COUNTING EXTREME POINTS FROM POISSON PROCESSES ON A HALF LINE

Eric Goodman

A DISSERTATION

in

Mathematics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2022

Supervisor of Dissertation

____

Robin Pemantle, Merriam Term Professor of Mathematics

Graduate Group Chairperson

____

Ron Donagi, Thomas A. Scott Professor of Mathematics

Dissertation Committee:
Bhaswar Bhattacharya, Assistant Professor of Statistics and Data Science
Ryan Hynd, Associate Professor of Mathematics

ACKNOWLEDGEMENTS

First, some academic thanks. At the top of the list, I owe many thanks to Robin Pemantle for sharing ideas and answering questions, for suggesting the problem here and several approaches, for providing friendly guidance and support, and for being extremely patient. This problem got going because of a springboard discussion with Da Wu and Kaitian Jin about one technique here, so thanks also go to the two of them. I also appreciate support from Bhaswar Bhattacharya, who listened kindly and suggested relevant literature, and from Ryan Hynd, who asked helpful, guiding questions.

Thank you also to friends within and around the department for talking, lunching together, even just saying hello.

On the personal side, I would never have completed this extensive process without the constant support I received daily from my family. Talking with them kept me calm, clear-headed, and motivated during these most strenuous and lonely times.

ABSTRACT

COUNTING EXTREME POINTS FROM POISSON PROCESSES ON A HALF LINE

Eric Goodman

Robin Pemantle

Run a Poisson process to generate points on the positive vertical axis, so that the counting process looks like an increasing arc with random jagged edges (see Figure 1.2). The outermost Poisson points—the extreme points—are those that sit on the boundary of the counting process' convex hull. How many extreme points are there? This thesis examines numerous approaches to this question with different styles of answers. Originally, the inspiration for this problem and the purpose of an answer was to guess a growth exponent for the extreme primes studied by McNew (2018), Tutaj (2018), and Pomerance (1979); from estimates here, one might guess 1/3. Upon exploration, the Poisson problem, certain results, and certain techniques herein have unmistakable ties to work by Groeneboom (2011) on a closely related problem about empirical distributions. In fact, the approach by Groeneboom (2011) would likely yield these 1/3 answers for our problem, as well (perhaps even with greater precision than we can provide), though we cannot say with complete certainty, since not all the details were laid out. Moreover, certain techniques here share features with the work by Groeneboom (2011), though the approach here begins from a slightly different point-by-point perspective. We also comment on these similarities and make use of this relationship. Aside from Poisson processes leading to growth exponent 1/3, we study other examples that have growth exponent 1 instead.

Table of Contents

List of Tables

List of Illustrations

Chapter 1

Introduction

In §1.2 we state the main question, but beforehand, we introduce the deterministic version that inspired it in §1.1. An overview of techniques and results plus their similarities to the literature is in §1.3. Unanswered questions are in §1.4, and §1.5 presents notation.

1.1 Inspiration: Extreme Primes

McNew [31], Tutaj [38], and Pomerance [33] studied the following problem and some variants. Plot the sequence of prime numbers $p_n = 2, 3, 5, \ldots$ as (n, p_n), like in Figure 1.1, shade the convex hull of the region to the left of these points, and mark each point (n, p_n) as an extreme point of this convex hull (blue) or as an interior point (red). Now, if $\mathcal{E}_{\text{prime}}(t)$ counts the number of these blue extreme points (n, p_n) with $p_n \leq t$, then how does $\mathcal{E}_{\text{prime}}(t)$ grow as $t \to \infty$? The theorem below states the best bounds known so far. The lower bounds were claimed without proof details by Pomerance [33, p. 407]. Details for these lower bounds as well as the upper bound were provided later by McNew [31]. The stated lower bound was also verified by Tutaj [38, §4.2, see esp. p. 145 and p. 148].

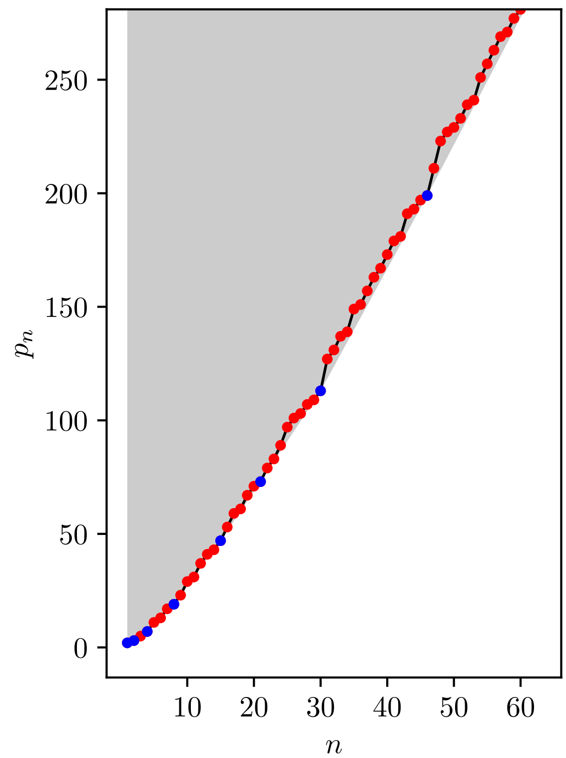

Figure 1.1: The convex hull of the region to the left of the primes, one context studied by McNew [31, cf. p. 126 Fig. 1], Tutaj [38, cf. p. 150], and Pomerance [33].

Theorem 1.1 (See McNew [31, Theorem 2 and Corollaries 1–2] and the references above)**.** As $t \to \infty$,

$$\Omega\left(\frac{t^{1/4}}{\log^{3/2} t}\right) \overset{*}{\leq} \mathcal{E}_{\text{prime}}(t) \leq O\left(\frac{t^{2/3}}{\log^{2/3} t}\right),$$

where $\overset{*}{\leq}$ uses the Riemann Hypothesis, although another bound is available without it.

From computer evidence, McNew [31, p. 138] estimates the exponent of t for the true growth rate is close to ≈ 0.285. Also using numerical evidence, Tutaj [38, Conjecture H] conjectured the exponent of t for the growth rate may be precisely half the Euler constant.

McNew [31, p. 140] also suggested that a random setup might help to clarify this rate, and so Robin Pemantle suggested we replace the primes with a Poisson process of intensity $f(x) = 1/\log x$. The idea here is that (a) this intensity approximates how the prime number theorem says roughly $t/\log t$ primes have $p_n \leq t$, and (b) properties of the primes often agree with properties of similar random sequences (see [37, Ch. 3 §2] or [6, Ch. 1]). Exponents 1/4 and 2/3 in Theorem 1.1 may gently guide reasonable answers for the Poisson process with intensity $1/\log x$, and conversely, knowing an exponent with the Poisson process would lead to an interesting comparison with the prime case. To preview, we will encounter a 1/3 exponent in the $1/\log x$ Poisson context (see Table 1.2, plus the discussion in §1.3.4–§1.3.6 for comparison to other results in the literature, especially some by Groeneboom [19]).

1.2 The Poisson Problem

We now repeat the problem setup of §1.1 using a Poisson process, as Robin Pemantle suggested. However, we consider several possible intensities. Choose a function $f(x) \geq 0$ to be the intensity function of a Poisson process on $(0, \infty)$, which randomly places points $X_1 < X_2 < \cdots$ onto the positive real axis (see [29] for background, or see Lemma 2.1 for a simple construction). The choices $f(x)$ of most interest are

$$\frac{1}{\log x}, \quad \frac{1}{x^q}, \quad \frac{1}{x}, \quad \frac{1}{x \log^q x}, \quad \text{or} \quad \frac{1}{x \log x}, \tag{1.1}$$

where from now on $q \in (0, 1)$ is fixed and $p = 1 - q$. Since these have asymptotes at $x = 0$ or $x = 1$, and since we are interested in tail behavior as $x \to \infty$, we usually set $f(x) = \mathbf{1}[x > C]/\log x$, and similarly in other cases, for some constant C. (Here $\mathbf{1}[\cdots]$ denotes the indicator function.) Sometimes this will be explicit, but usually not.

Now, as in the prime problem in §1.1, plot (j, X_j), fill in the convex hull of the region left of this graph, and color the points (j, X_j) blue if they are extreme points of this hull

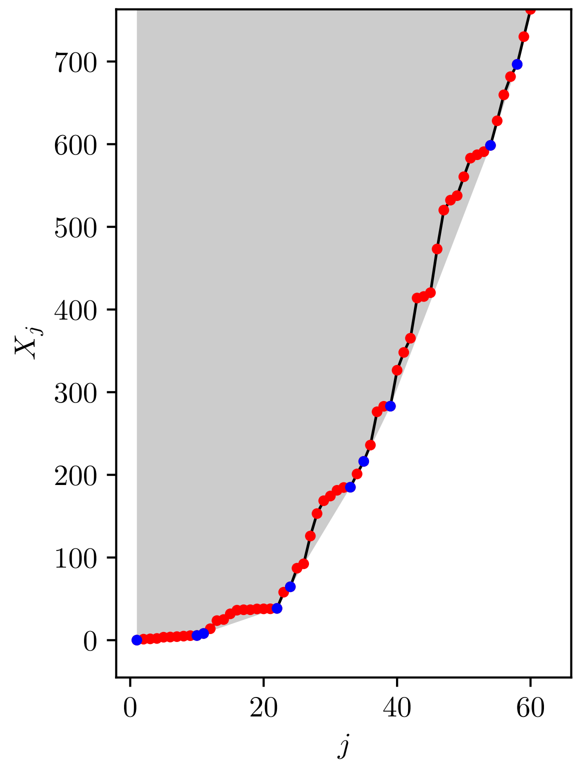

Figure 1.2: The beginning of a Poisson process with intensity $f(x) = 1/\sqrt{x}$.

---

or red otherwise. An example with $f(x) = 1/\sqrt{x}$ is shown in Figure 1.2. The convex hull extends upwards without end, so whether a point is extreme depends on all of the values $X_1, X_2, \ldots$, not just those $X_j \leq t$. Here is the question parallel to that of §1.1.

Question 1.2. How many extreme points (j, X_j) are below height t, as $t \to \infty$? In other words, as $t \to \infty$, what is the behavior of

$$\mathcal{E}_f(t) \stackrel{\text{def}}{=} \#\Big\{(j, X_j) \ : \ X_j \leq t \text{ and } (j, X_j) \text{ is a blue extreme point}\Big\}?$$

This is one kind of count, but here is another. Instead of the number of extreme points $\mathcal{E}_f(t)$ in a growing space, $[0, t]$ as $t \to \infty$, we could also ask for the number of extreme points by index: the random variable $\mathcal{E}_f(X_n)$ is the number of extremes among $X_1, X_2, \ldots, X_n$, which we can think about as $n \to \infty$. While §1.1 motivates counting by space, $[0, t]$, it will at times be useful to think about indices.

Remark 1.3. Given the question and the previewed 1/3 answer here, the reader might already notice similarities to work by Groeneboom in [19]. We have many comments below, especially in §1.3.4, §1.3.5, §1.3.6. See also §1.4. *

Although intensities in (1.1) are of most interest (particularly $1/\log x$ as noted in §1.1), it is reasonable to ask this question if

$$\int_0^t f(x)\,dx < \infty \text{ (when } t < \infty), \quad \int_0^\infty f(x)\,dx = \infty, \quad \text{and} \quad \lim_{x\to\infty} f(x) = 0. \tag{1.2}$$

The first two say the process has infinitely many points but only finitely many in any bounded region, and the last says points spread out as we look farther along the line.

1.3 Overview of Techniques, Results, & Comparisons to the Literature

The emphasis here is not any one example or answer to Question 1.2. Rather, the interest is in presenting various approaches, each of which can give a different kind of answer. With that in mind, we organize the main ideas by technique, instead of by example or result.

Before continuing we should mention some notational tidbits. (More detail is in §1.5.) First, given any particular intensity $f(x)$, denote the mean number of points X_j in $[0, t]$ by

$$\lambda(t) = \int_0^t f(x)\,dx. \tag{1.3}$$

Second, $\approx$ is an intuitive statement, not a rigorous one.

Intensity	Number nonextreme points (a.s.)	Location
$1/\log x$	$= \infty$	Example 3.9
$1/x^q$	$= \infty$	Example 3.8
$1/x$	$= \infty$	Example 3.7
$1/x \log^q x$		
$(0 < q \leq 1/2)$	(?)	Remarks 3.11 and 7.4
$(1/2 < q \leq 1)$	$< \infty$	Examples 3.6 and 3.10

Table 1.1: Finitely many or infinitely many nonextreme points.

