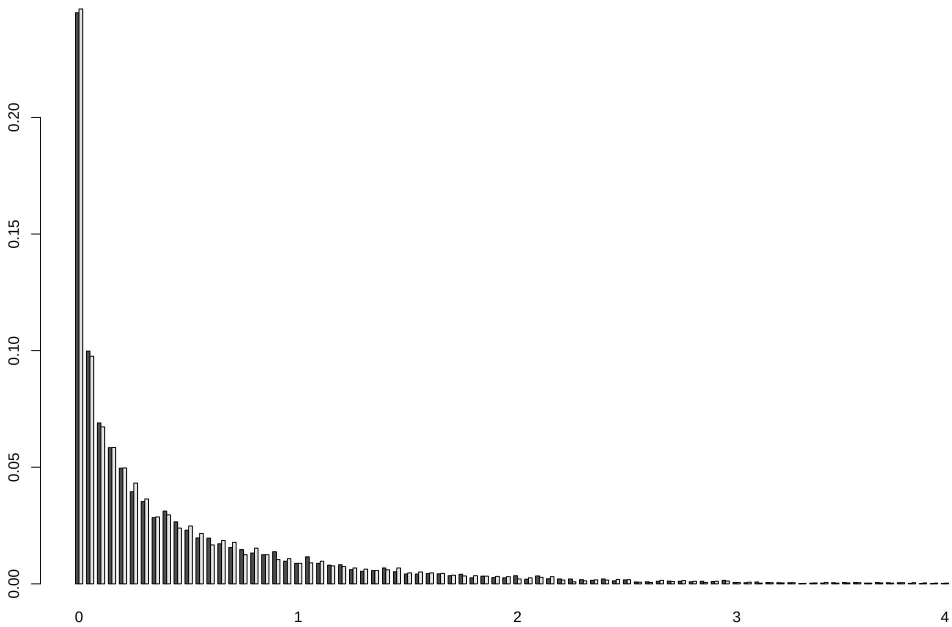


Figure 6.4: The empirical distribution of $D_n - \|\mu\|n$ for our random walk with drift and no variance in the first coordinate, unit mean plotted in the darker bars with the limiting distribution from Theorem 6.1.1 shown by the lighter bars for comparison.

We then demonstrate the claim in Remark 6.1.2 that the generalisation of considering the difference between the diameter and $S_n \cdot \hat{\mu}$ does not always follow the squared Normal distribution. First, in Figure 6.5, we have the empirical distribution of $D_n - S_n \cdot \hat{\mu}$ for our random walk with drift and all coordinates Normally distributed, unit mean.

Here, the simulation distribution is less concentrated near 0, and our Kolmogorov-Smirnov statistic with D_{emp} as the empirical cumulative distribution of $D_n - S_n$ is $\rho^{80}_{K-S}(D_{emp}, F) = 0.1734$ which suggests the limiting distribution is not the same as in Theorem 6.1.1. Before speculating on why this is, we consider a further walk outside of our usual set of examples.

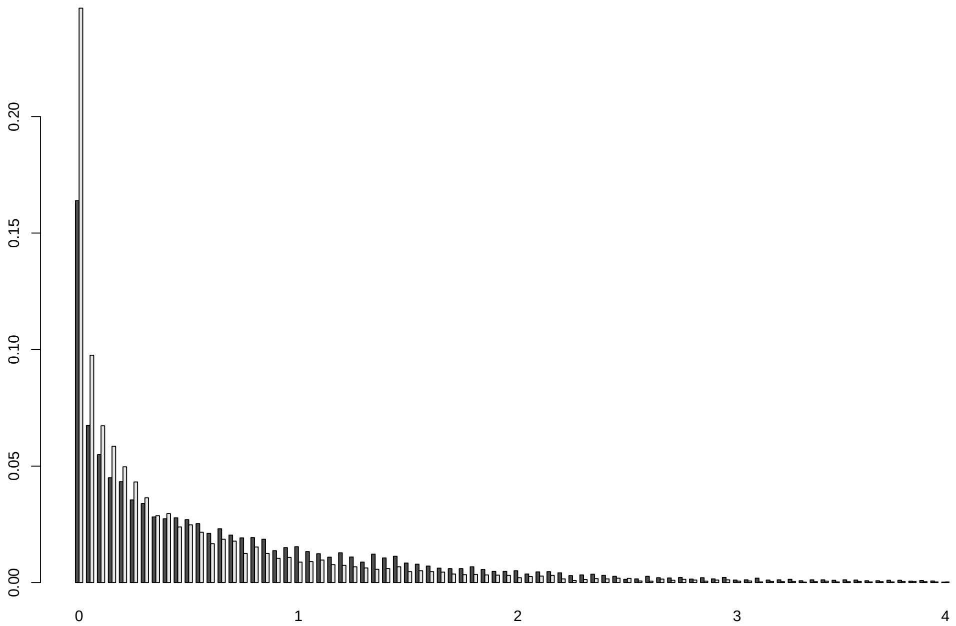


Figure 6.5: The empirical distribution of $D_n - S_n \cdot \hat{\mu}$ for our random walk with Normal drift, unit mean plotted in the darker bars with the limiting distribution from Theorem 6.1.1 shown by the lighter bars for comparison.

Here, in Figure 6.6, we will consider the walk where the increments are fixed unit jumps in the horizontal direction added to a jump on the unit disc, thus $Z = (1, 0) + \mathbf{e}_\theta$ where we recall $\mathbf{e}_\theta$ is the unit vector in direction θ which is chosen uniformly, $\theta \sim U[-\pi, \pi]$. Here we see the empirical distribution is much closer to that of the limiting distribution and this is supported by the now familiar Kolmogorov-Smirnov statistic $\rho^{80}_{K-S}(D_{emp}, F) = 0.0049$ where we have used D_{emp} to represent the empirical cumulative distribution of $D_n - S_n \cdot \hat{\mu}$ for this random walk.

Recalling that our walks, without loss of generality, have mean $(1, 0)$, so we can talk about the walk going to the right as the direction of the mean. One suggestion as to why the distributions look as they do, is that the Normal distribution increments have positive probability of jumping to the left of the origin which are likely to add to the diameter but won't add to $S_n \cdot \hat{\mu}$ which, by the law of large numbers will be approximately $\|\mu\|n$. Our second example, has variance in the direction of the mean but still, with probability 1, has increments which jump to the right. This can be summarized by the following conjecture.

Conjecture 6.3.1. *Suppose that* $\mathbb{E}(\|Z\|^p) < \infty$ *for some* $p > 2$*,* $\mu \neq \mathbf{0}$*, and* $\mathbb{P}(Z \cdot \hat{\mu} > 0) = 1$*. Then,*

$$D_n - S_n \cdot \hat{\mu} \xrightarrow{\mathrm{d}} \frac{\sigma_{\mu_\perp}^2 \zeta^2}{2\|\mu\|}.$$

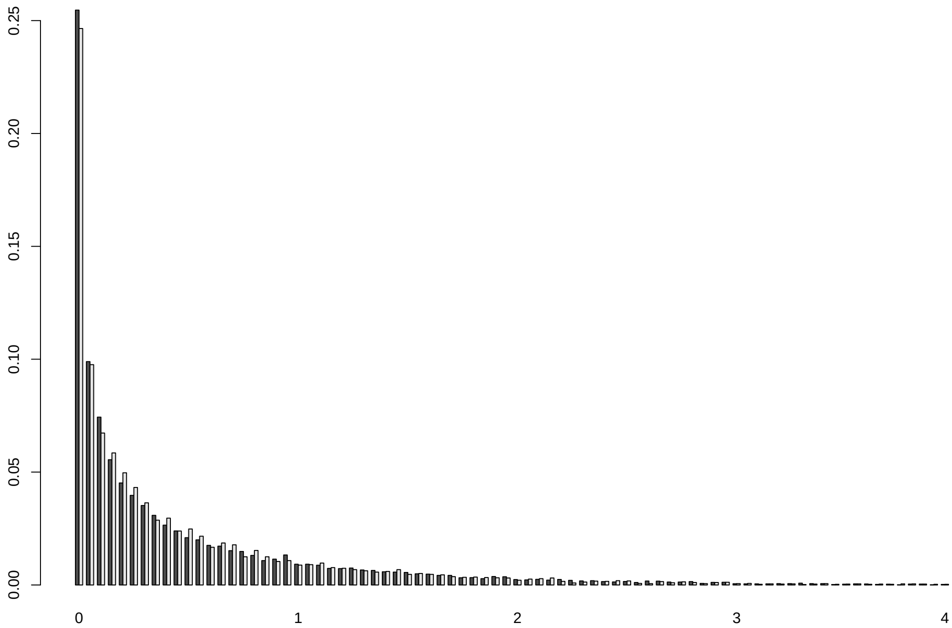


Figure 6.6: The empirical distribution of $D_n - S_n \cdot \hat{\mu}$ for a random walk with increments comprising of $(1,0) + \mathbf{e}_\theta$ plotted in the darker bars with the limiting distribution from Theorem 6.1.1 shown by the lighter bars for comparison.

Chapter 7

Open problems and further extensions

There are many ways in which these results and the work herein can be extended including generalizing the results by removing assumptions where possible, generalizing results to higher dimensions, or using further functionals to establish more information about the convex hulls of random walks. In this chapter we will briefly mention a few of these possibilities, state what is known in the literature where relevant, make some conjectures about further possible results, and provide some pictures which give evidence supporting some of the conjectures. We note, and do not repeat, the conjectures in Chapter 6 which were the result of Remark 6.1.2 and the motivation for the perimeter length section and thus better stated there.

7.1 Extending central limit theorems to trajectory convergence

In Chapter 3, we used the random walk's central limit theorem in establishing convergence of the random walk trajectories to Brownian motion. On the face of it, we may consider trying to perform the same calculation for the central limit theorems for the perimeter length and diameter in the case with drift, Corollary 5.0.2 and Theorem 1.2 of [WX15a].

Particularly suggestive are Theorem 5.1.1 and Theorem 1.3 of [WX15a], see (1.1.4), which show that the centralised and rescaled diameter and perimeter length respectively converge in L^2 to the random walk when centralised and projected in the direction of the mean. This walk is a one-dimensional random walk satisfying the conditions of Theorem 3.1.5, and so the associated trajectory converges to Brownian motion. Unfortunately the L^2 convergence is not enough to give convergence of the diameter or perimeter length processes. However, if the L^2 convergence statements could be strengthened to convergence in probability of maxima, then Slutsky's theorem in the context of weak convergence, Theorem 1.3.13, would indeed tell us that the diameter and perimeter length processes converge to Brownian motion.

7.2 Removing variance assumptions

Throughout, we refer to our condition that the second moment of the increment distribution is finite, $\mathbb{E}[\|Z\|^2] < \infty$. Whilst this is the case for a wide class of commonly used distributions, the theory can be extended to include other increment distributions.

Consider an increment distribution with heavy tails, for example the Pareto distribution where $\mathbb{P}(\|Z\| > t) = c \cdot t^{-\alpha}$ for $t > 1$, $\alpha > 0$ and some constant c. Here, if $\alpha < 2$ then $\mathbb{E}[\|Z\|^2] = \infty$. Therefore, our usual scaling limit of the Normal distribution or trajectory limit of Brownian motion will not pass over to this case.

However, there is much literature concerning such a distribution and it is known that a different scaling limit exists, namely the class of stable distributions. Amongst others, see [GK54; ST94; Nol15; Whi02; Pre72].

In particular, it should be noted that, whilst we do not delve into these broader classes of random walks, we have made use of the Skorokhod space and metric. This is important because it is the space in which discontinuities can be reasonably accounted for, and therefore a lot of the theory in Chapter 3 will be relevant if this wider view is to be considered. This is highlighted by Figure 7.1 which shows a trajectory with $\mathbb{E}\, Z = 0$ but $\mathbb{E}\, Z^2 = \infty$, with 10^5 steps with the colour changing every 10^4 steps to help determine the path of the walk. It is notable that there are several larger increments, shown by long straight lines, which could be thought of as the discontinuities in the process (true discontinuities are not possible to see in a finite simulation as picture here).

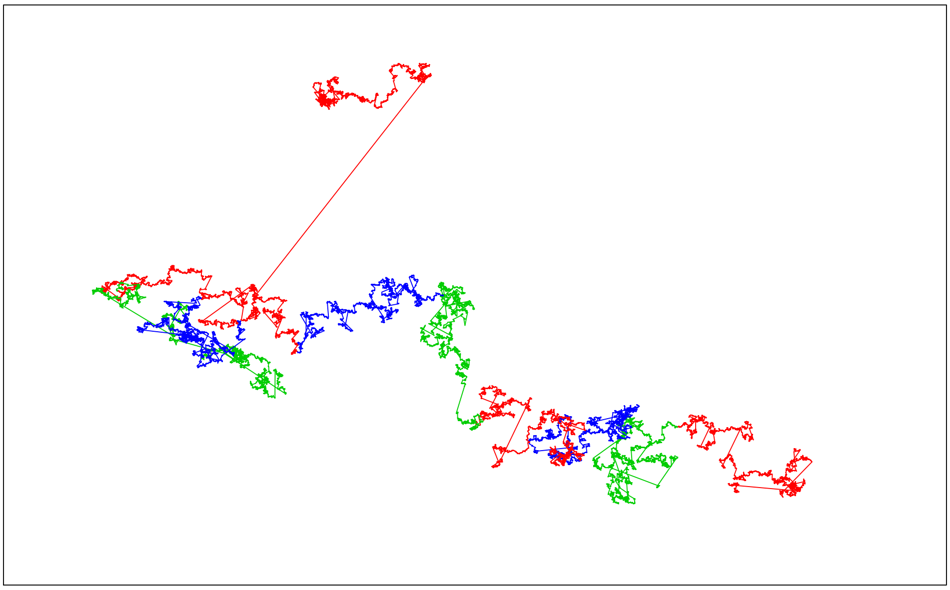


Figure 7.1: A random walk with zero drift with increments following a 2-dimensional version of the Pareto distribution.

7.3 Brownian motion functionals

In Chapter 3, we studied the diameter of planar Brownian motion run for unit time and noted that the expected diameter is still unknown. In order to get a better idea about this problem there are several further questions which could be of interest in their own right.

First, let $d_2(t)$ be the diameter of $b_2(t)$. Then we are interested in $\mathbb{E}\, d_2(1) = t^{-1/2}\, \mathbb{E}\, d_2(1)$ by Brownian scaling, but what does $d_2(t)$ look like as a process in itself? For which set of t does $d_2(t) = d_2(1)$? For this second question, considering $t \in [0, 1]$, one may be tempted to suggest that the process follows the arcsine law, however this seems to be incorrect. In one dimension, we have that both the time at which the minimum and maximum of Brownian motion are attained follow the arcsine law, however this is not a statement about the two times as a joint distribution. And in fact, if one is attained near the start of the process, it seems likely that the other will be attained near the end. Can we formalize this idea and find the exact distribution for the maximum of the two times?

Further to this question regarding the diameter, one could also ask more detailed questions about the convex hull of Brownian motion. For example, we know quite a lot of information about the perimeter length, but can we say something about the number of edges which contribute to the perimeter length? A clearer but equivalent set of questions are: what is the distribution of the length of the longest face, second longest face, and so on?

Although there are a lot of open questions in this area, the book of Borodin and Salmonin [BS02] is an extremely useful resource for finding all manners of distributions and theory related to Brownian motion.

7.4 Higher dimension extensions

Much of the work contained in this thesis is generalised to so-called d-dimensions, however there remains some results relating only to 2-dimensions.

The shape theorem, Theorem 4.2.1, in Chapter 4 could be generalised by considering convergence to $(d-1)$-dimensional shapes. In turn this would enable a generalisation of Corollary 4.1.3 by considering the ratio of the $(d-1)$-dimensional surface area and the diameter, which would have limiting objects of the unit vector with ratio 0, and the unit ball with ratio $\pi^{d/2}/2^{d-2}\Gamma(d/2)$ where Γ is the Euler gamma function, see for example [Som58, p. 136].

The martingale difference method does not seem to be specific to planar random walks, and in Chapter 5, when we are considering the diameter functional, the whole method seems to be easily generalised. The usual considerations of how to generalise other functionals applies but for the diameter at least, there is no issue.

Finally, it also appears likely that with care, one could extend the results of Section 6.1 pertaining to the diameter in the case $\sigma_\mu^2 = 0$ to higher dimensions. It seems reasonable to suggest the diameter is still determined by a point near the start of the walk and a point near the end of the walk, and with some work the technical lemmas could also be extended by taking d-dimensional norms instead of the fixed dimensional norms applied so far.

7.5 Other shape properties of random walks

The aim of this work was to improve the understanding of the shape and size of convex hulls of random walk for processes that have run for a long time. This may not always be the most appropriate feature of a process to study depending on the application. We conclude by mentioning two further properties of random walks which could provide further insight into the walks.

For walks in a discrete setting, it might be more interesting to simply consider how many different points are visited. This would give some crude information on the size of the walk, and maybe some information about the shape could be derived too. The functional described here has already been studied and is called the *range of the random walk*. For references see [DE51; Spi76; JO68; JP70a; JP70b; JP71; JP72b; JP72a; JP74; Fla76], and the introduction of [JP72b] provides a nice discussion of the contribution of some of these papers.

In a similar vein, one could consider the area enclosed by a random walk. Considering a walk in $\mathbb{R}^2$, this is defined as the set of points for which there is no line to infinity that does not intersect the trajectory of the walk. Informally, consider this as the sum of the areas of the polygons created when the trajectory intersects itself. It is a simple exercise to see that this set is a subset of the convex hull of the random walk, but it would be interesting to consider how much smaller this set is. Some simple examples, such as our time-space processes, have trivial solutions. In this case, no subset of $\mathbb{R}^2$ is enclosed by these walks because the trajectories are non-self-intersecting and so they never form polygons which could add to the area enclosed. What about walks with zero drift? In this case, is it reasonable to expect the area enclosed by the walk to converge to that of the convex hull under an appropriate scaling? The only reference known to the author studying this process is [Ham56], however, many papers in the natural sciences consider a slightly different process which they also term the area enclosed, see for example [MN98].

Bibliography

[APRUB11] J. Abramson, J. Pitman, N. Ross, and G. Uribe Bravo. "Convex minorants of random walks and Lévy processes". In: *Electron. Commun. Probab.* 16 (2011), pp. 423–434.

[AT78] S. G. Akl and G. T. Toussaint. "A fast convex hull algorithm". In: *Inform. Process. Lett.* 7 (1978), pp. 219–222.

[AV16] A. Akopyan and V. Vysotsky. "Large deviations for the perimeter of convex hulls of planar random walks". In: *Preprint arXiv:1606.07141* (2016).

[Bac00] L. Bachelier. "Théorie de la spéculation". In: *Ann. scientif. de l'École norm.* 17 (1900), pp. 21–89.

[Bar95] R. G. Bartle. *The Elements of Integration and Lebesgue Measure.* John Wiley & Sons, 1995.

[Bau96] H. Bauer. *Probability Theory.* Vol. 23. De Gruyter Studies in Mathematics. Translated from the fourth (1991) German edition by Robert B. Burckel and revised by the author. Walter de Gruyter & Co., Berlin, 1996.

[Bax61] G. Baxter. "A combinatorial lemma for complex numbers". In: *Ann. Math. Statist.* 32 (1961), pp. 901–904.

[BD88] N.H. Bingham and R.A. Doney. "On higher-dimensional analogues of the arc-sine law". In: *J. Appl. Probab.* 25 (1988), pp. 120–131.

[Bed+15] G. Bedi, F. Carrillo, G. A. Cecchi, D. F. Slezak, M. Sigman, N. B. Mota, S. Ribeiro, D. C. Javitt, M. Copelli, and C. M. Corcoran. "Automated analysis of free speech predicts psychosis onset in high-risk youths". In: *npj Schizophrenia* 1 (2015), p. 15030.

[Ben62] G. Bennett. "Probability inequalities for the sum of independent random variables". In: *J. Amer. Statist. Ass.* 57 (1962), pp. 33–45.

[Ber13] J. Bernoulli. *Ars conjectandi*. Impensis Thurnisiorum, fratrum, 1713.

[Bil12] P. Billingsley. *Probability and Measure*. Wiley Series in Probability and Statistics. Anniversary edition, with a foreword by Steve Lalley and a brief biography of Billingsley by Steve Koppes. John Wiley & Sons, Inc., New York, 2012.

[Bil99] P. Billingsley. *Convergence of Probability Measures*. Second edition. Wiley Series in Probability and Statistics: Probability and Statistics. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1999.

[BLM13] S. Boucheron, G. Lugosi, and P. Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press, 2013.

[BNB63] O. Barndorff-Nielsen and G. Baxter. "Combinatorial lemmas in higher dimensions". In: *Trans. Amer. Math. Soc.* 108 (1963), pp. 313–325.

[BR10] R. N. Bhattacharya and R. R. Rao. *Normal Approximation and Asymptotic Expansions*. Vol. 64. Classics in Applied Mathematics. Updated reprint of the 1986 edition, corrected edition of the 1976 original. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2010.

[Bro28] R. Brown. "XXVII. A brief account 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". In: *Philos. Mag.* 4 (1828), pp. 161–173.

[Bro89] H. Brozius. "Convergence in mean of some characteristics of the convex hull". In: *Adv. Appl. Probab.* 21 (1989), pp. 526–542.

[Bru64] H. D. Brunk. “A generalization of Spitzer’s combinatorial lemma”. In: *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 2 (1964), pp. 395–405.

[BS02] A. N. Borodin and P. Salminen. *Handbook of Brownian Motion—Facts and Formulae*. Second edition. Probab. Appl. Birkhäuser Verlag, Basel, 2002.

[BZ11] L. V. Bogachev and S. M. Zarbaliev. “Universality of the limit shape of convex lattice polygonal lines”. In: *Ann. Probab,* 39 (2011), pp. 2271 –2317.

[Bár95] I Bárány. “The limit shape of convex lattice polygons”. In: *Discrete Comput. Geom.* 13 (1995), pp. 279–295.

[Car64] G. Cardano. *Liber De Ludo Aleae*. 1564.

[CBM15a] M. Chupeau, O. Bénichou, and S. N. Majumdar. “Convex hull of a Brownian motion in confinement”. In: *Phys. Rev. E* 91 (2015), p. 050104.

[CBM15b] M. Chupeau, O. Bénichou, and S. N. Majumdar. “Mean perimeter of the convex hull of a random walk in a semi-infinite medium”. In: *Phys. Rev. E* 92 (2015), p. 022145.

[CE52] K. L. Chung and P. Erdös. “On the application of the Borel-Cantelli lemma”. In: *Trans. Amer. Math. Soc.* 72 (1952), pp. 179–186.

[CF49] K. L. Chung and W. Feller. “On fluctuations in coin-tossing”. In: *Proc. Nat. Acad. Sci. U.S.A.* 35 (1949), pp. 605–608.

[Che07] X. Chen. “A New Generalization of Chebyshev Inequality for Random Vectors”. In: *Preprint arXiv:0707.0805* (2007).

[Che67] P.L. Chebyshev. “On mean quantities”. In: *Mat. Sbornik* (1867).

[CHM15] G. Claussen, A. K. Hartmann, and S. N. Majumdar. “Convex hulls of random walks: large-deviation properties”. In: *Phys. Rev. E* 91 (2015), p. 052104.

[CMVW11] F. Comets, M.V. Menshikov, S. Volkov, and A.R. Wade. “Random walk with barycentric self-interaction”. In: *J. Stat. Phys.* 143 (2011), pp. 855–888.

[Dar51] D.A. Darling. "Sums of symmetrical random variables". In: *Proc. Amer. Math. Soc.* 2 (1951), pp. 511–517.

[DCHM16] T. Dewenter, G. Claussen, A. K. Hartmann, and S. N. Majumdar. "Convex hulls of multiple random walks: a large-deviation study". In: *Phys. Rev. E* 94 (2016), p. 052120.

[DE51] A. Dvoretzky and P. Erdös. "Some problems on random walk in space". In: *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950*. University of California Press, Berkeley and Los Angeles, 1951, pp. 353–367.

[Don51] M.D. Donsker. "An invariance principle for certain probability limit theorems". In: *Mem. Amer. Math. Soc.* 6 (1951), p. 12.

[Dur10] R. Durrett. *Probability: Theory and Examples. 4th ed.* Cambridge University Press, 2010.

[EB83] M. El Bachir. "L'enveloppe convex de mouvement brownien". Ph.D. thesis. Université Toulouse III – Paul Sabatier, 1983.

[Edd82] W. F. Eddy. "Convex hull peeling". In: *COMPSTAT 5th Symposium held at Toulouse 1982*. Springer. 1982, pp. 42–47.

[Efr65] B. Efron. "The convex hull of a random set of points". In: *Biometrika* 52 (1965), pp. 331–343.

[Ein05] A. Einstein. "Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen". In: *Annalen der physik* 322 (1905), pp. 549–560.

[EK09] S.N. Ethier and T.G. Kurtz. *Markov Processes: Characterization and Convergence*. John Wiley & Sons, 2009.

[EK47] P. Erdös and M. Kac. "On the number of positive sums of independent random variables". In: *Bull. Amer. Math. Soc.* 53 (1947), pp. 1011–1020.

[Eld14] R. Eldan. "Volumetric properties of the convex hull of an n-dimensional Brownian motion". In: *Electron. J. Probab.* 19 (2014), pp. 1–34.

[Eri73] K.B. Erickson. "The strong law of large numbers when the mean is undefined". In: *Trans. Amer. Math. Soc.* 185 (1973), pp. 371–381.

[ES35] P. Erdös and G. Szekeres. "A combinatorial problem in geometry". In: *Compositio Math.* 2 (1935), pp. 463–470.

[Fel51] W. Feller. "The asymptotic distribution of the range of sums of independent random variables". In: *Ann. Math. Statistics* 22 (1951), pp. 427–432.

[Fel68] W. Feller. *An Introduction to Probability Theory and its Applications. Vol. I.* Third edition. John Wiley & Sons, Inc., New York-London-Sydney, 1968.

[Fel71] W. Feller. *An Introduction to Probability Theory and its Applications. Vol. II.* Second edition. John Wiley & Sons, Inc., New York-London-Sydney, 1971.

[Fis11] H. Fischer. *A History of the Central Limit Theorem.* Sources and Studies in the History of Mathematics and Physical Sciences. Springer, New York, 2011.

[Fla76] Leopold Flatto. "The multiple range of two-dimensional recurrent walk". In: *Ann. Probab.* 4 (1976), pp. 229–248.

[FP93] S.-C. Fang and S. Puthenpura. *Linear Optimization and Extensions: Theory and Algorithms.* Prentice-Hall, Inc., 1993.

[Fre03] D.H. Fremlin. *Measure Theory, Vol. 4: Topological Measure Spaces.* Torres Fremlin, Colchester, 2003.

[GH08] S. Givant and P. Halmos. *Introduction to Boolean Algebras.* Springer Science & Business Media, 2008.

[GK54] B. Gnedenko and A. Kolmogorov. *Limit Distributions for Sums of Random Variables.* Cambridge, MA: Addison-Wesley, 1954.

[GLM17] D.S. Grebenkov, Y. Lanoiselée, and S. N. Majumdar. "Mean perimeter and mean area of the convex hull over planar random walks". In: *J. Stat. Mech. Theory Exp.* 10 (2017), p. 103203.

[Gol89] C. M. Goldie. “Records, permutations and greatest convex minorants”. In: *Math. Proc. Cambridge Philos. Soc.* 106 (1989), pp. 169–177.

[Gro12] P. Groeneboom. “Convex hulls of uniform samples from a convex polygon”. In: *Adv. in Appl. Probab.* 44 (2012), pp. 330–342.

[Gro73] H. Groemer. “On some mean values associated with a randomly selected simplex in a convex set”. In: *Pacific J. Math.* 45 (1973), pp. 525–533.

[Gru07] P.M. Gruber. *Convex and Discrete Geometry*. Springer, Berlin, 2007.

[Gut05] A. Gut. *Probability: A Graduate Course*. Springer Texts in Statistics. Springer, New York, 2005.

[Gut09] A. Gut. *An Intermediate Course in Probability*. Second. Springer, New York, 2009.

[GW88] P.W. Glynn and W. Whitt. “Ordinary CLT and WLLN versions of L=λW”. In: *Mathematics of operations research* 13 (1988), pp. 674–692.

[Hal50] P. Halmos. *Measure Theory*. First edition. Springer-Verlag New York, 1950.

[Hal81] P. Hall. “A comedy of errors: the canonical form for a stable characteristic function”. In: *Bulletin of the London Mathematical Society* 13 (1981), pp. 23–27.

[Hal93] E. Halley. “An Estimate of the Degrees of the Mortality of Mankind”. In: *Phil. Trans.* 196 (1693), pp. 596–610.

[Ham56] J. M. Hammersley. “The area enclosed by Pólya’s walk”. In: *Proc. Cambridge Philos. Soc.* 52 (1956), pp. 78–87.

[Haw06] S. Hawking. *God Created the Integers: The Mathematical Breakthroughs that Changed History*. Penguin Books, 2006.

[Hea97] T. L. Heath. *The Works of Archimedes*. Cambridge University Press, 1897.

[Hol71] P. Holgate. “Random walk models for animal behavior”. In: *Statistical Ecology* 2 (1971), pp. 1–12.

[Hus88] Z. Hussain. “A fast approximation to a convex hull”. In: *Pattern recognition letters* 8 (1988), pp. 289–294.

[HW16] J. Helfer and D. T. Wise. “A note on maxima in random walks”. In: *Electron. J. Combin.* 23 (2016).

[HW41] P. Hartman and A. Wintner. “On the law of the iterated logarithm”. In: *Amer. J. Math.* 63 (1941), pp. 169–176.

[JO68] N. Jain and S. Orey. “On the range of random walk”. In: *Israel J. Math.* 6 (1968), pp. 373–380.

[JP70a] Naresh C. Jain and William E. Pruitt. “The central limit theorem for the range of transient random walk”. In: *Bull. Amer. Math. Soc.* 76 (1970), pp. 758–759.

[JP70b] Naresh C. Jain and William E. Pruitt. “The range of recurrent random walk in the plane”. In: *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 16 (1970), pp. 279–292.

[JP71] Naresh C. Jain and William E. Pruitt. “The range of transient random walk”. In: *J. Analyse Math.* 24 (1971), pp. 369–393.

[JP72a] Naresh C. Jain and William E. Pruitt. “The law of the iterated logarithm for the range of random walk”. In: *Ann. Math. Statist.* 43 (1972), pp. 1692–1697.

[JP72b] Naresh C. Jain and William E. Pruitt. “The range of random walk”. In: *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. III: Probability Theory* (1972), pp. 31–50.

[JP74] Naresh C. Jain and William E. Pruitt. “Further limit theorems for the range of random walk”. In: *J. Analyse Math.* 27 (1974), pp. 94–117.

[JP75] N. C. Jain and W. E. Pruitt. “The other law of the iterated logarithm”. In: *Ann. Probab.* 3 (1975), pp. 1046–1049.