1.3.1 Technique 1. Almost-Sure Extreme or Nonextreme Counts

It is often possible to determine when almost surely all but finitely many points will be extreme, or when there will be infinitely many nonextreme points. Table 1.1 summarizes the results for each example. Thanks go to Da Wu and Kaitian Jin for their help in an initial discussion of this problem, during which this Borel-Cantelli strategy took off.

Details are in §3, but here is a taste of the arguments. Notice that if consecutive slopes between points are eventually increasing (Figure 1.3), then only finitely many points can be *nonextreme*; this ultimate behavior can be verified with Borel-Cantelli. This and some very similar observations lead to Table 1.1. Some of these related proofs are analogous to arguments about the primes in Tutaj [38] or Pomerance [33].

1.3.2 Technique 2. Comparison of Intensities

Let $f(x) = 1/x^q$ and $g(x) = 1/x^{q'}$ be two intensities with different exponents $q > q'$. A simple coupling of these two Poisson processes allows us to count the number of extreme

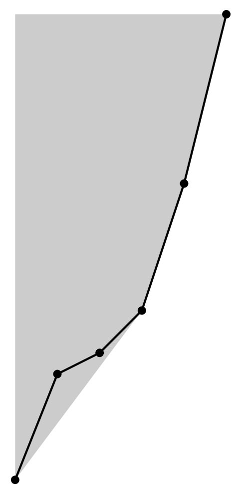

Figure 1.3: Slopes increase after one slope decrease. The fourth point onward is extreme.

points from $f(x)$ relative to the number of extreme points from $g(x)$, index by index. That is, if $X_1, X_2, \ldots$ are the points from $f(x)$ and $Y_1, Y_2, \ldots$ are the points from $g(x)$, then we can couple these so that

$$\mathcal{E}_f(X_n) \geq \mathcal{E}_g(Y_n).$$

The key proof idea is a convexity trick that Pomerance [33] used for the primes; this convexity step will work because of the form of the intensities $1/x^q$, and it will show that each point X_n is extreme whenever Y_n is. This also works with more than two intensities. An example where many exponent choices are all coupled together is shown in Figure 1.4, where we can see certain points switch from nonextreme to extreme as we lower the intensity (move upward along each gray line). Details are in §4.

1.3.3 Technique 3. Sampling Algorithm

As already mentioned, whether points $X_j \leq t$ are extreme depends also on those points $X_k > t$. This suggests that to generate a sample of $\mathcal{E}_f(t)$—or equivalently, to generate a

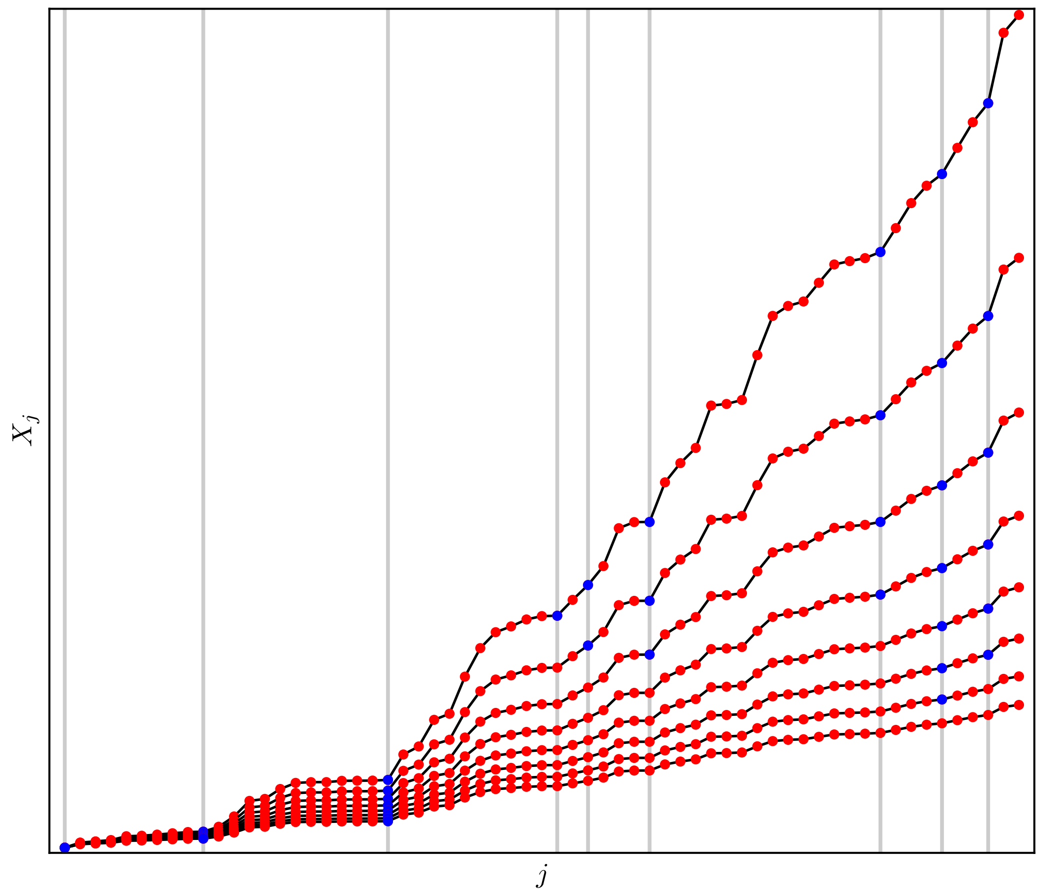

Figure 1.4: A coupling of intensities $1/x^q$ for $q = 0.05$ (increasing slowest), 0.1, 0.15, ..., 0.4 (increasing fastest).

---

bona fide picture of this process with points marked extreme or nonextreme accurately—we might need to know *all of the infinitely many* points. This is not so. There is a simple strategy that checks only finitely many points and works for any intensity satisfying (1.2). A short summary and then the details of the sampling algorithm are presented in §5.

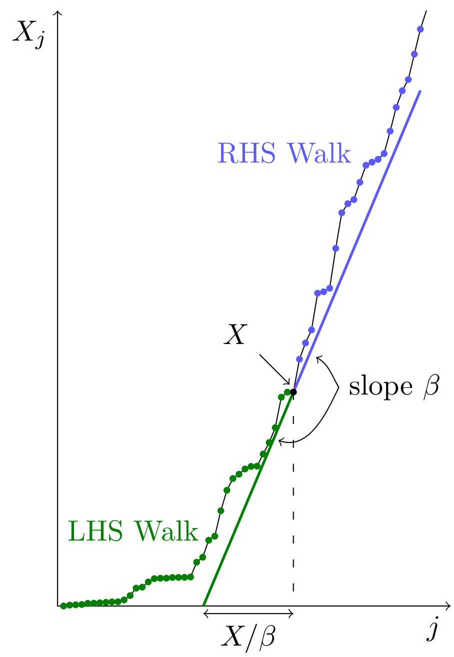

Figure 1.5: The two sides of the random walk.

1.3.4 Technique 4. Expectations Asymptotically

The Borel-Cantelli results in Table 1.1 are informative for low intensities. For higher intensities, Robin Pemantle suggested the following multi-step approach (a)–(d) to find asymptotic bounds on $\mathbb{E}\mathcal{E}_f(t)$ as $t \to \infty$:

(a) View the process around a point in location X (here ignoring the index) as a two-sided random walk, and observe that *if* the point at X is extreme with a particular support line, then that line is a boundary that neither side of this walk can cross (Figure 1.5).

(b) The probability of the event in (a)—where a specific point is extreme and has a given boundary line—may be computed as follows. Rescaling the random walk to Brownian motion, and likewise rescaling the given support-line boundary, the boundary should become roughly parabolic. Approximate with a second-order Taylor polynomial and apply results available in [17], [18], or [25], which provide the probability Brownian motion successfully avoids a given parabola.[1]

(c) Use many possible slopes (Figure 1.6) to estimate the probability a point in location X is extreme, that is, $\mathbb{P}(X \text{ is extreme} \mid X)$. (The conditioning will be explained later.) Robin actually suggested lower and upper bounds on this probability, as follows:

  (i) For a lower bound on the probability that X is extreme, check just one potential support line of some reasonable slope, β_*.

  (ii) Although no one single slope β can give an upper bound on the chance X is extreme, potential support lines may be grouped into zones, say by slopes

$$0 = \beta_0 < \beta_1 < \beta_2 < \cdots < \beta_J = \infty, \tag{1.4}$$

  as in Figure 1.7. If the process remains above *any* support line of slope β within a particular zone $\beta_j \le \beta \le \beta_{j+1}$, then it must avoid that zone's two lower boundary lines. This leads to an upper bound saying, informally,

$$\begin{aligned} \mathbb{P}(X \text{ is extreme} \mid X) &= \sum_j \mathbb{P}\Big(X \text{ is extreme with a support line in zone } j \;\Big|\; X\Big) \\ &\le \sum_j \Big[\mathbb{P}\Big(\text{LHS avoids line of slope } \beta_{j+1} \;\Big|\; X\Big) \\ &\qquad\qquad \cdot \mathbb{P}\Big(\text{RHS avoids line of slope } \beta_j \;\Big|\; X\Big)\Big]. \end{aligned} \tag{1.5}$$

In either case, compute as in (b) above.

(d) Finally, compute the expectation via the Mecke equation (see [29, Theorem 4.1])

$$\mathbb{E}\mathcal{E}_f(t) = \int_0^t f(x) \cdot \mathbb{P}(x \text{ is extreme} \mid x)\, dx. \tag{1.6}$$

[1] To aid intuition, the reader might find Figure 1.6 below reminiscent of a picture of Brownian motion below a parabola in Groeneboom's paper [17, Fig. 4.1]. Of course, the curves in Figure 1.6 are not exactly parabolas.

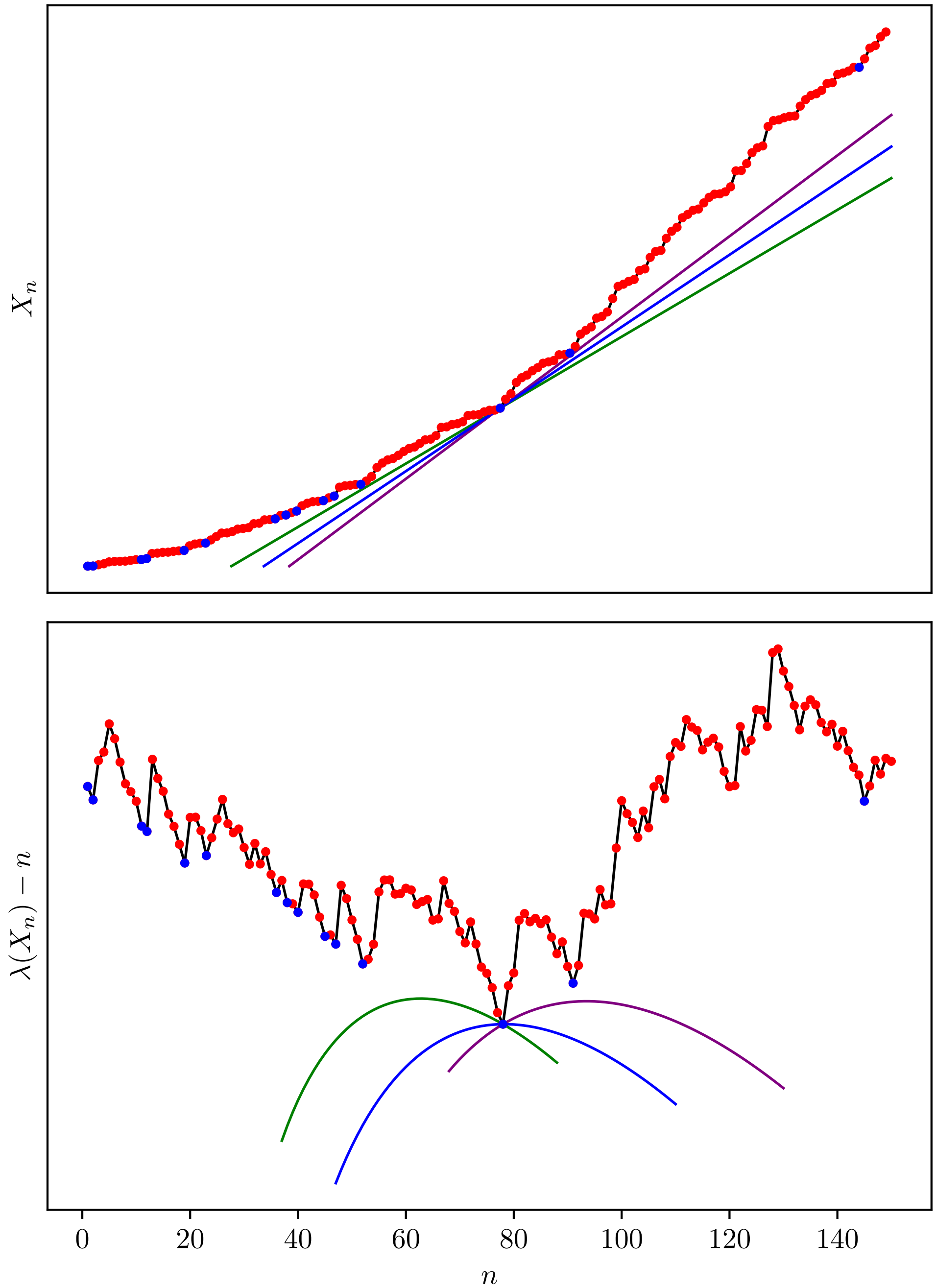

Figure 1.6: (Top) A Poisson process and three support lines through one point. (Bottom) The same process and lines after transforming to a random walk of i.i.d. steps.