[JR82] N. P. Jewell and J. P. Romano. “Coverage problems and random convex hulls”. In: *J. Appl. Probab.* 19 (1982), pp. 546–561.

[Kac54] M. Kac. “Toeplitz matrices, translation kernels and a related problem in probability theory”. In: *Duke Math. J.* 21 (1954), pp. 501–509.

[Kal02] O. Kallenberg. *Foundations of Modern Probability.* Second edition. Probability and its Applications (New York). Springer-Verlag, New York, 2002.

[Kel75] J. L. Kelley. *General Topology.* Springer-Verlag, New York-Berlin, 1975.

[Kes70] H. Kesten. “The limit points of a normalized random walk”. In: *Ann. Math. Statist.* 41 (1970), pp. 1173–1205.

[KF12] A.N. Kolmogorov and S.V. Fomin. *Introductory Real Analysis.* Courier Corporation, 2012.

[Khi29] A.Y. Khinchin. “Sur la loi des grandes nombres”. In: *Comptes Rendus Acad. Sci.* 188 (1929), pp. 477–479.

[Kho92a] D. Khoshnevisan. “Local asymptotic laws for the Brownian convex hull”. In: *Probab. Theory Related Fields* 93 (1992), pp. 377–392.

[Kho92b] D. Khoshnevisan. “Moment inequalities for functionals of the Brownian convex hull”. In: *Ann. Probab.* 20 (1992), pp. 627–630.

[Kin69] J. F. C. Kingman. “Random secants of a convex body”. In: *J. Appl. Probability* 6 (1969), pp. 660–672.

[KL98] J. Kuelbs and M. Ledoux. “On convex limit sets and Brownian motion”. In: *J. Theoret. Probab.* 11 (1998), pp. 461–492.

[KLM12] J. Kampf, G. Last, and I. Molchanov. “On the convex hull of symmetric stable processes”. In: *Proc. Amer. Math. Soc.* 140 (2012), pp. 2527–2535.

[KM63] M. G. Kendall and P. A. P. Moran. *Geometrical Probability.* Griffin’s Statistical Monographs & Courses, No. 10. Hafner Publishing Co., New York, 1963.

[Kol30] A.N. Kolmogorov. “Sur la loi forte des grands nombres”. In: *Comptes Rendus Acad. Sci.* 191 (1930), pp. 910–912.

[Kol56] A.N. Kolmogorov. “On Skorokhod convergence”. In: *Theor. Probab. Appl.* 1 (1956), pp. 215–222.

[KRR18] S. Kunis, B. Reichenwallner, and M. Reitzner. “Random approximation of convex bodies: monotonicity of the volumes of random tetrahedra”. In: *Discrete Comput. Geom.* 59 (2018), pp. 165–174.

[Kuo06] H. Kuo. *Introduction to Stochastic Integration.* Springer, New York, 2006.

[KVZ16] Z. Kabluchko, V. Vysotsky, and D. Zaporozhets. “A multidimensional analogue of the arcsine law for the number of positive terms in a random walk”. In: *Preprint arXiv:1610.02861* (2016).

[KVZ17a] Z. Kabluchko, V. Vysotsky, and D. Zaporozhets. “Convex hulls of random walks: expected number of faces and face probabilities”. In: *Adv. Math.* 320 (2017), pp. 595–629.

[KVZ17b] Z. Kabluchko, V. Vysotsky, and D. Zaporozhets. “Convex hulls of random walks, hyperplane arrangements, and Weyl chambers”. In: *Geom. Funct. Anal.* 27 (2017), pp. 880–918.

[KZ16] Z. Kabluchko and D. Zaporozhets. “Intrinsic volumes of Sobolev balls with applications to Brownian convex hulls”. In: *Trans. Amer. Math. Soc.* 368 (2016), pp. 8873–8899.

[Leh99] E. L. Lehmann. *Elements of Large-Sample Theory.* Springer-Verlag, New York, 1999.

[Let78] G. Letac. “Advanced Problem 6230”. In: *Am. Math. Mon.* 85 (1978), p. 686.

[Leu82] S. Leurgans. “Asymptotic distributions of slope-of-greatest-convex-minorant estimators”. In: *Ann. Statist.* 10 (1982), pp. 287–296.

[LGE13] M. Luković, T. Geisel, and S. Eule. “Area and perimeter covered by anomalous diffusion processes”. In: *New Journal of Physics* 15 (2013).

[Lip52] M. Lipschutz. “Generalization of a theorem of Chung and Feller”. In: *Proc. Amer. Math. Soc.* 3 (1952), pp. 659–670.

[Lip53] M. Lipschutz. “On strong laws for certain types of events connected with sums of independent random variables”. In: *Ann. of Math.* 57 (1953), pp. 318–330.

[LL10] G.F. Lawler and V. Limic. *Random Walk: A Modern Introduction*. Cambridge University Press, 2010.

[LMW18] C.H. Lo, J. McRedmond, and C. Wallace. “Functional limit theorems for random walks”. In: *Preprint arXiv:1810.06275* (2018).

[Luk70] E. Lukacs. *Characteristic Functions*. Griffin, 1970.

[LW17] C.H. Lo and A.R. Wade. “On the centre of mass of a random walk”. In: *Preprint arXiv:1708.04470* (2017).

[Lév40a] P. Lévy. “Le mouvement brownien plan”. In: *Amer. J. Math.* 62 (1940), pp. 487–550.

[Lév40b] P. Lévy. “Sur certains processus stochastiques homogènes”. In: *Compositio math* 7 (1940), pp. 283–339.

[Lév48] P. Lévy. *Processus Stochastiques et Mouvement Brownien. Suivi d’une note de M. Loève*. Gauthier-Villars, Paris, 1948.

[MAHPS95] R. Meier, F. Ackermann, G. Herrmann, S. Posch, and G. Sagerer. “Segmentation of molecular surfaces based on their convex hull”. In: *Image Processing, 1995. Proceedings., International Conference on*. Vol. 3. IEEE. 1995, pp. 552–555.

[Mar51] G. Maruyama. “Note on the arc sine law in the theory of probability”. In: *Nat. Sci. Rep. Ochanomizu Univ.* 2 (1951), pp. 25–27.

[Mas00] B. Massé. “On the LLN for the number of vertices of a random convex hull”. In: *Adv. in Appl. Probab.* 32 (2000), pp. 675–681.

[Mas99] B. Massé. “On the variance of the number of extreme points of a random convex hull”. In: *Statist. Probab. Lett.* 44 (1999), pp. 123–130.

[MCRF10] S. N. Majumdar, A. Comtet, and J. Randon-Furling. “Random convex hulls and extreme value statistics”. In: *J. Stat. Phys.* 138 (2010), pp. 955–1009.

[Med07] P. Medvegyev. *Stochastic Integration Theory*. Oxford University Press on Demand, 2007.

[ML64] R. MacArthur and R. Levins. "Competition, habitat selection, and character displacement in a patchy environment". In: *Proceedings of the National Academy of Sciences* 51 (1964), pp. 1207–1210.

[MMS13] S. N. Majumdar, P. Mounaix, and G. Schehr. "Exact statistics of the gap and time interval between the first two maxima of random walks and Lévy flights". In: *Physical review letters* 111 (2013), p. 070601.

[MMS14] S. N. Majumdar, P. Mounaix, and G. Schehr. "On the gap and time interval between the first two maxima of long random walks". In: *J. Stat. Mech. Theory Exp.* (2014), p. 09013.

[MN98] J. A. Mingo and A. Nica. "On the distribution of the area enclosed by a random walk in Z^2". In: *Journal of Combinatorial Theory.* A 84 (1998), pp. 55–86.

[MPW17] M. Menshikov, S. Popov, and A.R. Wade. *Non-Homogeneous Random Walks.* Vol. 209. Cambridge Tracts in Mathematics. Lyapunov function methods for near-critical stochastic systems. Cambridge University Press, Cambridge, 2017.

[MS86] Vitali D. Milman and Gideon Schechtman. *Asymptotic Theory of Finite-Dimensional Normed Spaces.* Vol. 1200. Lecture Notes in Mathematics. With an appendix by M. Gromov. Springer-Verlag, Berlin, 1986.

[MSM16] P. Mounaix, G. Schehr, and S. N. Majumdar. "On the gap and time interval between the first two maxima of long continuous time random walks". In: *J. Stat. Mech. Theory Exp.* (2016), p. 013303.

[MT85] M. M. Mcqueen and G. T. Toussaint. "On the ultimate convex hull algorithm in practice". In: *Pattern Recognition Letters* 3 (1985), pp. 29–34.

[MW16] I. Molchanov and F. Wespi. "Convex hulls of Lévy processes". In: *Electron. Commun. Probab.* 21 (2016), p. 69.

[MW18] J. McRedmond and A.R. Wade. "The convex hull of a planar random walk: perimeter, diameter, and shape". In: *Electron. J. Probab.* 23 (2018).

[MX17] J. McRedmond and C. Xu. "On the expected diameter of planar Brownian motion". In: *Statist. Probab. Lett.* 130 (2017), pp. 1–4.

[Nol08] J. Nolan. *An overview of multivariate stable distributions*. 2008.

[Nol15] J. Nolan. *Stable Distributions - Models for Heavy Tailed Data*. In progress, Chapter 1 online at academic2.american.edu/∼jpnolan. Boston, MA: Birkhauser, 2015.

[NS08] W. Nicholson and C. Snyder. *Microeconomic Theory: Basic Principles and Extensions, International Student Edition, 10th ed.* South-Western Cengage Learning, 2008.

[PB06] K. Pearson and J. Blakeman. *A Mathematical Theory of Random Migration*. Dulau and co., London, 1906.

[Pea05a] K. Pearson. "The problem of the random walk". In: *Nature* 72 (1905), p. 294.

[Pea05b] K. Pearson. "The problem of the random walk". In: *Nature* 72 (1905), p. 342.

[Pen03] M. Penrose. *Random Geometric Graphs*. Vol. 5. Oxford Studies in Probability. Oxford University Press, Oxford, 2003.

[Pey97] N. Peyerimhoff. "Areas and intersections in convex domains". In: *Amer. Math. Monthly* 104 (1997), pp. 697–704.

[Pfi89] R. E. Pfiefer. "The historical development of J. J. Sylvester's four point problem". In: *Math. Mag.* 62 (1989), pp. 309–317.

[Pit96] V.I. Piterbarg. *Asymptotic Methods in the Theory of Gaussian Processes and Fields*. AMS Trans. of Math., 1996.

[Poi37] S.D. Poisson. *Probabilité des jugements en matière criminelle et en matière civile, précédées des règles générales du calcul des probabilitiés*. Bachelier, 1837.

[Pol84] D. Pollard. *Convergence of Stochastic Processes*. Springer-Verlag, New York, 1984.

[PR69] Y.V. Prohorov and Y.A. Rozanov. *Probability Theory*. 1969.

[Pre72] S. J. Press. “Multivariate stable distributions”. In: *J. Multivariate Anal.* 2 (1972), pp. 444–462.

[PUB12] J. Pitman and G. Uribe Bravo. “The convex minorant of a Lévy process”. In: *Ann. Probab.* 40 (2012), pp. 1636–1674.

[QS05] Z. Qiao and J.M. Steele. “Random walks whose concave majorants often have few faces”. In: *Statist. Probab. Lett.* 75 (2005), pp. 97–102.

[Ray05] Lord Rayleigh. “The problem of the random walk”. In: *Nature* 72 (1905), p. 318.

[Ray80] Lord Rayleigh. “Xii. on the resultant of a large number of vibrations of the same pitch and of arbitrary phase”. In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 10 (1880), pp. 73–78.

[RC03] M. Rubinstein and R.H. Colby. *Polymer Physics.* Oxford University Press, 2003.

[Rei00] M. Reitzner. “Inequalities for convex hulls of random points”. In: *Monatsh. Math.* 131 (2000), pp. 71–78.

[Rei03] M. Reitzner. “Random polytopes and the Efron-Stein jackknife inequality”. In: *Ann. Probab.* 31 (2003), pp. 2136–2166.

[RF13] Julien Randon-Furling. “Convex hull of n planar Brownian paths: an exact formula for the average number of edges”. In: *J. Phys. A* 46 (2013), p. 015004.

[RF14] Julien Randon-Furling. “Universality and time-scale invariance for the shape of planar Lévy processes”. In: *Phys. Rev. E* 89 (2014), p. 052112.

[RFMC09] J. Randon-Furling, S. N. Majumdar, and A. Comtet. “Convex hull of N planar Brownian motions: exact results and an application to ecology”. In: *Phys. Rev. Lett.* 103 (2009), p. 140602.

[RFW17] Julien Randon-Furling and Florian Wespi. “Facets on the convex hull of d-dimensional Brownian and Lévy motion”. In: *Phys. Rev. E* 95 (2017), p. 032129.

[RMR11] A. Reymbaut, S. N. Majumdar, and A. Rosso. "The convex hull for a random acceleration process in two dimensions". In: *J. Phys. A* 44 (2011), p. 415001.

[Roc70] R. T. Rockafellar. *Convex Analysis*. Vol. 28. Princeton University Press, 1970.

[Rog78] L. C. G. Rogers. "The probability that two samples in the plane will have disjoint convex hulls". In: *J. Appl. Probab.* 15 (1978), pp. 790–802.

[RS63] A. Rényi and R. Sulanke. "Über die konvexe Hülle von n zufällig gewählten Punkten". In: *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 2 (1963), pp. 75–84.

[RS64] A. Rényi and R. Sulanke. "Über die konvexe Hülle von n zufällig gewählten Punkten. II". In: *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 3 (1964), pp. 138–147.

[RT87] WanSoo T. Rhee and Michel Talagrand. "Martingale inequalities and NP-complete problems". In: *Math. Oper. Res.* 12 (1987), pp. 177–181.

[Rud76] W. Rudin. *Principles of Mathematical Analysis*. Third edition. McGraw-Hill Book Co., New York-Auckland-Düsseldorf, 1976.

[Rén62] A. Rényi. "On the extreme elements of observations". In: *MTA III Oszt. Közl* 12 (1962), pp. 105–121.

[Rév13] P. Révész. *Random Walk in Random and Non-Random Environments*. Third. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2013.

[SA49] E. Sparre Andersen. "On the number of positive sums of random variables". In: *Skand. Aktuarietidskr.* 32 (1949), pp. 27–36.

[SA50] E. Sparre Andersen. "On the frequency of positive partial sums of a series of random variables". In: *Mat. Tidsskr. B.* (1950), pp. 33–35.

[SA53] E. Sparre Andersen. "On the fluctuations of sums of random variables". In: *Math. Scand.* 1 (1953), pp. 263–285.

[SA54] E. Sparre Andersen. "On the fluctuations of sums of random variables. II". In: *Math. Scand.* 2 (1954), pp. 195–223.

[Sai95] R. Saigal. *Linear Programming.* Vol. 1. International Series in Operations Research & Management Science. A modern integrated analysis. Kluwer Academic Publishers, Boston, MA, 1995.

[SHM17] H. Schawe, A. K. Hartmann, and S. N. Majumdar. "Convex hulls of random walks in higher dimensions: A large-deviation study". In: *Phys. Rev. E* 96 (2017), p. 062101.

[SHM18] H. Schawe, A. K. Hartmann, and S. N. Majumdar. "Large Deviations of Convex Hulls of Self-Avoiding Random Walks". In: *Preprint arXiv:1804.02371* (2018).

[Sin94] Ya. G. Sinai. "Probabilistic approach to the analysis of statistics for convex polygonal lines". In: *Funct. Anal. Appl.* 28 (1994), pp. 108–113.

[SM12] G. Schehr and S. N. Majumdar. "Universal order statistics of random walks". In: *Phys. Rev. Lett.* 108 (2012), p. 040601.

[Som58] D.M.Y. Sommerville. *An Introduction to the Geometry of n Dimensions.* Dover Publications, New York, 1958.

[Spi56] F. Spitzer. "A combinatorial lemma and its application to probability theory". In: *Trans. Amer. Math. Soc.* 82 (1956), pp. 323–339.

[Spi76] Frank Spitzer. *Principles of Random Walk.* Second. Graduate Texts in Mathematics, Vol. 34. Springer-Verlag, New York-Heidelberg, 1976.

[SS93] T.L. Snyder and J.M. Steele. "Convex hulls of random walks". In: *Proc. Amer. Math. Soc.* 117 (1993), pp. 1165–1173.

[ST94] G. Samorodnitsky and M.S. Taqqu. *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance.* CRC press, 1994.

[Sta83] A. J. Stam. "Cycles of random permutations". In: *Ars Combin.* 16 (1983), pp. 43–48.

[Ste02] J. M. Steele. "The Bohnenblust-Spitzer algorithm and its applications". In: *J. Comput. Appl. Math.* 142 (2002), pp. 235–249.

[Ste66] P. Stein. "Classroom Notes: A Note on the Volume of a Simplex". In: *Amer. Math. Monthly* 73 (1966), pp. 299–301.

[Ste90] J. Michael Steele. “Probabilistic and worst case analyses of classical problems of combinatorial optimization in Euclidean space”. In: *Math. Oper. Res.* 15 (1990), pp. 749–770.

[SW61] F. Spitzer and H. Widom. “The circumference of a convex polygon”. In: *Proc. Amer. Math. Soc.* 12 (1961), pp. 506–509.

[Syl64] J. J. Sylvester. “Question 1491”. In: *Educational Times, London* (1864).

[Tak80] L. Takács. “Expected perimeter length”. In: *Am. Math. Mon.* 87 (1980), pp. 142–142.

[TV16] E. Tsukerman and E. Veomett. “A Simple Proof of Cauchy’s Surface Area Formula”. In: *Preprint arXiv:1604.05815* (2016).

[Uda52] M. Udagawa. “On numbers of positive sums of independent random variables”. In: *Kōdai Math. Sem. Rep.* 4 (1952), pp. 45–50.

[Ver94] A. M. Vershik. “The limit form of convex integral polygons and related problems”. In: *Funct. Anal. Appl.* 28 (1994), pp. 13–20.

[VZ18] V. Vysotsky and D. Zaporozhets. “Convex hulls of multidimensional random walks”. In: *Trans. Amer. Math. Soc.* 370 (2018), pp. 7985–8012.

[Wat65] S. Watson. “Question No. 1229”. In: *Mathematical Questions and their Solutions from the Educational Times* 4 (1865), p. 101.

[Whi02] W. Whitt. *Stochastic-Process Limits*. Springer Series in Operations Research. An introduction to stochastic-process limits and their application to queues. Springer-Verlag, New York, 2002.

[Wil91] D. Williams. *Probability with Martingales*. Cambridge University Press, Cambridge, 1991.

[Woo67] W. S. B. Woolhouse. “Question no. 2471”. In: *Mathematical Questions and their Solutions from the Educational Times* 8 (1867), pp. 100–105.

[Wor87] B. J. Worton. “A review of models of home range for animal movement”. In: *Ecological Modelling* 38 (1987), pp. 277–298.

[Wor95] B. J. Worton. “A convex hull-based estimator of home-range size”. In: *Biometrics* (1995), pp. 1206–1215.

[WX15a] A. R. Wade and C. Xu. "Convex hulls of planar random walks with drift". In: *Proc. Amer. Math. Soc.* 143 (2015), pp. 433–445.

[WX15b] A. R. Wade and C. Xu. "Convex hulls of random walks and their scaling limits". In: *Stochastic Process. Appl.* 125 (2015), pp. 4300–4320.

[Ye95] Q.-Z. Ye. "A fast algorithm for convex hull extraction in 2D images". In: *Pattern Recognition Letters* 16 (1995), pp. 531–537.

[YMBH15] S. S.-T. Yau, W.-G. Mao, M. Benson, and R. L. He. "Distinguishing proteins from arbitrary amino acid sequences". In: *Scientific Reports* 5 (2015), p. 7972.

Appendix A

Functional limit theory

For the proofs in this section[1] we will the the extra notation for the canonical projection at time $t \in [0,1]$, $\pi_t : \mathcal{M}^d \to \mathbb{R}^d$, defined as $\pi_t f = f(t)$ for $f \in \mathcal{M}^d$. We start by describing some extra background theory relating to the metric space.

A.1 The space of trajectories - extra theory

A.1.1 The Skorokhod metric

We start by providing the motivation behind the Skorokhod metric, and then provide some technical results related to this metric which are required in some proofs. First, consider the following example.

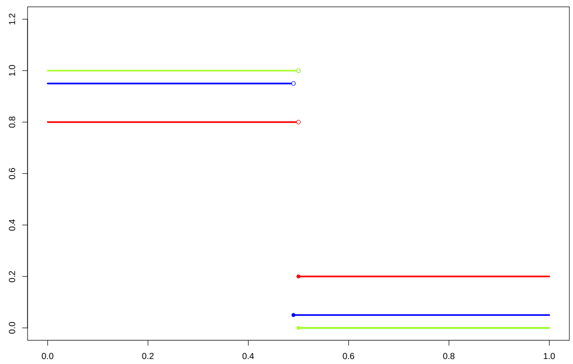
Example A.1.1. Consider the following three functions,

[1] Based on work in [LMW18] which is the generalisation of [Bil99] to higher dimensions. Lemma A.1.3, the Portmanteau theorem and generalisation of Etemadi's inequality were written by the first and third authors, whilst the rest of the work was all a joint collaboration.

$$f(t) = \begin{cases} 1 & \text{for } t \in [0, 1/2); \\ 0 & \text{for } t \in [1/2, 1]; \end{cases}$$

$$g(t) = \begin{cases} 0.8 & \text{for } t \in [0, 1/2); \\ 0.2 & \text{for } t \in [1/2, 1]; \end{cases}$$

$$h(t) = \begin{cases} 0.95 & \text{for } t \in [0, 0.49); \\ 0.05 & \text{for } t \in [0.49, 1]. \end{cases}$$



Taking an overview of the plot, it seems reasonable to suggest that the blue function, $h(t)$, is 'closer' to the light-green function, $f(t)$, than the red function, $g(t)$, is to the light-green function. However, if we consider the supremum metric, we find that $\rho_\infty(f, h) = 0.95$ whilst $\rho_\infty(f, g) = 0.2$. This is due to the slightly earlier jump at $t = 0.49$ for h, which, for a small interval of t, takes the function to the larger Euclidean distance of 0.95 from f. So for processes with jumps, we may wish to consider a different measure of distance.

As formally defined at (1.3.4), the metric we have used is the Skorokhod metric which we can also describe as follows.

For f and g in $\mathcal{M}^d$, define $\rho_S(f, g)$ to be the infimum of those positive ε for which there exists in Λ a λ satisfying

$$\sup_{0 \leq t \leq 1} |\lambda(t) - t| = \sup_{0 \leq t \leq 1} |t - \lambda^{-1}(t)| < \varepsilon \tag{A.1.1}$$

and

$$\sup_{0 \leq t \leq 1} \|f(t) - g(\lambda(t))\| = \sup_{0 \leq t \leq 1} \|f(\lambda^{-1}(t)) - g(t)\| < \varepsilon. \tag{A.1.2}$$

Applying this to the example above we see the difference in the metrics.

Example A.1.2. Consider the functions $f(t)$, $g(t)$ and $h(t)$ from Example A.1.1. The distance $\rho_S(f, g) = 0.2$ because there is no perturbation of the time which would decrease the Euclidean distance between f and g. However, when we consider f and h, we could define

$$\lambda(t) := \begin{cases} \frac{49}{50} t & \text{for } t \in [0, 1/2) \\ \frac{51}{50} t - \frac{1}{50} & \text{for } t \in [1/2, 1]. \end{cases}$$

It turns out that this λ is optimal giving $\rho_S(f,h) = 0.05$ because equation (A.1.1) gives us a lower bound of $\varepsilon = 0.01$ attained when $t = 0.5$ and $\lambda(t) = 0.49$, and equation (A.1.2) gives the lower bound of $\varepsilon = 0.05$, which is attained at any $t \in [0,1]$.

So this is why we have the Skorokhod metric, and as we have described just after (1.3.4), another metric, the Kolmogorov-Billingsley metric, is equivalent to the Skorokhod metric in the sense of Proposition 3.1.1. This metric is useful in some of our proofs but we will require the following technical observation about $\|\lambda\|^\circ$.

Lemma A.1.3. *Let $\lambda \in \Lambda$. Define $c(\lambda) := \max\{\mathrm{e}^{\|\lambda\|^\circ} - 1, 1 - \mathrm{e}^{-\|\lambda\|^\circ}\}$. Then we have*

$$|\lambda(t) - t| \leq tc(\lambda), \ \textit{for all}\ t \in [0,1]; \tag{A.1.3}$$

and

$$|\lambda'(t) - 1| \leq c(\lambda), \ \textit{almost everywhere on}\ t \in (0,1). \tag{A.1.4}$$

Proof. From the definition of $\|\lambda\|^\circ$, we have that for any $t \in [0,1)$ and $h > 0$ sufficiently small,

$$\log\left|\frac{\lambda(t+h) - \lambda(t)}{h}\right| \leq \|\lambda\|^\circ$$

so that

$$\mathrm{e}^{-\|\lambda\|^\circ} \leq \frac{\lambda(t+h) - \lambda(t)}{h} \leq \mathrm{e}^{\|\lambda\|^\circ}. \tag{A.1.5}$$

By Lebesgue's theorem on the differentiability of monotone functions, see [KF12, p. 321], $\lambda'(t)$ exists almost everywhere on $t \in (0,1)$, and when it does exist, we have from (A.1.5) that

$$\mathrm{e}^{-\|\lambda\|^\circ} \leq \lambda'(t) \leq \mathrm{e}^{\|\lambda\|^\circ}.$$

Hence we see that (A.1.4) holds as required. For the first assertion, since $\lambda(0) = 0$, another application of the definition of $\|\lambda\|^\circ$ shows that $\log|\lambda(t)/t| \leq \|\lambda\|^\circ$ for all $t \in (0,1)$, so that $|\lambda(t)| \leq t\mathrm{e}^{\|\lambda\|^\circ}$, hence

$$t\mathrm{e}^{-\|\lambda\|^\circ} \leq \lambda(t) \leq t\mathrm{e}^{\|\lambda\|^\circ}, \ \text{for all}\ t \in [0,1].$$

It follows that

$$-t\left(1 - \mathrm{e}^{-\|\lambda\|^\circ}\right) \leq \lambda(t) - t \leq t\left(\mathrm{e}^{\|\lambda\|^\circ} - 1\right),$$

and so we get (A.1.3) as required. Hence we completed the proof. □

Another reason we may wish to consider the Kolmogorov-Billingsley metric, ρ_S°, is that it has the advantage that it provides a metric with which the space $\mathcal{D}^d$ is a complete metric space.

Theorem A.1.4. *The space $\mathcal{C}^d$ is separable and complete under ρ_∞.*

Theorem A.1.5. *The space $\mathcal{D}^d$ is separable under ρ_S and ρ_S°, and complete under ρ_S°.*

The one-dimensional case of Theorem A.1.4 is discussed at [Bil99, p. 11]. The separability extends to higher dimensions by, for example [Fre03, §4A2Q]. This result also implies that there exists some measure for which the space is complete, and it is a simple exercise to see that every one-dimensional projection of a Cauchy sequence under ρ_∞ in d-dimensions is also a Cauchy sequence and therefore has a limit in the product space.

As mentioned above, the one-dimensional case of Theorem A.1.5 was proven by Kolmogorov in [Kol56], but is also discussed at [Bil99, Theorem 12.2]. The separability for higher dimensions extends as in the continuous case, using [Fre03, §4A2Q] and the completeness of the space under the measure ρ_S° also follows with a similar simple calculation.

A.1.2 Modulus of continuity

As well as the extra metric, we need to consider another way of comparing continuity of trajectories, in particular when they have discontinuities. First of all, for $f \in \mathcal{C}^d$, the associated modulus of continuity is defined by

$$w_f(\delta) := \sup_{|s-t|<\delta} \|f(s) - f(t)\|, \text{ for } 0 < \delta \leq 1.$$

In $\mathcal{D}^d$, the analogous concept is a little more involved (see [Bil99, p. 122]). A set $\{t_i : 0 \leq i \leq v\}$ which has $0 = t_0 < t_1 < \cdots < t_v = 1$ is called δ-sparse if it also satisfies $\min_{1\leq i\leq v}(t_i - t_{i-1}) > \delta$. Then define, for $0 < \delta \leq 1$,

$$w'_f(\delta) := \inf_{\{t_i\}} \max_{1\leq i\leq v} \sup_{t,s\in[t_{i-1},t_i)} \|f(t) - f(s)\|, \tag{A.1.6}$$

where the infimum extends over all δ-sparse sets $\{t_i\}$.

A.2 Functional laws of large numbers proofs

A.2.1 Almost-sure convergence and the strong law

We first provide a proof of the almost-sure mapping theorem, Theorem 3.1.4.

Proof. For any ω such that h is continuous at $X(\omega)$, $X_n(\omega) \to X(\omega)$ implies that $h(X_n(\omega)) \to h(X(\omega))$. Then

$$\begin{aligned}\mathbb{P}(\{\omega \in \Omega : \rho'(h(X_n(\omega)), h(X(\omega))) \to 0 \text{ as } n \to \infty\})& \\ \geq \mathbb{P}(\{\omega \in \Omega : h \text{ is continuous at } X(\omega)\})& \\ = \mathbb{P}(X \in D_h^{\mathrm{c}}) = 1,&\end{aligned}$$

so that $h(X_n) \xrightarrow{\text{a.s.}} h(X)$. □

A.2.2 The functional law of large numbers

Next we provide a proof of the functional law of large numbers. Theorem 3.1.2 is apparently stronger than Theorem 1.3.14 since convergence in the ρ_∞ metric implies convergence of the endpoints $X_n(1) = n^{-1}S_n \xrightarrow{\text{a.s.}} \mu = I_\mu(1)$ and $X'_n(1) = n^{-1}S_n \xrightarrow{\text{a.s.}} \mu = I_\mu(1)$. However, we will see that Theorem 3.1.2 is in fact just a recasting of Theorem 1.3.14, so the two results are equivalent. See, for example, [Whi02, p. 26] for a reference.

Proof of Theorem 3.1.2. Let $\varepsilon > 0$. By Theorem 1.3.14, there exists N_ε with $\mathbb{P}(N_\varepsilon < \infty) = 1$ such that, for all $n \geq N_\varepsilon$, $\|n^{-1}S_n - \mu\| \leq \varepsilon$. Then

$$\begin{aligned}\sup_{N_\varepsilon/n \leq t \leq 1} \|X'_n(t) - \mu t\| &\leq \sup_{N_\varepsilon/n \leq t \leq 1} \left\| X'_n(t) - \frac{\lfloor nt \rfloor}{n}\mu \right\| + \sup_{N_\varepsilon/n \leq t \leq 1} \left\| \frac{\lfloor nt \rfloor}{n}\mu - t\mu \right\| \\ &\leq \sup_{N_\varepsilon/n \leq t \leq 1} \left(\frac{\lfloor nt \rfloor}{n} \right) \left\| \frac{S_{\lfloor nt \rfloor}}{\lfloor nt \rfloor} - \mu \right\| + \sup_{0 \leq t \leq 1} \left| \frac{\lfloor nt \rfloor}{n} - t \right| \|\mu\| \\ &\leq \varepsilon + \frac{\|\mu\|}{n}. \end{aligned} \tag{A.2.1}$$

On the other hand,

$$\sup_{0 \leq t \leq N_\varepsilon/n} \|X'_n(t) - \mu t\| \leq \frac{1}{n} \max_{0 \leq k \leq N_\varepsilon} \|S_k\| + \frac{N_\varepsilon \|\mu\|}{n} \xrightarrow{\text{a.s.}} 0 \text{ as } n \to \infty, \tag{A.2.2}$$

since $\mathbb{P}(N_\varepsilon < \infty) = 1$. Thus combining (A.2.1) and (A.2.2) we obtain

$$\limsup_{n\to\infty} \sup_{0\le t\le 1} \|X_n'(t) - \mu t\| \le \varepsilon,$$

and since $\varepsilon > 0$ was arbitrary, we get $\rho_\infty(X_n', I_\mu) \xrightarrow{\text{a.s.}} 0$, proving part (b).