Overall, this approach computes probabilities *point-by-point* to estimate the expected number of extreme points. We should mention here that Groeneboom [19, see §3] studied a very similar problem with a related but seemingly more powerful jump-process technique. Actually, one important step claimed without a complete proof in [19, see p. 2255] strongly suggests that Groeneboom's work would also apply to our problem with general Poisson intensities. (More detail will come below in §1.3.5. See especially Remark 1.10.) A long time passed before realizing this approach by Groeneboom [19, p. 2255] very likely also works for our problem, and in the meanwhile, the point-by-point method was carried through for $1/\log x$ and $1/x^q$, leading to the lower and upper bounds given separately in §8–§9. So, if one were to carry out the steps omitted by Groeneboom in [19, p. 2255], then perhaps some of the work below in §8–§9 merely rederives in a weaker form what Groeneboom could show using his jump process. Nevertheless, these point-by-point arguments will be included because of their interest in connection with the prime problem of §1.1.

For $1/\log x$ and $1/x^q$, steps (b)–(c) above are fairly involved, and the upper bounds are especially so: for these we will use the Komlós, Major, Tusnády coupling [27] and lengthy estimates adjusted slightly from some by Groeneboom in [17]. This point-by-point approach shares the same overarching idea as Groeneboom's jump approach from [19]—in particular, rescale to Brownian motion and parabola. What is more, our approach is not as precise (or at least the estimates here are not) as what it sounds like Groeneboom's jump-process approach might yield [19, p. 2255] (see also Remark 1.8), so this point-by-point approach may not have much to encourage its use in all cases.

However, the point-by-point approach is quite useful and easy to apply with low intensities, where the walks simply do not look like Brownian motion with a parabola. (From the limited description Groeneboom gives in [19, p. 2255], it is not clear exactly how the jump approach might change.) This actually helps us. With the intensities $1/x$ and $1/x \log^q x$, the difficult steps of (b)–(c) can be completely skipped. These are treated in §7.

Resulting estimates are shown in Table 1.2. (Upper bounds are within logarithmic factors of lower bounds. See Note 1.25.) These calculations are done example by example. With the more involved examples, $1/x^q$ and $1/\log x$, a fair amount of work is just to adjust parabolic Taylor polynomials so that they actually satisfy inequalities (as opposed to being approximations only). So that these details do not obscure the ideas, §1.3.5 sketches a more general but very informal version of the lower bound argument. During this more general argument, one can see this method for $\mathbb{E}\mathcal{E}_f(t)$ is tightly intertwined with and is essentially redoing some work by Groeneboom [19, Lemma 3.1]. This is partly explained by:

Remark 1.4 (Rescaling here and in Groeneboom [19]). The rescaling factor used for the more general but informal calculation below in §1.3.5 seems to be the same as[2] the rescaling within [19, Lemma 3.1] to Brownain motion minus a parabola. We will see it is this rescaling factor which contributes significantly to the answers. (The rescaling is in (1.12); see also Remark 1.12 and just above it.) Beyond this one rescaling, as just mentioned, the work below and in [19] are similar in that both reduce to expressions derived by Groeneboom in [17] about Brownian motion minus a parabola. It is worth noting that this process—Brownian motion minus a parabola—was related to concave majorants of empirical distributions not only in [19] but also in earlier papers by Groeneboom [15] and Prakasa Rao [34].

The next section, §1.3.5, will be dedicated to explaining (i) the lower bound argument, (ii) connections with results in Groeneboom's work [19], and (iii) other results that help

[2]This is not so clear unless one compares back with [15, p. 542]; see especially the factor $a^{4/3}c^{-2/3}$ in the last displayed equation on that page.

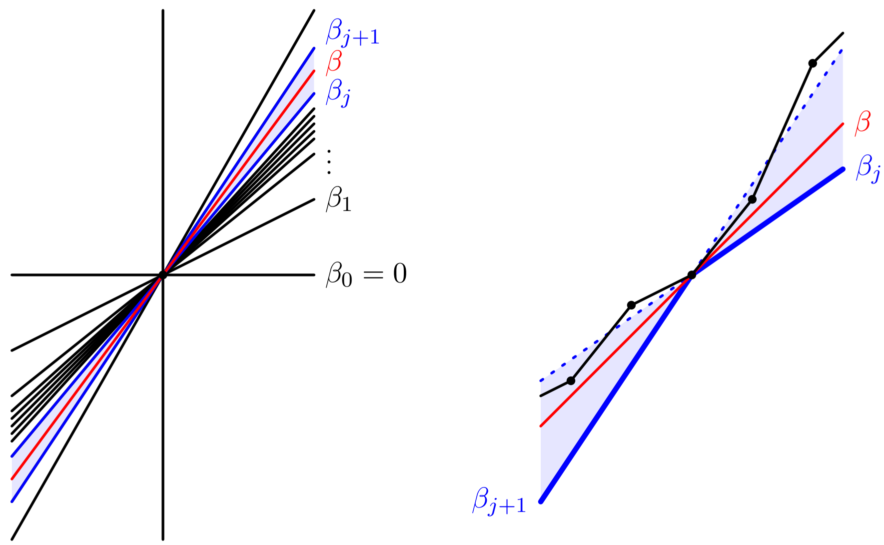

Figure 1.7: Slope Zones. (Left) Black lines of slopes $\beta_0, \beta_1, \ldots$ separate various zones. The red line falls into the zone highlighted in blue, between slopes β_j, β_{j+1}. (Right) If the process is above the red line, it is also above the two thickly drawn blue half-lines.

Intensity	$\mathbb{P}(X \text{ is extreme} \mid X)$	$\mathbb{E}\mathcal{E}_f(t)$	**Location**
$1/\log x$	$\widetilde{\Theta}(1/X^{2/3})$	$\widetilde{\Theta}(t^{1/3})$	Propositions 8.2 and 9.2
$1/x^q$	$\widetilde{\Theta}(1/X^{\frac{2}{3}p})$	$\widetilde{\Theta}(t^{p/3})$	Propositions 8.1 and 9.1
$1/x$	$\Theta(1)$	$\Theta(\log t)$	Proposition 7.1, Lemma 7.3, see also Proposition 7.2
$1/x \log^q x$	$1 - o(1)$	$\sim \log^p t/p$	Proposition 7.9

Table 1.2: Bounds on $\mathbb{P}(X \text{ is extreme} \mid X)$ as $X \to \infty$ and $\mathbb{E}\mathcal{E}_f(t)$ as $t \to \infty$. See Note 1.25 about $\widetilde{\Theta}$.

explain the connection. *

Remark 1.5 (Rescaling and parabolas)**.** Part (b) above tells us to estimate the chance a random walk avoids a curve by rescaling the walk to Brownian motion and simultaneously replacing the walk's boundary curve by a parabolic Taylor expansion. This local parabolic expansion and rescaling method was suggested to me by Robin Pemantle, who in turn learned the technique from Gregory Lawler. Note this description is fairly terse, and to do this, there are really two additional stages that must happen at the start and end of the walk:

- An extreme point sits on its convex support line, so our walk begins on its boundary curve. Before rescaling we need to boost the walk away from the boundary, otherwise after rescaling, the Brownian motion would hit the parabola immediately.
- Since Taylor approximations only work locally, we will need to replace them far away from the starting point. Once the walk is relatively far away from its boundary, we can just use a linear replacement.

In the end, calculations reduce more or less to the quantity $\max_t \left(B_t - \frac{1}{2}t^2\right)$, and to learn about this, (b) points to [17], [18], and [25]. References therein lead to other papers that dealt not only with this quantity but also techniques like those just mentioned—that is, a random walk rescaled alongside a boundary's parabolic Taylor approximation, or a linear-boundary argument far away from a parabolic maximum. The purpose here is not to overview the literature, and the reader should not assume the following list is exhaustive, but a few examples may be found in [7], [8], [3], [4, see esp. Lemmas 1(i), 5, 8], and in a sense also [34]. If interested, a discussion about this technique and about why $\max_t \left(B_t - \frac{1}{2}t^2\right)$ is often relevant (including further references) can be found in [25, §1.2]. *

Remark 1.6 (History)**.** Again, the approach in (a)–(d) above was suggested by Robin Pemantle. He had discussed this problem earlier with a past student who managed to draft certain results but never published them and ceased work on the problem. Although this unpublished manuscript was not shared with me, and although I was not told the name of this other student, it seems appropriate to acknowledge this preceding work, which perhaps included comparable techniques. *

1.3.5 Interlude. Extensive Comparison to Groeneboom [19], Informal Calculation, and Related Comments

We pause now to motivate, introduce, and compare with a result from Groeneboom [19].

Introduction to Groeneboom's Result

Earlier we contrasted $\mathcal{E}_f(t)$, which counts extreme points in $[0,t]$, and $\mathcal{E}_f(X_n)$, which counts extreme points among $X_1,\ldots,X_n$. Let us examine Table 1.2 in the second context. Intensity $f(x) = 1/x^q$ has $X_n \approx (pn)^{1/p}$, so let us set $t = (pn)^{1/p}$ and very, very loosely write $\mathbb{E}\mathcal{E}_f(X_n) \approx \mathbb{E}\mathcal{E}_f(t)$. Using $\mathbb{E}\mathcal{E}_f(t) = \widetilde{\Theta}(t^{p/3})$ from Table 1.2 then suggests we expect $\approx n^{1/3}$ extreme points (ignoring constants and logarithms) among $X_1,\ldots,X_n$. We can do the same thing with $f(x) = 1/\log x$, where $X_n \approx n\log n$ from (3.3). This rate $n^{1/3}$ is also in a very similar problem studied by Groeneboom, who stated the following result with a proof sketch in [19], but some details are forthcoming.

Theorem 1.7 (from Groeneboom [19, Lemma 3.1 and Theorem 3.1])**.** Plot the empirical distribution $G_n(x)$ of n i.i.d. points from a probability density $g(x)$ and count the number of extreme points, $\mathcal{N}_n$, on this plot (see Figure 1.8). If $g(x)$ is a decreasing density of finite support $[0,A]$, and if both $g(x), g'(x)$ are continuous and remain bounded away from zero, then as $n \to \infty$,[3]

$$\begin{cases} \mathbb{E}\mathcal{N}_n \sim k_1 I_g \cdot n^{1/3} \\ \operatorname{Var}\mathcal{N}_n \sim k_2 I_g \cdot n^{1/3} \end{cases} \quad \text{where} \quad \begin{cases} I_g = \displaystyle\int_0^A \left(\frac{(g'(u))^2}{4g(u)} \right)^{1/3} du \\ k_1 \approx 2.10848 \\ k_2 \approx 1.029, \end{cases} \tag{1.7/P.G.1}$$

and moreover

$$\frac{\mathcal{N}_n - k_1 I_g n^{1/3}}{\sqrt{k_2 I_g n^{1/3}}} \overset{d}{\Rightarrow} \mathrm{N}(0,1). \tag{1.8/P.G.2}$$

The resemblance between our Poisson problem and the setup for Groeneboom's result, Theorem 1.7, is hopefully clear: just reflect Figure 1.8 across the diagonal. Yes, some technical details differ, however, as we have been saying already, there are close similarities to what we do here, both in terms of results and arguments. This section compares these extensively.

Remark 1.8 (Results and precision)**.** To introduce Groeneboom's results we sketched an $n^{1/3}$ rate; this is one similarity we notice immediately. However, notice that (1.7/P.G.1) and (1.8/P.G.2) are much sharper than our calculations in Table 1.2, where we can only give expectations up to logarithms or constants. *

Remark 1.9 (Focus on slopes vs. points)**.** Compared to our viewpoint, the perspective in [19] from which Groeneboom proves Theorem 1.7 is slightly different. The focus in [19] is a process parameterized by the decreasing slope of a convex support line, which jumps every time the support line hits a new extreme point. Groeneboom then counts the jumps. To explain this pictorially, consider Figure 1.9, where we selected a density $g(x)$ fitting

[3]We refer to (1.7/P.G.1) and (1.8/P.G.2) repeatedly below. The extra labels P.G.1 and P.G.2 should help the reader quickly recall these are the results quoted from Piet Groeneboom.

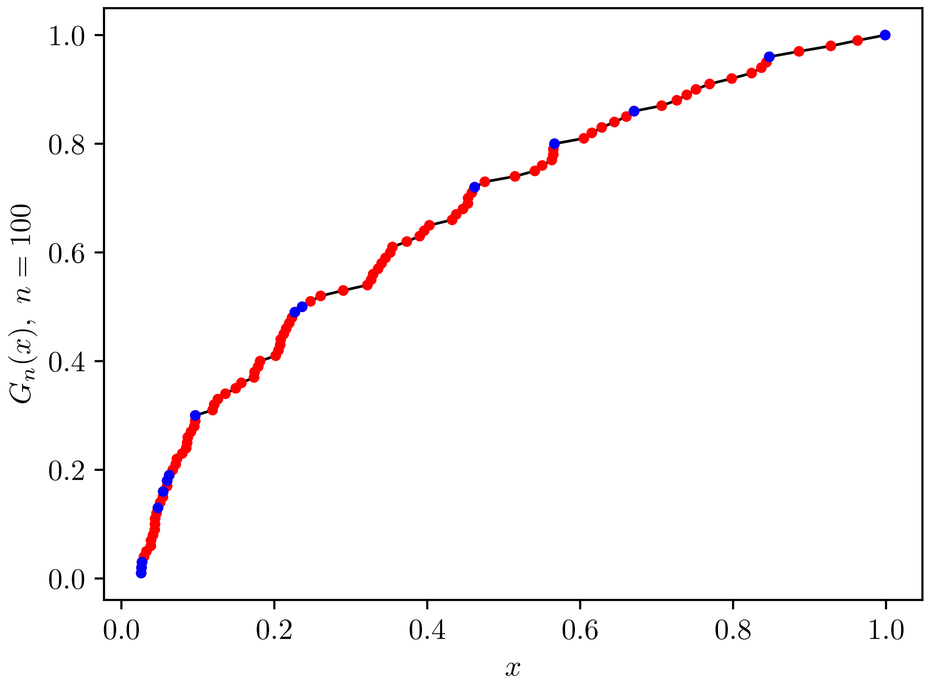

Figure 1.8: An empirical distribution of n samples with its extreme points marked, the context of Groeneboom [19, §3].

Groeneboom's theorem, sampled $n = 100$ points from $g(x)$, and plotted the empirical distribution $G_n(x)$. In this picture, focus first on the green line. If the green line's slope decreases, the line approaches the purple one, then the orange one. Since the orange line touches a later extreme point, Groeneboom's process would jump ahead to that point, and the slopes would continue decreasing while centered around that later point. In essence, Groeneboom asks: given a slope, which point is extreme with a support line of that slope? Contrast this to the point-by-point approach outlined in §1.3.4, which asks instead: given a particular point and no information about the surrounding process, what slopes are likely to make this point extreme, and how likely is it to be extreme with those slopes? With that said, in the

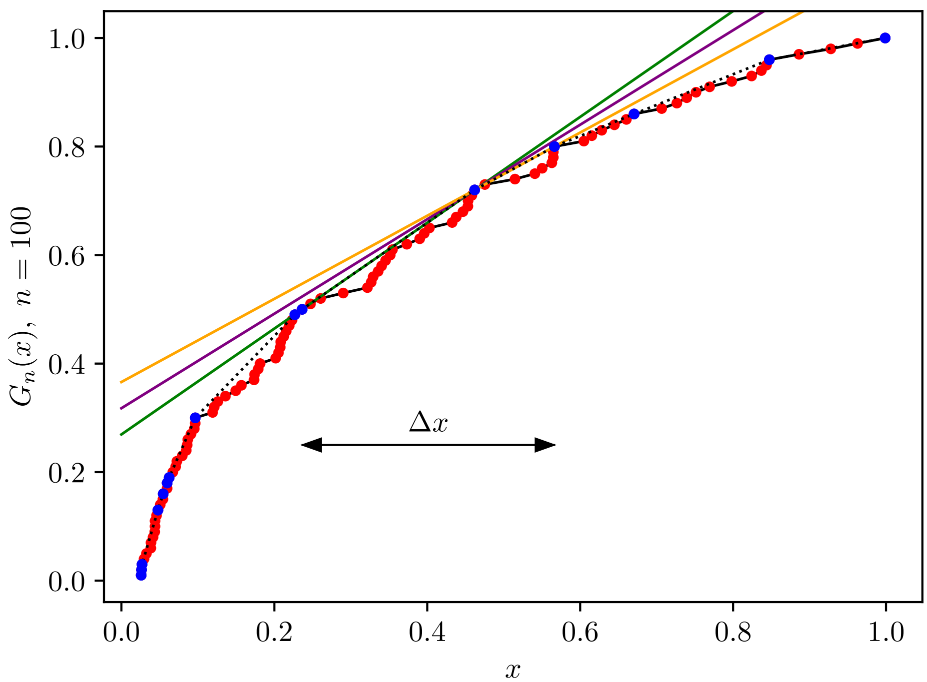

Figure 1.9: An empirical distribution of n samples with the concave majorant and some support lines, the context in Groeneboom [19, §3] and Prakasa Rao [34].

---

end, the work here and in [19] both reduce to Brownian motion with a parabola then build from calculations in [17]. *

Remark 1.10. Groeneboom announced the result quoted above in Theorem 1.7 without certain details that are forthcoming, in which he claimed to employ "a Poissonization argument together with a strong approximation" [19, see the top of p. 2255], the latter from [28], to more carefully compute how many extreme points the jump process hits. Changing what seems to be a typo—see Remark 1.11—it seems from [19, p. 2255] like this forthcoming work counts

$$\frac{2^{1/3}k_1 \log n}{\big(g(x)|g'(x)|\big)^{1/3}} \text{ extreme points as the slopes vary through a range } g(x) \pm \frac{\log n}{n^{1/3}}; \quad (1.9)$$

Groeneboom is also able to give variances. A key question is how broadly Groeneboom can prove (1.9) for Poisson processes. This would have interesting applications in our problem. We will say more about (1.9) later; for now we just focus on Groeneboom's description of the argument, quoted above. Despite the one minor difference in perspective (Remark 1.9), this sounds quite similar to the technique giving asymptotic expectations in §1.3.4. For one, we already promised in Remark 1.4 that a similar rescaling will come below. Moreover, in §9 we use the Komlós, Major, Tusnády coupling [27]; we actually use a minor adjustment of this coupling (Lemma 9.13), which can be compared to more substantial results in [28] (though our processes are in some sense sideways).

Later, in §10, we will also convert informally between the two problems using a rescaling that Robin Pemantle suggested. It is possible that Groeneboom's Poisson technique is related to this interchange. *

Remark 1.11 (Explanation of (1.9) above)**.** Above, (1.9) is slightly different than what Groeneboom claimed. Since details of the proof are forthcoming, it is uncertain, however enough of the proof is available to see [19, p. 2255] likely has a typo. Groeneboom's claim at the top of that page includes another factor of $n^{1/3}$ in the number of extreme points the jump process touches. Temporarily using Groeneboom's notation, here are three distinct suggestions that this $n^{1/3}$ is a typo:

- Since Groeneboom can compare his processes U_n and $\widetilde{V}_n$ on intervals that differ in length by a factor of $n^{-1/3}$, the number of jumps of U_n ought to lack the extra $n^{1/3}$.

- On the first line computing EN_n (what we call $\mathbb{E}\mathcal{N}_n$) on p. 2255 of [19], the $n^{\pm 1/3}$ cancel, leaving a sum of just $2k_1 c_1^{-1} \log n$ jumps from each interval.
- In the same computation of EN_n, the number of intervals considered in the sum (there called K_n) is of order $\dfrac{n^{1/3}}{\log n}$. If we expected order $n^{1/3} \log n$ extreme points in every single interval, and if we had $c_1 \asymp 1$, we would expect order $\dfrac{n^{1/3}}{\log n} \cdot n^{1/3} \log n = n^{2/3}$ extreme points overall, rather than $n^{1/3}$.

These should explain why we removed the $n^{1/3}$ in (1.9). *

Additional comparisons will come later, since many are better to discuss after an informal calculation.

Informal Calculation

Remark 1.4 and the comment above it gave two reasons to do an informal calculation: to give a simple presentation of how to bound $\mathbb{E}\mathcal{E}_f(t)$ and to show similarities to (1.7/P.G.1). We now calculate a lower bound for $\mathbb{E}\mathcal{E}_f(t)$ by ignoring most of the details, but still following §1.3.4(b)–§1.3.4(c)(i) and Remark 1.5. Much of this work is redone later, but it helps to present the rough idea. Let us focus specifically on the right-hand-side walk in Figure 1.5. Remaining above a line of slope β_* is equivalent (see §6.2) to the random-walk event

$$\mathcal{S}_n \geq \lambda(X + \beta_* n) - \lambda(X) - n \qquad \text{for all } n > 0, \tag{1.10}$$

where $\mathcal{S}_n$ is a walk of i.i.d. $\text{Exp}(1) - 1$ steps. We want to know the chance (1.10) happens as $X \to \infty$. The most reasonable choice of slope is $\beta_* = 1/f(X)$, as we will see later in (6.2). Since this is an informal calculation, just replace the quantity on the right by its second-order Taylor expansion in n around $n = 0$, meaning consider instead the event

$$\mathcal{S}_n \geq \underbrace{\frac{1}{2}\beta_*^2 \cdot f'(X)n^2 = \frac{f'(X)}{2[f(X)]^2}n^2}_{\approx \lambda(X+\beta_* n) - \lambda(X) - n} \qquad \text{for all } n > 0. \tag{1.11}$$

The lower bound here is negative if $f'(X) < 0$. Now rescale by $1/\sqrt{N}$ and let $n = Nt$ where

$$\left.\begin{aligned} &N = \left(\frac{-f'(X)}{[f(X)]^2}\right)^{-2/3} = \begin{cases} \dfrac{X^{\frac{2}{3}p}}{q} & \text{if } f(x) = 1/x^q \\ X^{\frac{2}{3}} & \text{if } f(x) = 1/\log x \end{cases} \qquad \text{(so } N \to \infty \text{ as } X \to \infty) \\ &\text{so that (1.11) says} \quad \frac{\mathcal{S}_n}{\sqrt{N}} \geq \frac{1}{2} \cdot \frac{f'(X)}{[f(X)]^2} \cdot N^{3/2} t^2 = -\frac{1}{2}t^2 \quad \text{for } n, t > 0. \end{aligned}\right\} \tag{1.12}$$

Remark 1.5 gave context for the last two steps, as well as for the next one—the random walk boost. To clarify why this boost is necessary, notice the event $\frac{\mathcal{S}_n}{\sqrt{N}} \geq -\frac{1}{2}t^2$ in (1.12) limits to the event $B_t \geq -\frac{1}{2}t^2 \;\forall\, t > 0$. This event never occurs, because $\mathcal{S}_0 = 0$ and $B_0 = 0$ begin right on the parabola, and B_t varies too widely (compare the modulus of continuity [32, Theorems 1.13–1.14]). Rather than rescale to Brownian motion immediately at $n = t = 0$, first let the walk move up to a height $> \sqrt{N}$, because that distance will not disappear when we apply Donsker's theorem:

$$\begin{aligned} &\mathbb{P}\left(\mathcal{S}_n \geq \frac{1}{2}\beta_*^2 \cdot f'(X)n^2\right) \\ &\qquad \geq \mathbb{P}\left(\mathcal{S}_n \geq 0 \text{ for } 0 < n < N \text{ and } \mathcal{S}_N > \sqrt{N}\right) \qquad \text{(teal path in Figure 1.10)} \\ &\qquad\quad \cdot \mathbb{P}\left(\mathcal{S}_n \geq \frac{1}{2}\beta_*^2 \cdot f'(X)n^2 \text{ for } n > N \;\middle|\; \mathcal{S}_N = \sqrt{N}\right) \qquad \text{(red path in Figure 1.10)} \end{aligned} \tag{1.13}$$