Let $X_n''(t) = S_{\lfloor nt\rfloor+1}$. A similar argument to that above shows that, for $n \ge 1$,

$$\begin{aligned}\sup_{N_\varepsilon/n\le t\le 1} \|X_n''(t) - \mu t\| &\le \sup_{N_\varepsilon/n\le t\le 1} \left(\frac{\lfloor nt\rfloor + 1}{n}\right) \left\| \frac{S_{\lfloor nt\rfloor+1}}{\lfloor nt\rfloor + 1} - \mu \right\| + \sup_{0\le t\le 1} \left| \frac{\lfloor nt\rfloor + 1}{n} - t \right| \|\mu\| \\ &\le 2\varepsilon + \frac{\|\mu\|}{n},\end{aligned}$$

and

$$\sup_{0\le t\le N_\varepsilon/n} \|X_n''(t) - \mu t\| \le \frac{1}{n} \max_{0\le k\le N_\varepsilon+1} \|S_k\| + \frac{N_\varepsilon\|\mu\|}{n} \xrightarrow{\text{a.s.}} 0 \text{ as } n \to \infty.$$

It follows that $\rho_\infty(X_n''(t), I_\mu) \xrightarrow{\text{a.s.}} 0$ as well. Let $\alpha_n(t) = nt - \lfloor nt\rfloor$; note that $\alpha_n(t) \in [0, 1)$ for all $n \ge 1$ and all $t \in [0, 1]$. Then

$$X_n(t) = X_n'(t) + n^{-1}\alpha_n(t)\xi_{\lfloor nt\rfloor+1} = (1 - \alpha_n(t))X_n'(t) + \alpha_n(t)X_n''(t),$$

so that

$$\begin{aligned}\rho_\infty(X_n, I_\mu) &= \sup_{0\le t\le 1} \|(1 - \alpha_n(t))(X_n'(t) - I_\mu(t)) + \alpha_n(t)(X_n''(t) - I_\mu(t))\| \\ &\le \sup_{0\le t\le 1} |1 - \alpha_n(t)|\|X_n'(t) - I_\mu(t)\| + \sup_{0\le t\le 1} |\alpha_n(t)|\|X_n''(t) - I_\mu(t)\| \\ &\le \rho_\infty(X_n', I_\mu) + \rho_\infty(X_n'', I_\mu),\end{aligned}$$

which tends to 0 a.s., establishing part (a). □

A.3 Functional central limit theorems

The final section is dedicated to proving the functional central limit theorem, Theorem 3.1.5, both in the space of continuous and discontinuous trajectories. This result in the one-dimensional case was proved by Donsker in 1951 [Don51]. We point the reader to [EK09, §5] for a comprehensive discussion of both d-dimensional Brownian motion and the steps leading to this result.

A.3.1 Proof overviews and a motivating example

In order to prove both the mappping theorem and Donsker's theorem, we will need to delve further into weak convergence theory. First, in Section A.3.2 we will present different characterisations of weak convergence and note Slutsky's theorem in this context, all of which will be necessary for the proofs. Then we will present the proof of the mapping theorem.

For the proof of Donsker's theorem, it could be suggested that a sufficient method would be to take some finite number of points on the trajectory, and see if the distribution of their location converges to the equivalent distribution for such points on a Brownian path. We now demonstrate why this will not be sufficient.

First, for $t \in [0,1]$ and $f \in \mathcal{M}^d$, recall the *projection* $\pi_t : \mathcal{M}^d \to \mathbb{R}^d$ is denoted $\pi_t f := f(t)$. More generally, for $k \in \mathbb{N}$ and $t_1, t_2, \dots, t_k \in [0,1]$, we define $\pi_{t_1,t_2,\dots,t_k} : \mathcal{M}^d \to (\mathbb{R}^d)^k$ by

$$\pi_{t_1,\dots,t_k} f := (f(t_1), \dots, f(t_k)).$$

We say the finite-dimensional distributions of a function converge if we have the following,

(FDD) (i) If $X, X_1, X_2, \dots$ is a sequence in $\mathcal{C}^d$ then, for all $t_i \in [0,1]$,

$$\begin{aligned}\pi_{t_1,t_2,\dots,t_k} X_n &= (X_n(t_1), X_n(t_2), \dots, X_n(t_k)) \\ &\Rightarrow (X(t_1), X(t_2), \dots, X(t_k)) = \pi_{t_1,t_2,\dots,t_k} X,\end{aligned}$$

where the convergence is on $(\mathbb{R}^d)^k$.

(ii) If $X, X_1, X_2, \dots$ is a sequence in $\mathcal{D}^d$ then,

$$\begin{aligned}\pi_{t_1,t_2,\dots,t_k} X_n &= (X_n(t_1), X_n(t_2), \dots, X_n(t_k)) \\ &\Rightarrow (X(t_1), X(t_2), \dots, X(t_k)) = \pi_{t_1,t_2,\dots,t_k} X,\end{aligned}$$

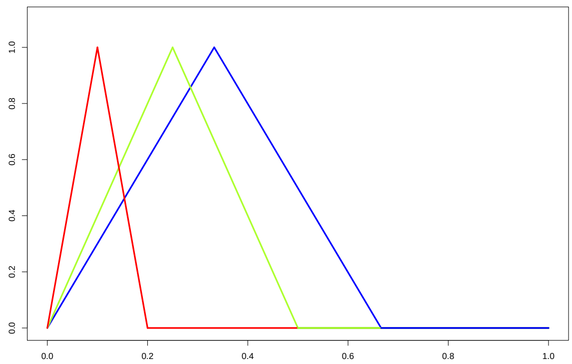
where the convergence is on $(\mathbb{R}^d)^k$ and holds for all $(t_1, t_2, \dots, t_k)$ such that each π_{t_i} is continuous.

Note that, in both cases, the weak convergence on $(\mathbb{R}^d)^k$ is convergence in distribution.

Noting that $\|\pi_t f - \pi_t g\| \leq \|f - g\|_\infty$ and so $\|\pi_{t_1,t_2,\ldots,t_k} f - \pi_{t_1,t_2,\ldots,t_k} g\| \leq \sqrt{k}\|f - g\|_\infty$, it follows that the projection is a continuous function from $(\mathcal{C}^d, \rho_\infty)$ to $(\mathbb{R}^d, \rho_E)$, hence it is a direct consequence of the mapping theorem, Theorem 3.1.6, that, if $X_n \Rightarrow X$ on $\mathcal{C}^d$, then the finite-dimensional distributions also converge. Unfortunately, the reverse is not necessarily true; there exist sequences of probability measures whose finite-dimensional distributions converge weakly, though the measures themselves do not.

Example A.3.1. Consider the following functions, with examples z_3 plotted in blue, z_4 plotted in light-green and z_{10} plotted in red;

$$z_n(t) = \begin{cases} nt & \text{for } t \in [0, 1/n); \\ 2 - nt & \text{for } t \in [1/n, 2/n); \\ 0 & \text{for } t \geq 2/n. \end{cases}$$



If we set $P_n = \delta_{z_n}$, the point mass at the function z_n, and $P = \delta_0$, then as soon as $t_i \geq 2n^{-1}$ for all i, $\pi_{t_1,\ldots,t_k} z_n = (0, \ldots, 0) = \pi_{t_1,\ldots,t_k} 0$, so weak convergence of finite-dimensional distributions holds; but, since $\rho_\infty(z_n, 0) = 1$ for all n, $z_n \not\to 0$ so $P_n \not\Rightarrow P$; we do not have weak convergence.

Based on this example, it is clear that we need a further condition on the family $\{P_n\}$. For trajectories in $\mathcal{C}_0^d$ it happens that such a sufficient condition is relative compactness, but it is hard to directly prove that a family of measures is relatively compact. However, Prokhorov's theorem tells us that tightness implies relative compactness, so we can work with tightness. Finally, we will use a couple of probability bounds on the running maximum of the trajectory to prove the tightness in $\mathcal{C}_0^d$. We complete the proof by showing the finite-dimensional distributions do in fact converge in this case.

Of course, the results for continuous trajectories are only enough to prove part (a) of Theorem 3.1.5. For part (b), we will show that tightness is still a sufficient condition to ensure the finite-dimensional limit is in fact the weak limit of the family of measures.

Then we note some relevant changes to the conditions for tightness which we will prove are satisfied in $\mathcal{D}_0^d$. Finally, we show the finite-dimensional distribution convergence enabling us to conclude the weak convergence statement of Theorem 3.1.5 part (b).

Remark A.3.2. It suffices to prove Donsker's theorem for the case $\Sigma = I_d$. To see this, consider the walk defined at ($\mathbf{W}_\mu$) with $\mu = \mathbf{0}$, for which ($\mathbf{V}$) holds with some arbitrary Σ. Assuming $\sigma^2 > 0$, if any of the eigenvalues of Σ are zero, then the walk is not truly d-dimensional and can be mapped to a walk with smaller dimension such that the covariance matrix for this walk is positive definite, see for example [LL10, p. 4]. Any results where this is the case, would of course then relate to weak convergence to the Brownian path in the lower dimension contained on the hypersurface, and this statement can be mapped back to the original space. Hence, it suffices to assume Σ is positive definite. Indeed, if Σ is positive definite, then the (unique) symmetric square-root $\Sigma^{1/2}$ is also positive definite, and $\Sigma^{1/2}$ has inverse $\Sigma^{-1/2}$. Then set $\zeta := \Sigma^{-1/2}\xi$, and let $\zeta_i = \Sigma^{-1/2}\xi_i$ for $i \in \mathbb{N}$. By linearity of expectation, $\mathbb{E}\,\zeta = \Sigma^{-1/2}\,\mathbb{E}\,\xi = \mathbf{0}$ and

$$\mathbb{E}[\zeta\zeta^\top] = \mathbb{E}[\Sigma^{-1/2}\xi\xi^\top\Sigma^{-1/2}] = \Sigma^{-1/2}\,\mathbb{E}[\xi\xi^\top]\Sigma^{-1/2} = I_d.$$

Let $\tilde{S}_n := \sum_{i=1}^n \zeta_i$ be the random walk associated with ζ. Then $\tilde{S}_n = \Sigma^{-1/2}S_n$, and $\tilde{S}_n$ satisfies ($\mathbf{W}_\mu$) and ($\mathbf{V}$) with $\mu = \mathbf{0}$ and $\Sigma = I_d$. The analogue of Y_n' for $\tilde{S}_n$ is

$$\tilde{Y}_n'(t) = n^{-1/2}\tilde{S}_{\lfloor nt \rfloor} = \Sigma^{-1/2}Y_n'(t),$$

so $Y_n' = \Sigma^{1/2}\tilde{Y}_n'$. The case of Theorem 3.1.5(b) where $\Sigma = I_d$ yields $\tilde{Y}_n' \Rightarrow b_d$. Since $\mathbf{x} \mapsto \Sigma^{1/2}\mathbf{x}$ is continuous, the mapping theorem, Theorem 3.1.6, shows that $Y_n' = \Sigma^{1/2}\tilde{Y}_n' \Rightarrow \Sigma^{1/2}b_d$, which is the conclusion of Theorem 3.1.5(b) in the general case. A similar argument holds for Theorem 3.1.5(a). Thus we can conclude that Donsker's theorem holds for general Σ following from the special case where $\Sigma = I_d$.

A.3.2 Some weak convergence theory

First we note the notion of convergence in distribution for random variables on $\mathbb{R}^d$. Given a random variable X taking values in $\mathbb{R}^d$, we write $X = (X_1, \ldots, X_d)^\top$ in components. The distribution function F of X is defined for $t = (t_1, \ldots, t_d)^\top \in \mathbb{R}^d$ by $F(t) := \mathbb{P}(X_1 \le t_1, \ldots, X_d \le t_d)$.

Definition A.3.3. Let $X, X_1, X_2, \ldots$ be a sequence of $\mathbb{R}^d$-valued random variables with corresponding distribution functions $F, F_1, F_2, \ldots$. Then we say that X_n converges in distribution to X, and write $X_n \xrightarrow{\mathrm{d}} X$, if $\lim_{n\to\infty} F_n(t) = F(t)$ for all continuity points t of F.

We also state a theorem of Pólya, see for example [Leh99, Theorem 2.6.1], which will allow us to take the convergence in the central limit theorem, Theorem 1.3.15, to be uniform convergence.

Theorem A.3.4. *Let $F_1, F_2, \ldots$ be a sequence of cumulative distribution functions such that $F_n \xrightarrow{\mathrm{d}} F$. If F is continuous, then $F_n(x)$ converges to $F(x)$ uniformly in x.*

The Portmanteau theorem (see e.g. [Bil99, Theorem 2.1]) gives several different characterisations of weak convergence. We only state them in terms of probability measures, but throughout consider random variables X_n to be endowed with the respective measure P_n, and hence statements like (ii) could be written as convergence of expectations of the respective random variables, notation we will use later.

Theorem A.3.5 (Portmanteau theorem)**.** *Let $P, P_1, P_2, \ldots$ be probability measures on metric measure space $(S, \mathcal{S}, \rho)$. The following statements are equivalent.*

(i) $P_n \Rightarrow P$.

(ii) $\int_S f \mathrm{d}P_n \to \int_S f \mathrm{d}P$ *for all bounded, uniformly continuous f.*

(iii) $\limsup_{n\to\infty} P_n(F) \leq P(F)$ *for all closed sets F.*

(iv) $\liminf_{n\to\infty} P_n(G) \geq P(G)$ *for all open sets G.*

(v) $\lim_{n\to\infty} P_n(A) = P(A)$ *for all A such that* $P(\partial A) = 0$.

Proof. First, note that (i) implies (ii) by definition.

Next we show that (ii) implies (iii). Let F be a closed set, let $\varepsilon > 0$ and recall $\mathbf{1}_A$ is the indicator function of the set A. Take f defined by $f(x) = (1 - \varepsilon^{-1}\rho(x, F))^+$, so $f(x) = 1$ for $x \in F$ and $f(x) = 0$ for $x \notin F^\varepsilon$, which gives $\mathbf{1}_F(x) \leq f(x) \leq \mathbf{1}_{F^\varepsilon}(x)$.

Thus f is bounded. A simple calculation also shows $|f(x) - f(y)| \leq \varepsilon^{-1}\rho(x, y)$, so f is also uniformly continuous. Then,

$$\limsup_{n\to\infty} P_n(F) = \limsup_{n\to\infty} \int \mathbf{1}_F \mathrm{d}P_n \leq \limsup_{n\to\infty} \int f \mathrm{d}P_n.$$

So by (ii) we get

$$\limsup_{n\to\infty} P_n(F) \leq \int f \mathrm{d}P \leq \int \mathbf{1}_{F^\varepsilon} \mathrm{d}P = P(F^\varepsilon).$$

Take $\varepsilon = 1/k$. Since F is closed, $F = \cap_{k\in\mathbb{N}} F^{1/k}$. Then continuity along monotone limits shows that $P(F^{1/k}) \downarrow P(F)$ as $k \to \infty$, and we obtain (iii).