The first random walk probability is at least $\Omega\left(1/\sqrt{N}\right)$ by Lemma 2.6 below (which is essentially a random walk result in Feller [13] combined with a well-known correlation inequality—

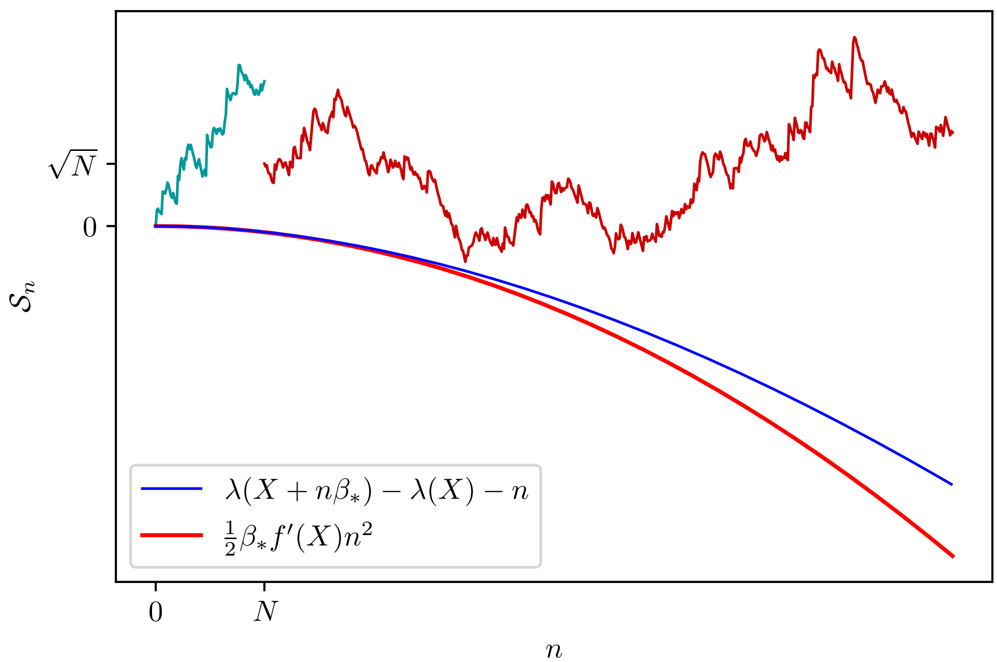

Figure 1.10: The informal probability calculation.

---

see the references below). For the final probability,

$$\mathbb{P}\left(\mathcal{S}_n \geq \frac{1}{2}\beta_*^2 \cdot f'(X)n^2 \text{ for } n > N \;\middle|\; \mathcal{S}_N = \sqrt{N}\right) \tag{1.14}$$

$$\to \mathbb{P}\left(B_t + 1 \geq -\frac{1}{2}t^2 \text{ for } t > 1 \;\middle|\; B_1 = 0\right) \qquad \text{(Donsker's Theorem)}$$

$$= \mathbb{P}\left(B_t - \frac{1}{2}t^2 \leq 1 \text{ for } t > 1 \;\middle|\; B_1 = 0\right) \qquad \text{(reflection, } B_t \overset{\mathrm{d}}{=} -B_t\text{)}$$

$$\geq \mathbb{P}\left(\max_{t>0}\left(B_t - \frac{1}{2}t^2\right) \leq 1 \;\middle|\; B_0 = 0\right) \qquad \text{(shift time back, require all } t > 0\text{).}$$

The last probability is constant-order, since $\max_t \left(B_t - \frac{1}{2}t^2\right)$ has a positive density above zero, as one can read off pictures given by Groeneboom [18, see Corollary 2.1 and Figure 2] or Janson, Louchard, and Martin-Löf [25, see Theorem 2.4 and Figure 2]. This suggests (1.10) happens with probability at least order $1/\sqrt{N}$. There is also the left-hand walk on the other side of Figure 1.5. It is about the same, so we expect by independence

$$\mathbb{P}(X \text{ is extreme} \mid X) \gtrapprox \Omega\left[\left(\frac{1}{\sqrt{N}}\right)^2\right] = \Omega\left[\left(\frac{-f'(X)}{[f(X)]^2}\right)^{2/3}\right]$$

as $X \to \infty$. Using this within (1.6), we see $\mathbb{E}\mathcal{E}_f(t)$ ought to be at least order

$$\int_0^t \left(\frac{(f'(x))^2}{f(x)}\right)^{1/3} dx \tag{1.15}$$

as $t \to \infty$, and with that we record:

Remark 1.12 (Recovery of integral, apart from constants)**.** The informal calculation above loosely recovers the integral I_g in (1.7/P.G.1) found by Groeneboom [19]. *

Remark 1.13 (Approximations vs. inequalities)**.** Tracing through the inequalities and glancing at Figure 1.10 shows replacing (1.10) by the approximation (1.11) was not okay: the parabolic approximation is *below* the walk's true boundary, so this calculation gave an unjustified lower bound. Remark 1.5 already mentioned the fix, which is done in §8. *

Remark 1.14 (Comments on Low Intensities)**.** With intensities $1/x$ or $1/x \log^q x$, the rescaling constant N in (1.12) is $N = O(1)$, so the rescaling above does not apply. However, (1.15) is still somewhat accurate with $1/x$. *

More Comparisons

Here we explain a few more relationships between the arguments in this paper and others. We have nearly all the context we need to discuss these now, but a few details from later will be used. It will help to have in mind the following result by Prakasa Rao:

Theorem 1.15 (from Prakasa Rao [34, Lemma 4.1])**.** Keep the setup above,[4] like in Figure 1.9, with the number of samples in the empirical distribution G_n growing, $n \to \infty$.

[4]We should say the hypotheses in [34, Lemma 4.1] are more general than those stated in Theorem 1.7 from Groeneboom. However, a loose understanding of this result is enough here.

Consider an interval of width Δx around any fixed point x. If $\Delta x \gg n^{-1/3}$, then with probability $1 - o(1)$, the slope of the concave majorant of G_n at x is determined by the behavior of G_n solely on the small interval $\left[x - \frac{1}{2}\Delta x, x + \frac{1}{2}\Delta x\right]$.

Remark 1.16 (Relevant Ranges of Slopes)**.** Theorem 1.15 is useful to explain why relevant rescalings and ranges of slopes are the same between our results and Groeneboom's. When looking for jumps, Groeneboom considers varying a prescribed slope by $\log n/n^{1/3}$ (see (1.9) above from [19, top of p. 2255]). Likewise, when checking if a particular point X_n from the $1/x^q$ process is extreme, yes we will need to check a list of slope zones as in (1.4), but in the end the most important slopes will vary on the order[5] $\log X_n/X_n^{p/3} \approx \log n/n^{1/3}$, and the rest of the slope zones will be unimportant (Lemma 9.3). Here is one explanation of this common $\log n/n^{1/3}$ range of slopes and why other slopes are unnecessary. As a technical assumption, Groeneboom requires that $g'(x)$ remain bounded and away from zero in [19, Lemma 3.1, Theorem 3.1], hence any interval of asymptotic width Δx around a point x corresponds to order-Δx variations in slopes around $g(x)$. In other words, the variations of slopes—not just the variations Δx of x values—should satisfy the $\Delta x \gg n^{-1/3}$ scaling in Theorem 1.15 from Prakasa Rao, and in that sense, $\log n/n^{1/3} \gg n^{-1/3}$ is just enough.

There is also a relation between Prakasa Rao's proof of Theorem 1.15 and how we check unimportant slope zones below Lemma 9.3. *

While Remark 1.16 emphasized that only a narrow range of *slopes* are key, there is a closely related observation to make about *points*.

Remark 1.17 (More on Relevant Ranges)**.** Sometimes we consider whether a point is extreme by looking only within a relatively small neighborhood of the point. This kind of

[5]This is (9.5) and the loose approximation $X_n = \lambda^{-1}(S_n) \approx \lambda^{-1}(n) \approx n^{1/p}$.

local focus has of course been used before. One perfect example is Theorem 1.15 above from Prakasa Rao. Another example discussed soon is a paper by Groeneboom [16] in stochastic geometry.[6] To clarify exactly what this means in our context, we outline two concrete examples now.

The first example we already saw in the informal calculation, above. More fully, the arguments in §8–§9 to determine bounds on $\mathbb{P}(X \text{ is extreme} \mid X)$ will essentially only look at the nearest $\approx N$ Poisson points—where N is order $X^{\frac{2}{3}p}$ or $X^{2/3}$, as in (1.12) above. The distant portion of the walk (called Stage III in §8) will barely affect the probability. There are already two ways we can see order N steps are enough.

- With either intensity, a range of N points around X_n corresponds loosely to the surrounding $\approx n^{2/3}$ points. This is the same number of points needed in the context of Theorem 1.15 from Prakasa Rao: out of n samples total, order $\Delta x \cdot n \gg n^{-1/3} \cdot n = n^{2/3}$ land in an interval of width $\Delta x \gg n^{-1/3}$ near x.
- Another way to understand the range of N steps is a scaling relation for the random time $t_{\max} = \operatorname{argmax}_t B_t - ct^2$. This relation is described extremely concisely in [24, equations (2.1)–(2.2)] or [19, Remark 2.5], for instance, and says $t_{\max}$ scales like $c^{-2/3}$. To use this, recall the technique in (1.12) chose the number of steps N to cancel the parabolic constant c, meaning $N^{3/2} \propto 1/c$. The scaling relation then says the time of the maximum scales like order $(N^{3/2})^{2/3} = N$.

To discuss the second example, we need to share an idea used in [19] and [16]: in both of these papers, Groeneboom works with processes jumping among extreme points and shows

[6] After a brief look at a survey article, [2], one gets the impression this same idea—looking locally or using long-range independence—crops up commonly in stochastic geometry. This broader direction was not explored, however it is worthwhile to say something more: Groeneboom saw parallels in [19, see pp. 2237–2238 (bottom to top) and p. 2257] between 1/3 exponents or logarithmic rates in concave majorant results and the stochastic geometry results he dealt with earlier in [16].

the distant past of these processes is nearly independent of the distant future, which enables him to use stationary sequence limit theorems. Intuitively, this says extreme points do not have long-lasting effects. With that said, we can now preview a related argument: to prove the extreme point count with intensity $1/x$ satisfies a limit (introduced soon, see (1.16) in §1.3.6) we will use a stationary sequence idea like in [19], [16]. The goal to do this is to show whether X_j is extreme has very limited influence on whether X_ℓ is extreme if the indices j, ℓ are distant. In other words, to decide if X_ℓ is extreme, a key is to show we only need to see a few indices around ℓ. (See Lemma 7.8, (7.8) and below.) *

1.3.6 Technique 5. A Limit Theorem, Plus Other Conjectures Based on Groeneboom's Work

With intensity $1/x$, among $X_1, \ldots, X_n$ we expect to see order n extreme points. This is noticeably different from the rate of $n^{1/3}$ discussed in §1.3.5 above. Nonetheless we can still use a method similar to proofs given by Groeneboom [19], namely to apply a central limit theorem for a stationary sequence, in order to learn there is a constant μ with which

$$\frac{\mathcal{E}_{1/x}(X_n) - \mu n}{\sqrt{n}} \tag{1.16}$$

either converges in probability to zero or converges in distribution to some $\mathrm{N}(0, \sigma^2)$. This is worked out in §7.

For other intensities, we cannot say much that is definite, however we can invoke Groeneboom's work from [19, §3] at least to guess what happens. The basic idea is to rescale a Poisson process of intensity $f(x)$ in a way Robin Pemantle suggested so that roughly n points of the squished Poisson process compare to n samples from a probability density $g(x)$ on $[0, 1]$. This is done in §10. This does not precisely translate between our Poisson context and the context in Groeneboom's paper [19, §3], but it makes it reasonable to think Groeneboom's result (1.8/P.G.2) might translate back into these tentative conjectures: with coefficients from (1.7/P.G.1) above,

$$\text{(a)}\quad \frac{\mathcal{E}_{1/x^q}(t) - k_1 \cdot \dfrac{3q^{2/3}}{2^{2/3}p} t^{p/3}}{\sqrt{k_2 \cdot \dfrac{3q^{2/3}}{2^{2/3}p} t^{p/6}}} \overset{d}{\Rightarrow} \mathrm{N}(0,1) \text{ and}$$

$$\text{(b)}\quad \frac{\mathcal{E}_{1/\log x}(t) - k_1 \cdot \dfrac{3}{4^{1/3}} \cdot \dfrac{t^{1/3}}{\log t}}{\sqrt{k_2 \cdot \dfrac{3}{4^{1/3}} \cdot \dfrac{t^{1/3}}{\log t}}} \overset{d}{\Rightarrow} \mathrm{N}(0,1).$$

It may be that Groeneboom would be able to prove these with the argument omitted from [19, see p. 2255], but we cannot say without knowing more (see Remark 1.10).