Next, observe that (iii) is equivalent to (iv) by complementation.

We next show that (iii) and (iv) together imply (v). Indeed,

$$\begin{aligned} P(\operatorname{cl} A) &\geq \limsup_{n\to\infty} P_n(\operatorname{cl} A) \geq \limsup_{n\to\infty} P_n(A) \\ &\geq \liminf_{n\to\infty} P_n(A) \geq \liminf_{n\to\infty} P_n(\operatorname{int} A) \geq P(\operatorname{int} A). \end{aligned}$$

If $P(\partial A) = 0$, then the extreme terms have the same value, and we obtain (v).

Finally, we show that (v) implies (i). Take f bounded and continuous; assume without loss of generality that $0 < f(x) < 1$ for all x. Let $t \geq 0$. Note that $\{x \in S : f(x) > t\}^{\mathrm{c}} = \{x \in S : f(x) \leq t\}$, and, since f is continuous, $\operatorname{cl}\{x \in S : f(x) > t\} \subseteq \{x \in S : f(x) \geq t\}$. Hence

$$\partial\{x \in S : f(x) > t\} \subseteq \{x \in S : f(x) = t\}.$$

Here we have that $P(\{x \in S : f(x) = t\}) = 0$ except for countably many t. To see this, consider $\{t : P(\{x \in S : f(x) = t\}) \in (1/(n+1), 1/n]\}$ for each $n \in \mathbb{N}$. The number of elements in each of these sets must be finite, or the law of total probability is contradicted, and thus we can label the set of t starting with those in the set with $n = 1$, then $n = 2$, and so on, hence there are only countably many of such t.

Using the short-hand $\{f > t\} = \{x \in S : f(x) > t\}$, we have by Fubini's theorem that

$$\int_S f \mathrm{d}P_n = \int_S \int_0^1 \mathbf{1}_{\{f>t\}} \mathrm{d}t \mathrm{d}P_n = \int_0^1 \int_S \mathbf{1}_{\{f>t\}} \mathrm{d}P_n \mathrm{d}t = \int_0^1 P_n(\{f > t\}) \mathrm{d}t,$$

and then by (v) and the bounded convergence theorem, we obtain

$$\int_S f \mathrm{d}P_n = \int_0^1 P_n(\{f > t\})\mathrm{d}t \to \int_0^1 P(\{f > t\})\mathrm{d}t = \int_S f \mathrm{d}P,$$

which completes the proof. □

Another useful consequence of the Portmanteau theorem is the following characterisation of weak convergence [Bil99, Theorem 2.6], which we do state in terms of random variables.

Theorem A.3.6. *$X_n \Rightarrow X$ if and only if every subsequence $\{X_{n_i}\}$ contains a further subsequence converging weakly to X.*

Proof. The 'only if' part is easy: if $X_n \Rightarrow X$, then for any bounded, continuous f we have $\mathbb{E} f(X_n) = \int f \mathrm{d}P_n \to \int f \mathrm{d}P = \mathbb{E} f(X)$, and then by properties of convergence of real numbers we have that any subsequence of $\mathbb{E} f(X_{n_i}) = \int f \mathrm{d}P_{n_i}$ also converges to $\int f \mathrm{d}P = \mathbb{E} f(X)$, i.e., $X_{n_i} \Rightarrow X$.

For the 'if' part, we prove the contrapositive. Suppose that $X_n \not\Rightarrow X$, then $\mathbb{E} f(X_n) = \int f \mathrm{d}P_n \not\to \int f \mathrm{d}P = \mathbb{E} f(X)$ for some bounded, continuous f. We then have that for some subsequence n_i of $\mathbb{N}$ and some $\varepsilon > 0$, $|\mathbb{E} f(X_{n_i}) - \mathbb{E} f(X)| = |\int f \mathrm{d}P_{n_i} - \int f \mathrm{d}P| > \varepsilon$ for all i, so that X_{n_i} has no weakly convergent subsequence. □

A.3.3 Proof of the mapping theorem

The Portmanteau theorem is enough for us to prove the mapping theorem.

Proof of Theorem 3.1.6. Given that $P_n \Rightarrow P$, it follows that for any $F \in S'$,

$$\limsup_{n\to\infty} P_n\left(h^{-1}F\right) \le \limsup_{n\to\infty} P_n\left(\mathrm{cl}(h^{-1}F)\right) \le P\left(\mathrm{cl}(h^{-1}F)\right), \tag{A.3.1}$$

by the equivalence of parts (i) and (iii) of the Portmanteau theorem, Theorem A.3.5. Also, let $F \in \mathcal{S}'$ be closed; then, since h is measurable, $h^{-1}F \in \mathcal{S}$. If $x \in \mathrm{cl}(h^{-1}F)$, then there exist $x_n \in h^{-1}F$ such that $\rho(x_n, x) \to 0$. Since $h(x_n) \in F$, we have $h(x_n) \to h(x) \in \mathrm{cl}\, F = F$ if h is continuous at x. We therefore have

$$D_h^{\mathrm{c}} \cap \mathrm{cl}(h^{-1}F) \subseteq h^{-1}F. \tag{A.3.2}$$

Combining (A.3.2) and (A.3.1) gives

$$\limsup_{n\to\infty} P_n\left(h^{-1}F\right) \le P(\mathrm{cl}(h^{-1}F)) = P\left(D_h^c \cap \mathrm{cl}(h^{-1}F)\right) \le P(h^{-1}F),$$

since $P(D_h^c) = 1$. This holds true for all closed F, thus another application of parts (i) and (iii) of the Portmanteau theorem yields weak convergence of $P_n h^{-1}$ to Ph^{-1}. □

A.3.4 Weak convergence conditions for continuous trajectories

In order to show weak convergence in the case of $\mathcal{C}^d$ we need to show a collection of probability measures on $\mathcal{C}^d$ is *relatively compact* for which we have the following definition stated for an arbitrary measure space.

(RC) A collection of probability measures Π on $(S, \mathcal{S})$ is called relatively compact if for every sequence P_n of elements of Π, there exists a weakly convergent subsequence P_{n_m}.

We say that (**RC**) holds for random variables $X_1, X_2, \ldots$ if (**RC**) holds for probability measures $P_1, P_2, \ldots$ and the random variables and probability measures are associated as described at (1.3.6).

Considering Theorem A.3.6, it seems that the two concepts of relative compactness and convergence of finite-dimensional distributions would be sufficient to determine weak convergence. The following result confirms that this is in fact the case. We state the result for random variables, the result for probability measures can be found as Example 5.1 from [Bil99].

Theorem A.3.7. *For elements $X, X_1, X_2, \ldots$ of $\mathcal{C}^d$, if* (**FDD**) *and* (**RC**) *hold, then $X_n \Rightarrow X$.*

Proof. By (**RC**) we have that any subsequence X_{n_m} has a further subsequence $X_{n_{m_i}}$ such that $X_{n_{m_i}} \Rightarrow Y$ for some random variable Y, possibly depending on the subsequences chosen. Then the mapping theorem implies $\pi_{t_1,\ldots,t_k} X_{n_{m_i}} \Rightarrow \pi_{t_1,\ldots,t_k} Y$. But by (**FDD**), we have $\pi_{t_1,\ldots,t_k} X_{n_{m_i}} \Rightarrow \pi_{t_1,\ldots,t_k} X$, so $\pi_{t_1,\ldots,t_k} X$ has the same distribution as $\pi_{t_1,\ldots,t_k}Y$. Since the class of finite-dimensional sets is a separating class for $\mathcal{C}^d$, see [Bil99, p. 12], this implies that X and Y have the same distribution, and since the subsequences were arbitrary, we have that all such subsequences contain a further subsequence which weakly converges to X. By the 'only if' statement in Theorem A.3.6, we complete the proof. □

It is difficult to prove relative compactness directly; however, a more convenient condition that we can work with and which implies relative compactness in certain spaces is *tightness*. For a family of probability measures tightness is defined as follows.

(T) A family Π of probability measures on metric measure space $(S, \mathcal{S}, \rho)$ is called *tight* if for every $\varepsilon > 0$ there exists a compact $K \in \mathcal{S}$ such that for all $P \in \Pi$, $P(K) > 1 - \varepsilon$.

Again, we use the terminology in the natural way for random variables: a collection $(X_\alpha, \alpha \in I)$ of random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in a metric measure space $(S, \mathcal{S}, \rho)$ is tight if the collection of probability measures $(P_\alpha, \alpha \in I)$, defined by $P_\alpha(B) = \mathbb{P}(X_\alpha \in B)$ for $B \in \mathcal{S}$, is tight.

To formalise the statement tightness implies relative compactness we state the following theorem of Prokhorov [Bil99, Theorems 5.1 & 5.2].

Theorem A.3.8 (Prokhorov's theorem)**.** (**T**) *implies* (**RC**)*. If S is separable and complete, and Π satisfies* (**RC**)*, then Π also satisfies* (**T**)*.*

Here we only need the implication (**T**) implies (**RC**), however note that Theorem A.1.4 tells us $\mathcal{C}^d$ is separable and complete, so we do indeed have that tightness and relative compactness are equivalent in this space.

Instead of replicating the full proof of Billingsley here [Bil99, pp. 59–63], we only give an outline of the proof of the first statement, the proof of the second is brief so we do provide that here.

Proof. Using the tightness, one can construct a sequence of increasing compact sets which cover deterministically large amounts of the probability mass for all the probability measures P_n. Then a measure theory result states that we can use this sequence to construct a countable class of sets for which any element of an arbitrary open set G must lie in one of these sets. Taking the σ-algebra of the compact sets and these countable sets, we get a countable class of compact sets which contain good approximating sets of the arbitrary set G, we will call this class $\mathcal{H}$.

Now, since the class was countable, a Cantor diagonal method allows us to be sure that there exists a subsequence P_{n_i} for which $\lim_{i\to\infty} P_{n_i}H$ exists for all $H \in \mathcal{H}$. Then we will try to find a probability measure P such that

$$P(G) = \sup_{H\subset G} \lim_{i\to\infty} P_{n_i}H.$$

If this was true, then since the supremum is over $H \subset G$ we have $P(G) \leq \liminf_{i\to\infty} P_{n_i}G$, which is condition (iv) of the Portmanteau theorem, Theorem A.3.5 so we have $P_{n_i} \Rightarrow P$ as desired. The proof that such a measure exists can be found at [Bil99, pp. 61–63], we move on to the reverse implication.

Consider a non-decreasing sequence of open sets G_n with $\lim_{n\to\infty} G_n = S$. For each ε, there exists an n for which $P(G_n) > 1 - \varepsilon$ for all $P \in \Pi$, otherwise the relative compactness assumption would mean the limit of this subsequence of bad measures is the whole space but with non-total probability.

Now consider a sequence of open balls $A_{k_1}, A_{k_2}, \ldots$ with radius $1/k$ which cover S, and take n_k such that $P(\cup_{i\leq n_k} A_{k_i}) > 1 - 2^{-k}\varepsilon$ for all $P \in \Pi$ which we can do by the previous fact. Then by completeness of S, there exists a compact set $K \in S$ defined by

$$K = \cap_{k\geq 1} \cup_{i\leq n_k} A_{k_i},$$

with $P(K) > 1 - \varepsilon$ for all $P \in \Pi$, hence tightness holds. □

Corollary A.3.9. *For elements $X, X_1, X_2, \ldots$ of $\mathcal{C}^d$, if* (**FDD**) *and* (**T**) *hold, then $X_n \Rightarrow X$.*

A.3.5 Tightness conditions for continuous trajectories

Having proven that tightness is sufficient, we need to find a way to prove the tightness holds. In order to do this, we first need to state the Arzelà-Ascoli theorem in d-dimensions. The proof at [Rud76, Theorem 7.25] is not dimension dependent so carries across. Recall a subset A of a topological subspace is relatively compact if it has a compact closure.

Theorem A.3.10. *A set A in $\mathcal{C}^d$ is relatively compact if and only if*

$$\sup_{f \in A} \|f(0)\| < \infty \quad \text{and} \quad \lim_{\delta \to 0} \sup_{f \in A} w_f(\delta) = 0.$$

This allows us to generalise the conditions for tightness at [Bil99, Theorem 7.3] to d-dimensions.

Lemma A.3.11. *Let P_n be a sequence of probability measures on $\mathcal{C}^d$. Then* (**T**) *holds if and only if the following two conditions hold.*

(i) We have

$$\lim_{a \to \infty} \limsup_{n \to \infty} P_n(\{f : \|f(0)\| \geq a\}) = 0. \tag{A.3.3}$$

(ii) For each $\varepsilon > 0$,

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} P_n\left(\{f : w_f(\delta) \geq \varepsilon\}\right) = 0. \tag{A.3.4}$$

Proof. For the 'only if' case, given some $\gamma > 0$, consider a compact K such that $P_n(K) > 1 - \gamma$ for all n; such a K exists by the tightness. Since K is compact, Theorem A.3.10 tells us that $\sup_{f \in K} \|f(0)\| < \infty$ so $K \subseteq \{f : \|f(0)\| \leq a\}$ for a large enough choice of a. Further, $\lim_{\delta \to 0} \sup_{f \in K} w_f(\delta) = 0$ so for a small enough choice of δ, $K \subseteq \{f : w_f(\delta) \leq \varepsilon\}$. These two facts imply (A.3.3) and (A.3.4) respectively.

For the reverse implication, we start by recalling Theorem A.1.4 which says that $\mathcal{C}^d$ is separable and complete under ρ_∞. Noting that a single measure clearly satisfies (**RC**), it follows from Prokhorov's theorem, Theorem A.3.8 that a single measure is tight. Then, using the 'only if' part of this lemma, for a fixed probability measure P, and a given $\gamma > 0$ there is an a such that $P(\{f : \|f(0)\| \geq a\}) \leq \gamma$, and for a given ε and γ there is a δ such that $P(\{f : w_f(\delta) \geq \varepsilon\}) \leq \gamma$.

If we have (A.3.3) and (A.3.4), then there exists a finite n_0 such that, for all $n > n_0$,

$$P_n(\{f : \|f(0)\| \geq a\}) \leq \gamma, \tag{A.3.5}$$

holds for some large enough a and

$$P_n(\{f : w_f(\delta) \geq \varepsilon\}) \leq \gamma, \tag{A.3.6}$$

holds for some small enough δ. Then, for each of the finitely many measures $P_1, P_2, \ldots, P_{n_0}$ we have tightness so (A.3.5) and (A.3.6) still hold for these measures, possibly requiring a larger choice of a or smaller choice of δ. Using this, we can assume there exists some a and some δ for which (A.3.5) and (A.3.6) hold for all n.

Using this assumption, given γ, we can choose a and δ_k such that the sets $B = \{f : \|f(0)\| \leq a\}$ and $B_k = \{f : w_f(\delta_k) < 1/k\}$ have probabilities $P_n(B) \geq 1 - \gamma$ and $P_n(B_k) \geq 1 - \gamma 2^{-k}$ for all n. Consider the set $K = \mathrm{cl}(B \cap (\cap_{k\geq 1} B_k))$ which has $P_n(K) \geq 1 - \gamma - \gamma 2^{-k} \geq 1 - 2\gamma$ for all n. This closed set satisfies both conditions of Theorem A.3.10, so it is compact, hence the $\{P_n\}$ are tight. $\square$

The next ingredient we need is a theorem bounding the modulus of continuity which is the d-dimensional equivalent to [Bil99, Theorem 7.4].