Remark 1.18. When trying to translate between problems, an immediate distinction between our problem and Theorem 1.7 is that Groeneboom's work in [19, §3] concerns a finite sample—that is, there are only n points in Figure 1.8, then the picture stops—whereas each Poisson process above has a tail of infinitely many points. If we had not already seen Tables 1.1–1.2, we might ask a few questions about this infinite tail. Are many of these tail points extreme? Do these tail points grow slowly enough to make many of the earlier points nonextreme? It is possible to argue, at least for $1/x^q$ (using Theorem 1.15 from Prakasa Rao and (1.9) from Groeneboom to see these are local), that these tail points really do not matter. The details are not interesting so are omitted. *

There is something interesting about the guess (b). Once translated, $1/\log x$ does not formally fit Groeneboom's context to apply (1.8/P.G.2). More precisely, when we rescale as in §10.1, this intensity turns not into a fixed density $g(x)$ but rather into a sequence of densities $g_n(x)$ that change with n. In fact, whenever $0 < x < 1$ is fixed as $n \to \infty$, (10.4) will check

$$g_n(x) \sim 1 \qquad \text{and} \qquad g_n'(x) \sim 0,$$

showing in a sense $g_n(x)$ limits to the uniform density away from $x = 0$. (This is not true if $x \to 0$ with n, where there is a spike and $g_n(x) \approx \log n$.) This limit is interesting because we know $1/\log x$ has $\mathbb{E}\mathcal{E}_f(t) = \widetilde{\Theta}(t^{1/3})$, corresponding to a guess that $\approx n^{1/3}$ points are extreme among $X_1, \ldots, X_n$, but on the other hand the following is known:

Remark 1.19 (Uniform case)**.** Groeneboom pointed out in [19] that when $g(x) = 1$ is uniform on $[0, 1]$, the answer is no longer (1.7/P.G.1) or (1.8/P.G.2). In this case, the number of extreme points $\mathcal{N}_n$ on $G_n(x)$ has a normal limit with mean and variance of only $\log n$; a brief proof of this fact is in the paper [20], however the result is attributed originally to Sparre Andersen. Sparre Andersen [35, see §7–8] managed to find the exact distribution of the number of extreme points in a finite sequence of sums $S_1, \ldots, S_n$ of exchangeable continuous steps. If the steps are Exp(1), this would be relevant to a constant-intensity Poisson process (again, apart from this being a finite sequence of points). This case will not be considered here, but see also Groeneboom's paper [14]. *

Remark 1.20 (Compatibility with point-by-point probabilities)**.** At least informally, the probabilities $\mathbb{P}(X \text{ is extreme} \mid X)$ for $1/\log x$ and $1/x^q$ in Table 1.2 seem compatible with this rescaling and with Groeneboom's claim (1.9). To see this, recall the loose substitutions $X_n \approx (pn)^{1/p}$ with $1/x^q$ or $X_n \approx n \log n$ with $1/\log x$ (see (3.3) below), and substitute these into the probabilities from Table 1.2; then for each $k \asymp n$, the chance the kth largest point is extreme should be around order $1/n^{2/3}$, ignoring logarithms. Now consider what Groeneboom can show in (1.9), the fact from [19, p. 2255] with the correction in Remark 1.11.

If slopes vary through $g(x) \pm \frac{\log n}{n^{1/3}}$, then x varies on the order $\Delta x \asymp \frac{\log n / n^{1/3}}{|g'(x)|}$. In an interval of this width, there are order $n\Delta x$ points total, out of which (1.9) tells us the fraction that are extreme:

$$\asymp \frac{\log n \cdot \Big(g(x)|g'(x)|\Big)^{-1/3}}{n\Delta x} \asymp \frac{|g'(x)|^{2/3}}{n^{2/3} \cdot [g(x)]^{1/3}}.$$

Evaluated the rescaled density $g(x) = p/x^q$ from $f(x) = 1/x^q$ and (10.2), this quantity is order $1/n^{2/3}$, agreeing with above. We already said that $1/\log x$ does not rescale to a single density, but when evaluated along the sequence g_n, the fraction above is still order $1/(n \log n)^{2/3}$. *

1.4 Remaining Questions

Recall that Theorem 1.7 from Groeneboom assumed the density $g(x)$ had a finite support. Groeneboom conjectured this finite support assumption may be unnecessary if in (1.7/P.G.1) the integral I_g is replaced by an integral over $[0, \infty)$, and he simulated with an exponential distribution to show the $n^{1/3}$ mean and variance results are still believable [19, pp. 2255–2256, below the proof of Lemma 3.1]. The exponential distribution, however, is very concentrated near zero. One may ask in light of the similarities, and in light of our order-n answers for Poisson intensities $1/x$ and $1/x \log^q x$ $(q > 1/2)$ where points spread out quickly:

Question 1.21. Do we ever encounter the rate n in Groeneboom's problem? What if points sampled from $g(x)$ are well spread in a long tail, too?

One may want to look at a density like $g_r(x) = \dfrac{r}{x \log^{1+r} x}$ (some $0 < r \leq 2$) on $x \geq e$ which has CDF $G_r(x) = 1 - \dfrac{1}{\log^r x}$. The first and more interesting reason is that, with this density, an integral like I_{g_r} from (1.7/P.G.1) but integrated over $[e,\infty)$ is infinite if $r \le 2$:

$$\int_e^\infty \left(\frac{(g_r'(x))^2}{g_r(x)}\right)^{1/3} dx = \int_e^\infty \frac{r^{1/3}\big(\log^{1+r} x + (1+r)\log^r x\big)^{2/3}}{x\log^{1+r} x}\, dx$$
$$> \int_e^\infty \frac{r^{1/3}}{x(\log x)^{(1+r)/3}}\, dx = \infty.$$

A second reason is if $\widetilde{U}_1,\ldots,\widetilde{U}_n \overset{\text{iid}}{\sim} \text{Unif}(0,1)$ and $U_1 < \cdots < U_n$ are their order statistics, then samples $X_j \overset{\text{def}}{=} G_r^{-1}(U_j) = \exp\left(\frac{1}{(1-U_j)^{1/r}}\right)$ grow very quickly, like the Poisson intensities above. Small computer simulations were not so enlightening about what should happen with g_r.

Other unresolved questions relate to the Borel-Cantelli and coupling strategies.

Question 1.22. What answer goes in Table 1.1 for $1/x\log^q x$ when $0 < q \le 1/2$?

Question 1.23. Is there a cutoff intensity $f_*(x)$ below which there are finitely many nonextreme points and above which there are infinitely many nonextreme points? That is, are there finitely many nonextreme points whenever $f(x) \ll f_*(x)$ and infinitely many nonextreme points whenever $f(x) \gtrsim f_*(x)$? Two reasonable guesses seem to be $1/x$ and $1/x\sqrt{\log x}$, and maybe the latter is more believable.

Question 1.24. Can a relative counting property be proven for intensities besides $1/x^q$?

1.5 Notation

Whatever $f(x)$ we consider, we denote the mean number of points in $[0,t]$ or $A \subset [0,\infty)$ by

$$\lambda(t) = \int_0^t f(x)\,dx \qquad \text{and} \qquad \mu(A) = \int_A f(x)\,dx. \tag{1.17}$$

The notation $\mathbf{1}[\cdots]$ is an indicator function.

Many arguments consider walks with exponential steps. The exponential distribution of mean 1 will be written Exp(1), while the same distribution shifted down by a constant M will be written $\text{Exp}(1) - M$. To clarify, the associated distribution function is

$$F(x) = \begin{cases} 1 - e^{-(x+M)} & \text{if } x \geq -M \\ 0 & \text{otherwise.} \end{cases}$$

Nearly everywhere, we will reserve

- S_n for a random walk with i.i.d. Exp(1) steps,
- $\mathcal{S}_n$ for a random walk with i.i.d. $\text{Exp}(1) - 1$ steps, and
- B_t for a Brownian motion.

LHS and RHS abbreviate left-hand side and right-hand side, and in superscripts and subscripts, L and R (such as in $\mathcal{S}_n^R$) denote the same. To emphasize both sides simultaneously, sometimes we write $*$ instead (as in $\mathcal{S}_n^*$, where $* = L, R$).

Most commonly, asymptotic statements will be as $X \to \infty$ (where X denotes a point of the process) or as $t \to \infty$. Besides the well-known big-O notations (see, for instance, [36] or [9]) we also use a tilde to indicate omitted logarithms. For convenience, Table 1.3 summarizes what all of these mean—both the notations with tildes and the notations that are standard (though inequalities might not be in style). A tilde may hide either a $\log X$ or a $\log t$, depending on the context.

Note 1.25. To clarify, $f = \widetilde{\Theta}(g)$ does *not* guarantee $f = \Theta[g(X) \cdot \log^r X]$. Said another way, $\widetilde{\Theta}$ allows the logarithm exponents to differ (in the table, $r > s$).

Note 1.26. The notation $\approx$ means nothing except intuition.

	Notation	Meaning
Common Notations	$f \sim g$	$f(X)/g(X) \to 1$
	$f \leq o(g)$ or $f \ll g$	$f(X)/g(X) \to 0$
	$f \leq O(g)$ or $f \lesssim g$	$\limsup f(X)/g(X) < \infty$
	$f \geq \Omega(g)$ or $f \gtrsim g$	$\liminf f(X)/g(X) > 0$
	$f = \Theta(g)$ or $f \asymp g$	$\Omega(g) \leq f \leq O(g)$
With Logs	$f \leq \widetilde{O}(g)$	$f(x) = O[g(X) \cdot \log^r X]$ for some $r \in \mathbb{R}$
	$f \geq \widetilde{\Omega}(g)$	$f(x) = \Omega[g(X) \cdot \log^s X]$ for some $s \in \mathbb{R}$
	$f = \widetilde{\Theta}(g)$	$\widetilde{\Omega}(g) \leq f \leq \widetilde{O}(g)$

Table 1.3: Clarification of asymptotic notations used here.

Chapter 2

Setup

This section is a catalog of facts for use later. Before continuing, it would be useful to read the statement of Lemma 2.1, which has a simple way to build Poisson processes. The rest can be read as necessary, if desired.

Lemma 2.1. Let $Z_1, Z_2, \ldots \overset{\text{iid}}{\sim} \mathrm{Exp}(1)$ with partial sums $S_n = \sum_{j=1}^{n} Z_j$. One can generate a Poisson process of intensity $f(x)$ by setting $X_n = \lambda^{-1}(S_n)$, where $\lambda(t)$ is defined in (1.17).

For instance, a well-known property of a rate-1 Poisson process is that the interarrival times $X_{j+1} - X_j$ are already Exp(1), and in this case we see $X_n = S_n$. To find this property in general, argue as follows.

Sketch for Lemma 2.1. If we condition on $X_j = x$ (or if we begin with a nonexistent point $X_0 = 0$), then the distribution function for the interarrival time $X_{j+1} - X_j$ is

$$y \mapsto \mathbb{P}(X_{j+1} - x < y \mid X_j = x) = 1 - e^{-\mu(x, x+y)}.$$

Inverting this function and evaluating with a Unif(0, 1) shows a useful property: during the interarrival waiting times, the process must accumulate an Exp(1) number of points *in* *expectation*, or in other words $(\mu(X_j, X_{j+1}) \mid X_j) \sim \text{Exp}(1)$. Using this fact and adding up the waiting times concludes the argument. □

Remark 2.2. The way interarrival times are generated in the sketch above is essentially the same way jump times are generated by Groeneboom in [19, pp. 2249–2250]. *

We need a few estimates that will be used repeatedly. The first may be found in [12, Ch. VII.1, Lemma 2] or [10, Theorem 1.2.6]:

$$\int_t^\infty \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx \le \frac{1}{\sqrt{2\pi} \cdot t} e^{-t^2/2} \qquad (t > 0). \tag{2.1}$$

The second estimate we will use is a large deviations result that may be found in [13, Ch. XVI.7 Theorem 1], specialized here to our exponential-step case: if $\mathcal{S}_n$ is a sum of n i.i.d. $\text{Exp}(1) - 1$ random variables, and if $0 < \delta < 1/6$, then as $n \to \infty$,

$$\mathbb{P}\left(\frac{\mathcal{S}_n}{\sqrt{n}} > n^\delta\right) \sim \int_{n^\delta}^\infty \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx \le \frac{1}{\sqrt{2\pi} n^\delta} e^{-n^{2\delta}/2} \tag{2.2}$$

(simply concatenating (2.1) at the end). As Robin Pemantle pointed out to me, summing (2.2) over $n \in \mathbb{N}$ and applying Borel-Cantelli quickly gives a useful estimate:

Lemma 2.3. Select $0 < \delta < 1/6$. If S_n and X_n are as in Lemma 2.1, then $S_n = n + o\left(n^{\frac{1}{2}+\delta}\right)$ and $X_n = \lambda^{-1}\left(n + o\left(n^{\frac{1}{2}+\delta}\right)\right)$ as $n \to \infty$.

Applying exactly the same reasoning in the previous paragraph but now to the number of points, rather than their locations, gives:

Lemma 2.4. Select $0 < \delta < 1/6$. The number of Poisson points in $[0, t]$ satisfies $N(t) \sim \lambda(t) + o\left([\lambda(t)]^{\frac{1}{2}+\delta}\right)$ as $t \to \infty$.

We will need several exponential random walk results. This first is available in [13]:

Lemma 2.5. Let $\widetilde{\mathcal{S}}_n$ be a walk started at $\widetilde{\mathcal{S}}_0 = 0$ and taking i.i.d. $\mathrm{Exp}(1) - M$ steps, where $0 < M < 1$. Then $\mathbb{P}(\widetilde{\mathcal{S}}_n > 0 \; \forall \, n > 0) = 1 - M$.

Proof. This probability is derived in [13], Ch. XII.4 Example (a), equation (4.5) on p. 405, where our answer in notation there is $1 - \rho(0) = \alpha\mu$ with $\alpha = 1$ and $\mu = 1 - M$. □

Lemma 2.6. Let $\mathcal{S}_n$ be a walk from $\mathcal{S}_0 = 0$ taking i.i.d. $\mathrm{Exp}(1) - 1$ steps. As $N \to \infty$,

$$\mathbb{P}\Big(\mathcal{S}_n \geq 0 \text{ when } 0 \leq n \leq N \text{ and } \mathcal{S}_N \geq \sqrt{N}\Big) \geq \Omega\bigg(\frac{1}{\sqrt{N}}\bigg).$$

The same is true if $\mathcal{S}_n$ instead takes (reflected) i.i.d. $1 - \mathrm{Exp}(1)$ steps.

Proof. By the central limit theorem, $\mathbb{P}\Big(\mathcal{S}_N \geq \sqrt{N}\Big) = \Theta(1)$. Separately,

$$\mathbb{P}(\mathcal{S}_n \geq 0 \text{ when } 0 \leq n \leq N) \geq \Omega\bigg(\frac{1}{\sqrt{N}}\bigg)$$

by referring to [13], pp. 414–415, Ch. XII.7 Theorem 1a and the preceding discussion. (To verify the required condition, that $\sum_n \big(\mathbb{P}(\mathcal{S}_n > 0) - \frac{1}{2}\big)/n$ converges, Feller points to another result; however, since $\mathrm{Exp}(1) - 1$ steps have a third moment, it is easy to verify convergence using the Berry-Esseen theorem, see [13], p. 542, Ch. XVI.5 Theorem 1.) Robin Pemantle observed both of the events above are increasing, so we can immediately multiply the two probabilities using a correlation inequality quoted below in Lemma 2.7. □

The next result quoted is a continuous-variable version of a well-known correlation inequality of Harris [22]. The continuous version can be found in [21, Theorem 1] or [26, eq. (1.10), Theorems 2.1, 2.3], which we specialize to events of Exp(1) variables below:

Lemma 2.7. Suppose $Z_1, \ldots, Z_n \overset{\text{iid}}{\sim} \mathrm{Exp}(1)$ are defined as the coordinates of $\Omega = \mathbb{R}^n$. If $A, B \subset \Omega$ are increasing events, meaning

$$x_1 \geq 0, \ldots, x_n \geq 0 \text{ and } (z_1, \ldots, z_n) \in A \quad \Rightarrow \quad (z_1 + x_1, \ldots, z_n + x_n) \in A$$

and likewise for B, then $\mathbb{P}(A \cap B) \geq \mathbb{P}(A) \cdot \mathbb{P}(B)$.

Lemma 2.8. Suppose $\mathcal{S}_n$ satisfies Donsker's theorem, meaning $\mathcal{S}_{Nt}/\sqrt{N} \overset{d}{\Rightarrow} B_t$ as $N \to \infty$. If $a \to \infty$ and $ab \to \infty$, then $\mathbb{P}(\mathcal{S}_n \geq -a - bn \ \ \forall n > 0) \to 1$.

Proof. Let $N = a^2$ with $n = Nt$, and suppose $ab > C$ for some constant C. Then

$$\begin{aligned}
\mathbb{P}(\mathcal{S}_n \geq -a - bn \ \ \forall n \geq 0) &= \mathbb{P}\left(\frac{\mathcal{S}_n}{\sqrt{N}} \geq \frac{-a - bNt}{\sqrt{N}} \ \ \forall t = \frac{n}{N} > 0 \right) \\
&= \mathbb{P}\left(\frac{\mathcal{S}_n}{\sqrt{N}} \geq -1 - abt \ \ \forall t = \frac{n}{N} > 0 \right) \\
&\geq \mathbb{P}\left(\frac{\mathcal{S}_n}{\sqrt{N}} \geq -1 - Ct \ \ \forall t = \frac{n}{N} > 0 \right).
\end{aligned}$$

Applying Donsker's theorem as $a \to \infty$, then using a well-known calculation about Brownian motion—see [10, Exercise 7.5.2]—this lower bound becomes

$$\mathbb{P}(\mathcal{S}_n \geq -a - bn \ \ \forall n \geq 0) \gtrsim \mathbb{P}(B_t \geq -1 - Ct \ \ \forall t > 0) = 1 - e^{-2C}.$$

At last send $C \to \infty$. □

Chapter 3

Borel-Cantelli Slope Observations

Results for general intensities $f(x)$ will be stated and proved in §3.1, then these will be applied to one example at a time in §3.2. Once again, I appreciate help from Da Wu and Kaitian Jin during an initial discussion of this problem, when we looked together at particular examples and used a prototype for the proof below of (3.1).

3.1 General Observations

Let L_j denote the line segment connecting (j, X_j) to $(j+1, X_{j+1})$, and let $\text{slope}(L_j) = X_{j+1} - X_j$ denote its slope. In §1.3.1 and Figure 1.3 we made the following observation: if eventually $\text{slope}(L_j) < \text{slope}(L_{j+1}) < \text{slope}(L_{j+2}) < \cdots$, then only finitely many points are nonextreme. We now use Borel-Cantelli with this and some similar observations.

Proposition 3.1. The following events are equal almost everywhere:

$$\left\{ \sum_{j=1}^{\infty} \int_{X_{j+1}}^{2X_{j+1}-X_j} f(x)\,dx < \infty \right\} \stackrel{\text{a.s.}}{=} \left\{ \text{slope}(L_j) \geq \text{slope}(L_{j+1}) \text{ holds only for finitely many } j \right\}. \tag{3.1}$$

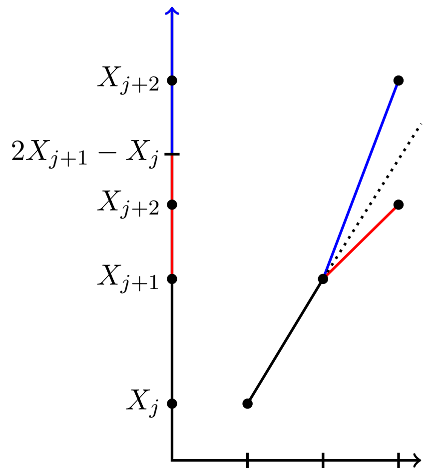

Figure 3.1: The key distance for a slope increase or decrease.

When this event occurs, all but finitely many points are extreme. When it does not occur, there are infinitely many nonextreme points. The event (3.1) also satisfies a zero-one law.

Proof. The event that the slopes decrease around X_{j+1} is the event that X_{j+2} lands too close to X_{j+1}; more precisely it is the event that $X_{j+2} \leq 2X_{j+1} - X_j$ (see Figure 3.1). By a conditional version of Borel-Cantelli (see [10, Theorem 4.3.4]), the event that there are finitely many slope decreases is almost surely the event that

$$\sum_{j=1}^{\infty} \mathbb{P}\Big(X_{j+2} \leq 2X_{j+1} - X_j \;\Big|\; X_j, X_{j+1}\Big) < \infty.$$

The summands here are

$$\mathbb{P}\Big(X_{j+2} \leq 2X_{j+1} - X_j \;\Big|\; X_j, X_{j+1}\Big) = 1 - e^{-\mu(X_{j+1},\ 2X_{j+1} - X_j)}.$$

For the sum of these terms to converge, the exponent must go to zero, and if so, using $\lim_{\mu \to 0} \frac{1 - e^{-\mu}}{\mu} = 1$ shows the sum converges if and only if

$$\sum_{j=1}^{\infty} \mu(X_{j+1},\ 2X_{j+1} - X_j) = \sum_{j=1}^{\infty} \int_{X_{j+1}}^{2X_{j+1} - X_j} f(x)\, dx$$

converges.

For the zero-one law, write $X_n = \lambda^{-1}(S_n)$ as in Lemma 2.1, with S_n the partial sum of $Z_1, Z_2, \ldots \overset{\text{iid}}{\sim} \text{Exp}(1)$. A finite permutation of $Z_1, \ldots, Z_n$ can only change $X_1, \ldots, X_{n-1}$, and so the permutation cannot affect convergence of the sum of integrals. This means the Hewitt-Savage 0–1 law (see [10, Theorem 2.5.4]) applies to the left-hand event in (3.1). $\square$

Because we consider decreasing intensities $f(x)$, simple bounds on the integral in (3.1) are usually convenient: $(X_{j+1} - X_j) f(2X_{j+1} - X_j) \leq \int_{X_{j+1}}^{2X_{j+1}-X_j} f(x)\,dx \leq (X_{j+1} - X_j) f(X_{j+1})$.

The next few results present related techniques that are often useful to find extreme or nonextreme points. These turn out to be exactly analogous to some results and arguments about the primes by Tutaj [38] or Pomerance [33]. The first technique is whether $\operatorname{slope}(L_j)$ generally grows but is also, however infrequently, relatively low. This is roughly the same observation as in Tutaj [38, Proposition 2].

Proposition 3.2. Let $0 < M < \infty$ and assume f is decreasing. Suppose the following two events both occur almost surely: $\sum_{j=1}^{\infty} f(X_j + M) = \infty$ and $X_j \gg j$. Then

$$\liminf_{j\to\infty} \operatorname{slope}(L_j) \leq M < \infty = \limsup_{j\to\infty} \operatorname{slope}(L_j),$$

and consequently there are infinitely many slope decreases.

This will be broken into two lemmas.

Lemma 3.3. With $0 < M < \infty$, these events are almost everywhere equal:

$$\left\{ \sum_{j=1}^{\infty} \int_{X_j}^{X_j+M} f(x)\,dx = \infty \right\} \overset{\text{a.s.}}{=} \Big\{ \operatorname{slope}(L_j) < M \text{ holds for infinitely many } j \Big\}. \tag{3.2}$$

Proof. Argue as for (3.1). First calculate

$$\sum_{j=1}^{\infty} \mathbb{P}\Big(\operatorname{slope}(L_j) < M \;\Big|\; X_j\Big) = \sum_{j=1}^{\infty} \mathbb{P}\Big(X_{j+1} - X_j < M \;\Big|\; X_j\Big) = \sum_{j=1}^{\infty} \Big(1 - e^{-\mu(X_j,\ X_j+M)}\Big).$$

The last sum converges if and only if

$$\sum_{j=1}^{\infty} \mu(X_j,\ X_j + M) = \sum_{j=1}^{\infty} \int_{X_j}^{X_j+M} f(x)\,dx$$

does. $\square$

The proof of the next lemma can be compared to Tutaj [38, p. 131].

Lemma 3.4. If $X_j \gg j$ almost surely, then $\limsup_{j\to\infty} \operatorname{slope}(L_j) = \infty$.