Theorem A.3.12. *Suppose that* $0 = t_0 < t_1 < \ldots < t_k = 1$ *and* $\min_{1<i<k}(t_i - t_{i-1}) \geq \delta$. *Then, for arbitrary* $f \in \mathcal{C}^d$,

$$w_f(\delta) \leq 3 \max_{1\leq i\leq k} \sup_{t_{i-1}\leq s\leq t_i} \|f(s) - f(t_{i-1})\|, \tag{A.3.7}$$

and, for any probability measure P *on* $\mathcal{C}^d$,

$$P\{f : w_f(\delta) \geq 3\varepsilon\} \leq \sum_{i=1}^{k} P\left\{f : \sup_{t_{i-1}\leq s\leq t_i} \|f(s) - f(t_{i-1})\| \geq \varepsilon\right\}. \tag{A.3.8}$$

Proof. Let m be the maximum in (A.3.7). If s and t lie in the same interval $I_i = [t_{i-1}, t_i]$, then $\|f(s) - f(t)\| \leq \|f(s) - f(t_{i-1})\| + \|f(t) - f(t_{i-1})\| \leq 2m$. If s and t lie in adjacent intervals I_i and I_{i+1}, then $\|f(s) - f(t)\| \leq \|f(s) - f(t_{i-1})\| + \|f(t_{i-1}) - f(t_i)\| + \|f(t_i) - f(t)\| \leq 3m$. If $|s - t| \leq \delta$ then s and t must either lie in the same interval, or adjacent ones, which proves (A.3.7). The second statement follows by Boole's inequality. $\square$

Next we present a lemma that gives a sufficient condition for tightness in $\mathcal{C}_0^d$.

Lemma A.3.13. *Suppose that we have a random walk as defined at* ($\mathbf{W}_\mu$), *and define* Y_n *as at* (3.1.3). *Then a sufficient condition for* $\{Y_n : n \in \mathbb{N}\}$ *to be tight is*

$$\lim_{\lambda\to\infty} \limsup_{n\to\infty} \lambda^2 \mathbb{P}\left(\max_{0\leq j\leq n} \|S_j\| \geq \lambda\sqrt{n}\right) = 0. \tag{A.3.9}$$

Proof. We will show the two conditions in Lemma A.3.11 hold. The first, (A.3.3), clearly holds, since $Y_n(0) = 0$. For the second condition, we use the bound in (A.3.8). In particular, we take $t_i = m_i/n$ for integers m_i satisfying $0 = m_0 < m_1 < \ldots < m_k = n$. Then the supremum in (A.3.8) becomes a maximum of differences as follows,

$$\begin{aligned}\mathbb{P}\left(w_{Y_n}(\delta) \geq 3\varepsilon\right) &\leq \sum_{i=1}^{k} \mathbb{P}\left(\max_{m_{i-1} \leq j \leq m_i} \frac{\|S_j - S_{m_{i-1}}\|}{\sqrt{n}} \geq \varepsilon\right) \\ &= \sum_{i=1}^{k} \mathbb{P}\left(\max_{0 \leq j \leq m_i - m_{i-1}} \|S_j\| \geq \varepsilon\sqrt{n}\right),\end{aligned}$$

where the equality is due to the identical distribution of the increments. For this to hold, of course we need the choice of m_i to satisfy the condition $\min_{1<i<k}(m_i - m_{i-1})n^{-1} \geq \delta$. We can further simplify this choice by taking $m_i = im$ for each $i < k$ and some $m > 1$. In order to satisfy the criterion we take $m = \lceil n\delta \rceil$. By this choice, we naturally fix $k = \lceil n/m \rceil$, with $m_k = n$. Note that this means, for large enough n, $|k - \delta^{-1}| \leq 1$, so for large enough n and $\delta < 1$, we have $k < 2\delta^{-1}$. Also, for large enough n, $|n/m - \delta^{-1}| < 1$ so for large enough n and $\delta < 1/2$, we have $n > m/2\delta$. Using these inequalities, we have, for large enough n and small enough δ,

$$\mathbb{P}\left(w_{Y_n}(\delta) \geq 3\varepsilon\right) \leq \sum_{i=1}^{k} \mathbb{P}\left(\max_{0 \leq j \leq m_i - m_{i-1}} \|S_j\| \geq \varepsilon\sqrt{n}\right) \leq \frac{2}{\delta}\mathbb{P}\left(\max_{0 \leq j \leq m} \|S_j\| \geq \frac{\varepsilon\sqrt{m}}{\sqrt{2\delta}}\right).$$

If we now take $\lambda = \varepsilon/\sqrt{2\delta}$, we get,

$$\limsup_{n\to\infty} \mathbb{P}\left(w_{Y_n}(\delta) \geq 3\varepsilon\right) \leq \frac{4\lambda^2}{\varepsilon^2} \limsup_{m\to\infty} \mathbb{P}\left(\max_{0 \leq j \leq m} \|S_j\| \geq \lambda\sqrt{m}\right).$$

Now, under the suggested condition (A.3.9), for a fixed ε and any $\gamma > 0$, there exists a λ such that

$$\frac{4\lambda^2}{\varepsilon^2} \limsup_{m\to\infty} \mathbb{P}\left(\max_{0 \leq j \leq m} \|S_j\| \geq \lambda\sqrt{m}\right) < \gamma.$$

Fixing ε and a large enough λ means fixing δ to be small enough. The second condition in Lemma A.3.11 follows, and the proof is complete. $\square$

A.3.6 Donsker's theorem for d-dimensional continuous trajectories - proof

We need two final pieces before completing the proof of Donsker's theorem in d-dimensions. First, this is the d-dimensional version of the inequality of Etemadi [Bil12, Theorem 22.5].

Lemma A.3.14. *Let $S_n = \sum_{i=1}^n Z_i$ be a random walk on $\mathbb{R}^d$. Then for any $x \geq 0$,*

$$\mathbb{P}\left(\max_{0\leq j\leq n} \|S_j\| \geq 3x\right) \leq 3 \max_{0\leq j\leq n} \mathbb{P}(\|S_j\| \geq x).$$

Proof. For given x and fixed n, let

$$B_k := \left\{\max_{0\leq j\leq k-1} \|S_j\| \leq 3x\right\} \cap \{\|S_k\| \geq 3x\}$$
$$B := \bigcup_{k=1}^{n} B_k = \left\{\max_{0\leq k\leq n} \|S_k\| \geq 3x\right\}$$

Then the B_k are disjoint for $x > 0$, and for $k \leq n$, by the triangle inequality,

$$B_k \cap \{\|S_n\| \leq x\} \subseteq B_k \cap \{\|S_n - S_k\| > 2x\},$$

and the terms on the right hand side are independent of each other. We therefore have that,

$$\begin{aligned}
\mathbb{P}(B) &= \mathbb{P}\left(B \cap \{\|S_n\| > x\}\right) + \mathbb{P}\left(B \cap \{\|S_n\| \leq x\}\right) \\
&\leq \mathbb{P}\left(\|S_n\| > x\right) + \mathbb{P}\left(B \cap \{\|S_n\| \leq x\}\right) \\
&= \mathbb{P}\left(\|S_n\| > x\right) + \sum_{k=1}^{n} \mathbb{P}\left(B_k \cap \{\|S_n\| \leq x\}\right) \\
&\leq \mathbb{P}\left(\|S_n\| > x\right) + \sum_{k=1}^{n} \mathbb{P}\left(B_k \cap \{\|S_n - S_k\| > 2x\}\right) \\
&\leq \mathbb{P}\left(\|S_n\| > x\right) + \sum_{k=1}^{n} \mathbb{P}(B_k)\mathbb{P}\left(\|S_n - S_k\| > 2x\right) \\
&\leq \mathbb{P}\left(\|S_n\| > x\right) + \max_{k\leq n} \mathbb{P}\left(\|S_n - S_k\| > 2x\right) \\
&\leq \mathbb{P}\left(\|S_n\| > x\right) + \max_{k\leq n} \left[\mathbb{P}\left(\|S_n\| > x\right) + \mathbb{P}\left(\|S_k\| > x\right)\right] \\
&\leq 3\max_{k\leq n} \mathbb{P}\left(\|S_k\| > x\right),
\end{aligned}$$

as required. □

Using Etemadi's inequality, we can state the following estimate, which we have as a separate lemma because it will be useful in the proof of both parts of Theorem 3.1.5, not just in the continuous case.

Lemma A.3.15. *Suppose that we have a random walk as defined at* ($\mathbf{W}_\mu$) *with* $\mu = \mathbf{0}$, *and satisfying* ($\mathbf{V}$) *with* $\Sigma = I_d$. *Then there exists a constant* $C \in \mathbb{R}_+$ *such that for all* $k \in \mathbb{N}$ *and all* $\lambda \geq 0$,

$$\limsup_{n\to\infty} \mathbb{P}\left(\max_{0\leq j\leq \lfloor n/k \rfloor} \|S_j\| \geq \lambda\sqrt{n} \right) \leq Ck^{-2}\lambda^{-4}.$$

Proof. Let $Z \sim \mathcal{N}(\mathbf{0}, I_d)$. Then by Markov's inequality there is a constant $C \in \mathbb{R}_+$ depending only on d such that, for all $a \geq 0$,

$$\mathbb{P}\left(\|Z\| \geq \frac{a}{3} \right) = \mathbb{P}\left(\|Z\|^4 \geq \left(\frac{a}{3}\right)^4 \right) \leq Ca^{-4}. \tag{A.3.10}$$

We apply the d-dimensional version of Etemadi's inequality, see Lemma A.3.14, to obtain, for $\lambda \geq 0$,

$$\mathbb{P}\left(\max_{0\leq j\leq \lfloor n/k \rfloor} \|S_j\| \geq \lambda\sqrt{n} \right) \leq 3 \max_{0\leq j\leq \lfloor n/k \rfloor} \mathbb{P}\left(\|S_j\| \geq \frac{\lambda\sqrt{n}}{3} \right).$$

Now for any $n_0 \in \mathbb{N}$,

$$\begin{aligned}
\max_{0\leq j\leq \lfloor n/k \rfloor} \mathbb{P}\Big(\|S_j\| \geq & \frac{\lambda\sqrt{n}}{3} \Big) \\
&\leq \max_{0\leq j\leq n_0} \mathbb{P}\left(\|S_j\| \geq \frac{\lambda\sqrt{n}}{3} \right) + \max_{n_0\leq j\leq \lfloor n/k \rfloor} \mathbb{P}\left(\|S_j\| \geq \frac{\lambda\sqrt{n}}{3} \right) \\
&\leq \max_{0\leq j\leq n_0} \mathbb{P}\left(\|S_j\| \geq \frac{\lambda\sqrt{n}}{3} \right) + \max_{n_0\leq j\leq \lfloor n/k \rfloor} \mathbb{P}\left(j^{-1/2}\|S_j\| \geq \frac{\lambda\sqrt{n/j}}{3} \right) \\
&\leq \max_{0\leq j\leq n_0} \mathbb{P}\left(\|S_j\| \geq \frac{\lambda\sqrt{n}}{3} \right) + \max_{n_0\leq j\leq \lfloor n/k \rfloor} \mathbb{P}\left(j^{-1/2}\|S_j\| \geq \frac{\lambda\sqrt{k}}{3} \right). \\
&\leq \max_{0\leq j\leq n_0} \mathbb{P}\left(\|S_j\| \geq \frac{\lambda\sqrt{n}}{3} \right) + \max_{j\geq n_0} \mathbb{P}\left(j^{-1/2}\|S_j\| \geq \frac{\lambda\sqrt{k}}{3} \right).
\end{aligned}$$

Now if we consider Theorem 1.3.15 in conjunction with Theorem A.3.4, and the $a = \lambda\sqrt{k}$ case of (A.3.10), then we can choose n_0 sufficiently large so that for all $k \in \mathbb{N}$ and all $\lambda \geq 0$,

$$\max_{j\geq n_0} \mathbb{P}\left(j^{-1/2}\|S_j\| \geq \frac{\lambda\sqrt{k}}{3} \right) \leq 2Ck^{-2}\lambda^{-4}.$$

Therefore,

$$\limsup_{n\to\infty}\left\{\max_{0\le j\le n_0}\mathbb{P}\left(\|S_j\|\ge\frac{\lambda\sqrt{n}}{3}\right)+\max_{j\ge n_0}\mathbb{P}\left(j^{-1/2}\|S_j\|\ge\frac{\lambda\sqrt{k}}{3}\right)\right\}\le 2Ck^{-2}\lambda^{-4},$$

which gives the claimed result. □

Now we are ready to complete the statement that the measures associated with trajectories in $\mathcal{C}_0^d$ are tight, so we must turn our attention to showing that the finite-dimensional distributions do in fact converge to those of Brownian motion. The following lemma will again be useful for both the continuous and discontinuous cases, hence we state it as a separate result.

Lemma A.3.16. *Suppose that we have a random walk as defined at* $(\mathbf{W}_\mu)$ *with* $\mu = \mathbf{0}$, *and satisfying* $(\mathbf{V})$ *with* $\Sigma = I_d$. *Then for any* $0 \le t_1 < t_2 < \cdots < t_k \le 1$, *we have that as* $n \to \infty$,

$$n^{-1/2}\Big(S_{\lfloor nt_1\rfloor}, S_{\lfloor nt_2\rfloor}-S_{\lfloor nt_1\rfloor},\ldots,S_{\lfloor nt_k\rfloor}-S_{\lfloor nt_{k-1}\rfloor}\Big)$$
$$\xrightarrow{\mathrm{d}} \left(b_d(t_1), b_d(t_2)-b_d(t_1),\ldots,b_d(t_k)-b_d(t_{k-1})\right).$$

Proof. The idea is contained already in the case $k = 2$, so for simplicity we present that case here. By the Markov property, $S_{\lfloor nt_2\rfloor}-S_{\lfloor nt_1\rfloor}$ and $S_{\lfloor nt_1\rfloor}$ are independent. By the multidimensional central limit theorem, Theorem 1.3.15, we have

$$\frac{1}{\sqrt{n}}S_{\lfloor nt_1\rfloor}=\left(\frac{\sqrt{\lfloor nt_1\rfloor}}{\sqrt{n}}\right)\frac{1}{\sqrt{\lfloor nt_1\rfloor}}S_{\lfloor nt_1\rfloor}\xrightarrow{\mathrm{d}} t_1^{1/2}Z_1,$$

where $Z_1 \sim \mathcal{N}(\mathbf{0}, I_d)$, using the fact that, if $\alpha_n \to \alpha$ in $\mathbb{R}$ and $\zeta_n \xrightarrow{\mathrm{d}} \zeta$ in $\mathbb{R}^d$, then $\alpha_n\zeta_n \xrightarrow{\mathrm{d}} \alpha\zeta$ in $\mathbb{R}^d$. Similarly,

$$\frac{1}{\sqrt{n}}\left(S_{\lfloor nt_2\rfloor}-S_{\lfloor nt_1\rfloor}\right)\xrightarrow{\mathrm{d}}(t_2-t_1)^{1/2}Z_2,$$

where $Z_2 \sim \mathcal{N}(\mathbf{0}, I_d)$. Here Z_2 is independent of Z_1 because if $X_n \xrightarrow{\mathrm{d}} X$ and $Y_n \xrightarrow{\mathrm{d}} Y$, and X_n and Y_n are pairwise independent, then $(X_n, Y_n) \xrightarrow{\mathrm{d}} (X, Y)$ where (X, Y) are independent. □

Now we can complete the proof of part (a) of Donsker's theorem.

Proof of Theorem 3.1.5(a). We follow [Bil99, §8], and aim to apply Corollary A.3.9. Recall from Remark A.3.2 that it suffices to consider the case where $\Sigma = I_d$.

First we must establish convergence of the finite-dimensional distributions of Y_n. We need to show that for any $0 \le t_1 < t_2 < \cdots < t_k \le 1$ we have

$$(Y_n(t_1), Y_n(t_2), \ldots, Y_n(t_k)) \xrightarrow{\text{d}} (b(t_1), b(t_2), \ldots, b(t_k)).$$

By continuity of the function $(x_1, x_2, \ldots, x_k) \mapsto \left(x_1, x_1 + x_2, \ldots, \sum_{i=1}^k x_i\right)$, it is sufficient to prove that

$$(Y_n(t_1), Y_n(t_2)-Y_n(t_1), \ldots, Y_n(t_k)-Y_n(t_{k-1})) \xrightarrow{\text{d}} (b(t_1), b(t_2)-b(t_1), \ldots, b(t_k)-b(t_{k-1})).$$

Lemma A.3.16 provides the main step here, but there is a little more work due to the definition of Y_n in terms of interpolation. Again, the main idea is contained in the case $k = 2$ so we describe only that case here. Let $0 \le t_1 < t_2 \le 1$. Using (3.1.3) we may write

$$(Y_n(t_2), Y_n(t_2) - Y_n(t_1)) = \frac{1}{\sqrt{n}} \left(S_{\lfloor nt_1 \rfloor}, S_{\lfloor nt_2 \rfloor} - S_{\lfloor nt_1 \rfloor}\right) + (\psi_{n,t_1}, \psi_{n,t_2} - \psi_{n,t_1}),$$

where $\psi_{n,t} := \frac{nt - \lfloor nt \rfloor}{\sqrt{n}} \xi_{\lfloor nt \rfloor + 1}$. Using Markov's inequality, we have that for $r > 0$,

$$\mathbb{P}(\|\xi\| \ge r) \le \frac{\mathbb{E}[\|\xi\|^2]}{r^2} = \frac{\text{tr } \Sigma}{r^2} = \frac{d}{r^2},$$

since $\mu = 0$ and $\Sigma = I_d$. Since $\|\psi_{n,t}\| \le n^{-1/2} \|\xi_{\lfloor nt \rfloor + 1}\|$, we get

$$\mathbb{P}(\|\psi_{n,t}\| > r) \le \mathbb{P}\left(\|\xi_{\lfloor nt \rfloor + 1}\| \ge r\sqrt{n}\right) \le \frac{d}{r^2 n}.$$

It follows that $\psi_{n,t_1} \xrightarrow{\text{P}} 0$, and similarly for $\psi_{n,t_2} - \psi_{n,t_1}$. Hence $(\psi_{n,t_1}, \psi_{n,t_2} - \psi_{n,t_1}) \xrightarrow{\text{P}} 0$. Thus by Lemma A.3.16 and Theorem 1.3.13, we get

$$(Y_n(t_2), Y_n(t_2) - Y_n(t_1)) \xrightarrow{\text{d}} \left(t_1^{1/2} Z_1, (t_2 - t_1)^{1/2} Z_2\right),$$

which is exactly the distribution of $(b(t_1), b(t_2) - b(t_1))$, as required.

Next we use Lemma A.3.13 to establish tightness. The $k = 1$ case of Lemma A.3.15 shows that

$$\limsup_{n\to\infty} \lambda^2 \mathbb{P}\left(\max_{0 \le j \le n} \|S_j\| \ge \lambda\sqrt{n}\right) \le C\lambda^{-2},$$

which converges to 0 as $\lambda \to \infty$. Thus Lemma A.3.13 gives tightness, and Theorem A.3.9 completes the proof of part (a) of Theorem 3.1.5. □

A.3.7 Weak convergence conditions in the Skorokhod topology

Now we turn to part (b) of Theorem 3.1.5.

The first difference for trajectories with discontinuities is that the spaces $\mathcal{D}$ and $\mathcal{D}^d$ do not automatically have the class of finite-dimensional sets as a separating class. This means the proof of Theorem A.3.7 does not translate to this setting. However, we extract the following result from Theorem 12.5 of [Bil99] which will help us.

Theorem A.3.17. *Let $T \subseteq [0,1]$ with $1 \in T$ such that T is dense in $[0,1]$, then the class of finite-dimensional sets taking values in T is a separating class of $\mathcal{D}^d$.*

To prove this result we recall, without proof, some standard results from measure theory, see e.g. [Dur10, Theorem A.1.4].

Definition A.3.18. Any non-empty collection of sets $\mathcal{P}$ is a π-system if for any $A, B \in \mathcal{P}$, then $A \cap B \in \mathcal{P}$.

Theorem A.3.19. *[Bil12, Theorem 3.3] Suppose that P_1 and P_2 are probability measures on $\sigma(\mathcal{P})$, where $\mathcal{P}$ is a π-system and $\sigma(\mathcal{P})$ is the σ-algebra generated by $\mathcal{P}$. If P_1 and P_2 agree on $\mathcal{P}$ then they agree on $\sigma(\mathcal{P})$.*

We omit the proof of this result because it would require a considerable diversion into Dynkin's $\pi - \lambda$ theorem which is already well covered ground in the literature, see [Bil12, Theorem 3.2].

Proof of Theorem A.3.17. For the duration of this proof, let $\mathcal{B}$ denote the Borel subsets of $(\mathcal{D}^d, \rho_S)$, and recall that $\mathcal{B}_d$ denotes the Borel subsets of $\mathbb{R}^d$. Let $\mathcal{C}$ denote the finite cylinder sets over T, that is, the collection of all subsets of $\mathcal{D}^d$ of the form

$$\left\{ f \in \mathcal{D}^d : \pi_{t_0,t_1,\dots,t_k} f \in \prod_{i=1}^{k} A_i \right\}, \tag{A.3.11}$$

where $k \in \mathbb{Z}_+$, $t_1, \ldots, t_k \in T$, and $A_1, A_2, \ldots, A_k \in \mathcal{B}_d$. If $C_1, C_2 \in \mathcal{C}$ are of the form (A.3.11) with $k = k_1, k_2$ respectively, then $C_1 \cap C_2$ is also a set of the form (A.3.11) with $k = k_1 + k_2$. Thus $\mathcal{C}$ is a π-system. It generates the σ-algebra $\sigma(\mathcal{C})$.

By the assumption that T is dense, there is a sequence $t_1 > t_2 > \cdots$ of elements of T such that $t_n \downarrow 0$ as $n \to \infty$, and then any $f \in \mathcal{D}^d$ has $\pi_0 f = \lim_{n\to\infty} \pi_{t_n} f$ by right continuity. Hence $\pi_0 = \lim_{n\to\infty} \pi_{t_n}$ pointwise, and so π_0 is a limit of functions measurable with respect to $\sigma(\mathcal{C})$, and hence is itself measurable with respect to $\sigma(\mathcal{C})$. Thus we may assume that $0 \in T$. Then, for a given $m \in \mathbb{N}$, choose a positive integer k and points $s_0, s_1, \ldots, s_k$ of T such that $0 = s_0 < \cdots < s_k = 1$ and $\max_{1\le i\le k}(s_i - s_{i-1}) < m^{-1}$. For $\alpha = (\alpha_0, \ldots, \alpha_k)$ in $(\mathbb{R}^d)^{k+1}$, let $V_m\alpha$ be the element of $\mathcal{D}^d$ such that $V_m\alpha(t) = \alpha_{i-1}$ for $t \in [s_{i-1}, s_i)$ for each $1 \le i \le k$, and $V_m\alpha(1) = \alpha_k$. Since $V_m : (\mathbb{R}^d)^{k+1} \to \mathcal{D}^d$ is continuous, it is measurable, i.e., $V_m^{-1}(B) \in \mathcal{B}_{d(k+1)}$ for each $B \in \mathcal{B}$. Since $\pi_{s_0,\ldots,s_k}$ is measurable from $(\mathcal{D}^d, \sigma(\mathcal{C}))$ to $((\mathbb{R}^d)^{k+1}, \mathcal{B}_{d(k+1)})$, the composition $V_m\pi_{s_0,\ldots,s_k}$ is measurable from $(\mathcal{D}^d, \sigma(\mathcal{C}))$ to $(\mathcal{D}^d, \mathcal{B})$. It is a straightforward exercise to show that $\rho_S(f, V_m\pi_{s_0,\ldots,s_k} f) \le \max(m^{-1}, w'_f(m^{-1}))$ for any $f \in \mathcal{D}^d$, which implies that $f = \lim_{m\to\infty} V_m\pi_{s_0,\ldots,s_k} f$. Hence the identity function on $\mathcal{D}^d$ is a limit of a sequence of functions measurable from $(\mathcal{D}^d, \sigma(\mathcal{C}))$ to $(\mathcal{D}^d, \mathcal{B})$ and hence is itself measurable from $(\mathcal{D}^d, \sigma(\mathcal{C}))$ to $(\mathcal{D}^d, \mathcal{B})$. It follows that $\sigma(\mathcal{C}) = \mathcal{B}$, i.e., the π-system $\mathcal{C}$ generates the full Borel σ-algebra. Theorem A.3.19 now completes the proof. □

Now, we can take $T \subseteq [0, 1]$ to be the set of continuity points of $X \in \mathcal{D}^d$, which must contain 1 by the right continuity of $\mathcal{D}^d$ and must be dense because the set of discontinuity points has measure 0. Thus, we have the following replacement of Corollary A.3.9, with the proof now being identical to that of Theorem A.3.7, with the use of Prokhorov's theorem to allow us to claim the result for tightness not relative compactness.

Theorem A.3.20. *For elements $X, X_1, X_2, \ldots$ of $\mathcal{D}^d$, if* (**FDD**) *and* (**T**) *hold, then $X_n \Rightarrow X$.*

A.3.8 Tightness conditions in Skorokhod topology

First we need to state a generalised form of the Arzelà-Ascoli theorem, not only for the Skorokhod topology case, but also in d-dimensions. The proof of the Skorokhod case in 1-dimension was done at [Bil99, Theorem 12.3], but the proof has no dimensional dependency so we refrain from copying it here. Recall the definition of w'_f from (A.1.6).

Theorem A.3.21. *A set A in $\mathcal{D}^d$ is relatively compact if and only if*

$$\sup_{f\in A}\|f\|_\infty < \infty \quad \text{and} \quad \lim_{\delta\to 0}\sup_{f\in A} w'_f(\delta) = 0.$$

Now we can also generalize the tightness conditions of [Bil99, Theorem 13.2] to d-dimensions, the proof reads the same as that for Lemma A.3.11 with the modulus of continuity w_f replaced with w'_f so we omit it.

Lemma A.3.22. *Let P_n be a sequence of probability measures on $\mathcal{D}^d$. Then* (**T**) *holds if and only if the following two conditions hold.*

(i) We have

$$\lim_{a\to\infty}\limsup_{n\to\infty} P_n(\{f : \|f\|_\infty \geq a\}) = 0. \tag{A.3.12}$$

(ii) For each $\varepsilon > 0$,

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty} P_n\left(\{f : w'_f(\delta) \geq \varepsilon\}\right) = 0. \tag{A.3.13}$$

A.3.9 Donsker's theorem in d-dimensional Skorokhod space - proof

Proof of Theorem 3.1.5(b). The convergence of the finite-dimensional distributions is a consequence of Lemma A.3.16 and the continuous mapping theorem, Theorem 3.1.4, which is applicable because the mapping $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k) \mapsto (\mathbf{x}_1, \mathbf{x}_1 + \mathbf{x}_2, \ldots, \sum_{i=1}^k \mathbf{x}_i)$ defined for $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{R}^d$ is continuous. For tightness, it will be sufficient to check the conditions in Lemma A.3.22 applied to the measures P_n defined by $P_n(B) = \mathbb{P}(Y'_n \in B)$. The condition (A.3.12) then becomes

$$\lim_{a\to\infty}\limsup_{n\to\infty}\mathbb{P}\left(\max_{0\leq j\leq n}\|S_j\| \geq a\sqrt{n}\right) = 0,$$

which is easily verified by the $k = 1$ case of Lemma A.3.15.

The condition (A.3.13) becomes

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\mathbb{P}\left(\inf_{\{t_i\}}\max_{1\le i\le v}\sup_{t,s\in[t_{i-1},t_i)}\|Y_n'(s)-Y_n'(t)\|\ge\varepsilon\right)=0,$$

where the infimum is over all δ-sparse sets $\{t_0, t_1, \ldots, t_v\}$. It suffices to suppose $\delta = 1/2k$, with $k \in \mathbb{N}$, and then choose $t_i = i/k$ and $v = k$ to obtain an upper bound for the probability. This gives

$$\begin{aligned}\mathbb{P}\left(\inf_{\{t_i\}}\max_{1\le i\le v}\sup_{t,s\in[t_{i-1},t_i)}\|Y_n'(s)-Y_n'(t)\|\ge\varepsilon\right) &\le \mathbb{P}\left(\max_{1\le i\le v}\sup_{t,s\in[\frac{i-1}{k},\frac{i}{k})}\|Y_n'(s)-Y_n'(t)\|\ge\varepsilon\right)\\ &= \mathbb{P}\left(\bigcup_{i=1}^{k}\left\{\sup_{t,s\in[\frac{i-1}{k},\frac{i}{k})}\|Y_n'(s)-Y_n'(t)\|\ge\varepsilon\right\}\right)\\ &\le \sum_{i=1}^{k}\mathbb{P}\left(\sup_{t,s\in[\frac{i-1}{k},\frac{i}{k})}\|Y_n'(s)-Y_n'(t)\|\ge\varepsilon\right).\end{aligned}$$

Here we have $\|Y_n'(s)-Y_n'(t)\| = \sum_{j=\lfloor ns\rfloor+1}^{\lfloor nt\rfloor}\xi_j$ if $s < t$ (and we can restrict the supremum to such t, s) so that the distribution of $\sup_{t,s\in[\frac{i-1}{k},\frac{i}{k})}\|Y_n'(s)-Y_n'(t)\|$ is the same for each i. Hence

$$\begin{aligned}&\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\mathbb{P}\left(\inf_{\{t_i\}}\max_{1\le i\le v}\sup_{t,s\in[t_{i-1},t_i)}\|Y_n'(s)-Y_n'(t)\|\ge\varepsilon\right)\\ &\qquad\le \lim_{k\to\infty}\limsup_{n\to\infty}k\mathbb{P}\left(\sup_{t,s\in[0,\frac{1}{k})}\|Y_n'(s)-Y_n'(t)\|\ge\varepsilon\right).\end{aligned}$$

Here we have that

$$\begin{aligned}\mathbb{P}\left(\sup_{t,s\in[0,\frac{1}{k})}\|Y_n'(s)-Y_n'(t)\|\ge\varepsilon\right) &\le \mathbb{P}\left(\sup_{t,s\in[0,\frac{1}{k})}(\|Y_n'(s)\|+\|Y_n'(t)\|)\ge\varepsilon\right)\\ &= \mathbb{P}\left(\sup_{t\in[0,\frac{1}{k})}\|Y_n'(t)\|\ge\varepsilon/2\right)\\ &= \mathbb{P}\left(\max_{0\le j\le\lfloor n/k\rfloor}\|S_j\|\ge(\varepsilon/2)\sqrt{n}\right).\end{aligned}$$

Then by Lemma A.3.15 we have that

$$\lim_{k\to\infty}\limsup_{n\to\infty}k\mathbb{P}\left(\sup_{t,s\in[0,\frac{1}{k})}\|Y_n'(s)-Y_n'(t)\|\ge\varepsilon\right)\le\lim_{k\to\infty}Ck^{-1}(\varepsilon/2)^{-4}=0,$$

which verifies condition (A.3.13). This completes the proof of tightness which, with the convergence of the finite-dimensional distributions and Theorem A.3.20, completes the proof. □