Proof. If instead $\ell = \limsup_{j\to\infty} \operatorname{slope}(L_j) < \infty$, wait long enough that $\operatorname{slope}(L_j) < \ell + 1$ when $j \geq W$. Then with $J > W$,

$$\frac{X_J - X_1}{J} = \frac{\sum_{j=1}^{J} X_{j+1} - X_j}{J} < \frac{\sum_{j=1}^{W-1} \operatorname{slope}(L_j) + \sum_{j=W}^{J} \ell + 1}{J}.$$

As $J \to \infty$, the final expression limits to $\ell + 1 < \infty$, meaning X_j/j must be bounded. $\square$

Proof of Proposition 3.2. Recall we assume f is decreasing. This and $\sum_{j=1}^{\infty} f(X_j + M) = \infty$ implies that (3.2) holds, so $\liminf_{j\to\infty} \operatorname{slope}(L_j) \leq M$. Lemma 3.4 is the rest. $\square$

Instead of looking for slope decreases and nonextreme points, we can also look for extreme points as follows. The next result and proof very closely parallel the result and proof for the prime case given in Tutaj [38, Proposition 1] and Pomerance [33, Theorem 2.1].

Proposition 3.5. If $X_j \gg j$ almost surely, then there are almost surely infinitely many extreme points.

Proof. If in some realization X_r is an extreme point, and if there *does* exists another extreme point afterwards, call it X_s $(s > r)$, then we would know

$$\frac{X_s - X_r}{s - r} = \min_{j>r} \left\{ \frac{X_j - X_r}{j - r} \right\}.$$

On the other hand, if X_r happens to be the last extreme point, and if no extreme X_s exists later, then necessarily the minimum does not exist; rather

$$\inf_{j>r}\left\{\frac{X_j - X_r}{j-r}\right\} = \liminf_{j\to\infty}\left\{\frac{X_j - X_r}{j-r}\right\} < \infty,$$

and the convex hull ends with a line of finite slope. By assumption this does not happen: for any fixed r, $\lim_{j\to\infty}\dfrac{X_j - X_r}{j-r} = \infty$. This shows a later extreme point X_s exists. $\square$

3.2 Examples

This section justifies Table 1.1. Throughout, $X_n = \lambda^{-1}(S_n)$ as in Lemma 2.1.

Example 3.6 ($\mathbf{1}[x > e]/x \log x$)**.** We have $\lambda(t) = \log\log t$ so that $X_n = \exp(\exp(S_n))$. By the law of large numbers, there is an N for which $S_n > \frac{1}{2}n$ once $n > N$. From here we see (3.1) almost surely holds, because once $j + 1 > N$,

$$\int_{X_{j+1}}^{2X_{j+1}-X_j} f\,dx \le (X_{j+1} - X_j)f(X_{j+1}) < \frac{1}{\log(X_{j+1})} = \frac{1}{e^{S_{j+1}}} < e^{-\frac{1}{2}j}.$$

Example 3.7 ($\mathbf{1}[x > 1]/x$)**.** Here $\lambda(t) = \ln t$ and $X_n = e^{S_n}$. Then

$$\begin{aligned}\int_{X_{j+1}}^{2X_{j+1}-X_j} f\,dx &\ge (X_{j+1} - X_j)\cdot f(2X_{j+1} - X_j) = \frac{X_{j+1} - X_j}{2X_{j+1} - X_j}\\ &\ge \frac{1}{2}\left(1 - \frac{X_j}{X_{j+1}}\right) = \frac{1}{2}\big(1 - e^{-Z_{j+1}}\big) \overset{\mathrm{d}}{=} \frac{1}{2}\cdot \mathrm{Unif}(0,1).\end{aligned}$$

A sum of uniforms diverges by Kolmogorov's Three-Series Theorem, so (3.1) never occurs. (Note with $1/x$ we can count slope decreases more carefully in §7.)

Example 3.8 ($\mathbf{1}[x > 0]/x^q$ with $0 < q < 1$ and $p = 1 - q$)**.** Now $\lambda(t) = t^p/p$, hence $X_n = (pS_n)^{1/p}$. This means

$$\begin{aligned}\int_{X_{j+1}}^{2X_{j+1}-X_j} f\,dx &\ge (X_{j+1} - X_j)f(2X_{j+1} - X_j)\\ &\ge \frac{X_{j+1} - X_j}{(2X_{j+1})^q} \ge \frac{1}{2^q}\left(X_{j+1}^p - X_j^p\right) = \frac{p}{2^q}Z_{j+1}.\end{aligned}$$

Again $\sum_{j=1}^{\infty} Z_{j+1}$ diverges by Kolmogorov's Three-Series Theorem, so (3.1) does not hold.

Example 3.9 ($\mathbf{1}[x > e]/\log x$)**.** We check the two conditions for Proposition 3.2. As $t \to \infty$, we have the well-known estimates (see [36, §2.6 and §3.4])

$$\lambda(t) = \int_e^t \frac{1}{\log x}\, dx \sim \frac{t}{\log t} \quad \text{and} \quad \lambda^{-1}(t) \sim t \log t, \tag{3.3}$$

but it is enough to use $t \ll \lambda^{-1}(t) < t^2$. The law of large numbers guarantees $n/2 < S_n < 2n$ once $n > N$, for some N; selecting any fixed, positive number M, we see

$$\frac{1}{\log(X_j + M)} > \frac{1}{\log(S_j^2 + M)} > \frac{1}{\log(4j^2 + M)} \quad \text{when } j > N.$$

This shows $\sum_{j=1}^{\infty} f(X_j + M) = \infty$ almost surely. Also $X_j > \lambda^{-1}(j/2) \gg j$.

Example 3.10 ($\mathbf{1}[x > 1]/x \log^q x$ with $1/2 < q < 1$ and $p = 1 - q$)**.** With this intensity, $\lambda(t) = \frac{1}{p} \log^p t$ and $\lambda^{-1}(s) = \exp\left((ps)^{1/p}\right)$. Then

$$\int_{X_{j+1}}^{2X_{j+1}-X_j} f\, dx \le \frac{X_{j+1} - X_j}{X_{j+1} \log^q X_{j+1}} \le \frac{1}{\log^q X_{j+1}} = \frac{1}{(pS_{j+1})^{q/p}}.$$

Almost surely, we eventually see $S_{j+1} > j/2$, and then

$$\int_{X_{j+1}}^{2X_{j+1}-X_j} f\, dx < \left(\frac{2}{p}\right)^{q/p} \frac{1}{j^{q/p}}.$$

Now $\sum_j 1/j^{q/p} < \infty$ if $q/p > 1$. So, if $q > 1/2$, then (3.1) holds.

Remark 3.11. So far, simple bounds like these do not seem precise enough when $0 < q \le 1/2$. Another approach is discussed later, in Remark 7.4 and Proposition 7.9. *

Chapter 4

Relative Counting

In this section we simultaneously consider two Poisson processes, $X_1, X_2, \ldots$ from intensity $f(x)$ and $Y_1, Y_2, \ldots$ from intensity $g(x)$. The goal is to make relative statements about the number of extreme points.

Before continuing, we give some vague intuition. Recall there are two ways we can count extreme points and two ways we might try to compare intensities:

(a) number of extreme points in a region of space, $[0, t]$, comparing $\mathcal{E}_f(t)$ and $\mathcal{E}_g(t)$, or

(b) number of extreme points among certain indices, comparing $\mathcal{E}_f(X_n)$ and $\mathcal{E}_g(Y_n)$.

We count like (b) here. Really, this is to recycle a trick Pomerance [33] used with the primes. Nevertheless, counting by index is also intuitive, because when we count by space there are competing factors. Suppose, for instance, that $f(x) \leq g(x)$ and these intensities are chosen from Table 1.2. The probabilities in that table tell us that, counting by space,

- $f(x)$ has fewer points in $[0, t]$, but each one is more likely to be extreme, while
- $g(x)$ has more points in $[0, t]$, but each one is less likely to be extreme.

Proposition 4.1. It is possible to couple Poisson processes $\left\{X_n^{(q)} : n = 1, 2, \ldots\right\}$ of all intensities $1/x^q$ simultaneously so that whenever $0 < q' < q < 1$,

$$\left(n, X_n^{(q')}\right) \text{ is extreme} \Rightarrow \left(n, X_n^{(q)}\right) \text{ is extreme.}$$

This coupling gives $\mathcal{E}_{1/x^q}\left(X_n^{(q)}\right) \geq \mathcal{E}_{1/x^{q'}}\left(X_n^{(q')}\right)$.

Proof. First we run the proof with only two values, $0 < q' < q < 1$. Simplify the notation somewhat by setting $X_n = X_n^{(q)}$ and $X'_n = X_n^{(q')}$. Couple these with the same random walk S_n of Exp(1) steps, as in Lemma 2.1:

$$X_n = (pS_n)^{1/p} \qquad \text{and} \qquad X'_n = (p'S_n)^{1/p'},$$

where $p = 1 - q$ and $p' = 1 - q'$ as usual. This coupling starts both processes at $x = 0$.

Now here is where we adapt the convexity trick that Pomerance used for the primes; compare with [33, bottom to top of pp. 400–401]. If (m, X_m) is *not* an extreme point in the q process, then there are indices ℓ, n, with $\ell < m < n$, so that

$$X_m > \frac{n-m}{n-\ell} X_\ell + \frac{m-\ell}{n-\ell} X_n,$$

or when we divide by X_m,

$$1 > \frac{n-m}{n-\ell}\left(\frac{S_\ell}{S_m}\right)^{1/p} + \frac{m-\ell}{n-\ell}\left(\frac{S_n}{S_m}\right)^{1/p}. \tag{4.1}$$

With ℓ, m, n fixed, and treating the random walk as temporarily fixed, too, define a function of x by replacing the exponent $1/p$ with x:

$$h(x) = \frac{n-m}{n-\ell}\left(\frac{S_\ell}{S_m}\right)^{x} + \frac{m-\ell}{n-\ell}\left(\frac{S_n}{S_m}\right)^{x}.$$

Note $h(0) = 1$ and $h(1/p) < 1$ by (4.1). Also, $h''(x) \geq 0$, so $h(x)$ is convex. From there it must also be that $h(x) < 1$ when $0 < x < 1/p$, in particular for $x = 1/p' < 1/p$. Yet,

$h(1/p') < 1$ means

$$X'_m > \frac{n-m}{n-\ell} X'_\ell + \frac{m-\ell}{n-\ell} X'_n,$$

showing (m, X'_m) is not extreme. As described, this coupling works with all $0 < q < 1$ simultaneously. □

Chapter 5

Sampling

In this section, we describe an algorithm to generate accurate samples of $\mathcal{E}_f(t)$. An equivalent goal is to generate a picture of the process inside a window of arbitrary height $t = W$ with all extreme points marked accurately. The difficulty here is, again, that extremeness of $X_j \leq W$ inside the window may depend on points $X_k > W$ far outside of this window.

Proposition 5.1. Suppose $f(x)$ is such that $\lambda^{-1}(s) \gg s$ is superlinear, where as usual $\lambda(t)$ is given by (1.3). There is an algorithm to sample a Poisson process of intensity $f(x)$ inside any window $[0, W]$ with all points accurately marked extreme or nonextreme. Equivalently, there is an algorithm to generate samples of $\mathcal{E}_f(t)$.

Remark 5.2. The assumption here is no real restriction. It holds for any $f(x)$ satisfying (1.2). Since $\lambda(t) \to \infty$ as $t \to \infty$, substitute $s = \lambda(t)$ and use $\lambda'(t) = f(t)$ to check

$$\lim_{s\to\infty} \frac{\lambda^{-1}(s)}{s} = \lim_{t\to\infty} \frac{t}{\lambda(t)} = \lim_{t\to\infty} \frac{1}{f(t)} = \infty.$$

We emphasize $\lambda^{-1}(s) \gg s$ since it is the key allowing us to ignore most of the process.

One might also think about this as $\lambda(t) \ll t$. In the prime context (compare (3.3) and the prime number theorem [6, p. 10]), this is like saying the prime counting function is sublinear, so it seems very likely that McNew [31] and Tutaj [38] checked numerical estimates on extreme primes using some procedure like what is described below. *

To summarize the algorithm in just a few sentences, we use standard rejection-sampling with conditioned walks to generate a sequence S_n—as usual of i.i.d. Exp(1) steps—coupled with a time τ after which $S_n > n/2$. Because $\lambda^{-1}(s)$ is superlinear, we will see that $X_n = \lambda^{-1}(S_n) > \lambda^{-1}(n/2)$ will eventually sit above any potential support line to any point inside the window; that means we only need to see *finitely many* points to correctly select extreme points inside the window.

Details omitted in the last paragraph are filled in by §5.1, assuming we can generate S_n coupled with τ. What makes this difficult is τ is not a stopping time. An algorithm to generate S_n with τ is in §5.2, assuming it is possible to compute a particular probability exactly. Computing this probability reduces to solving a recurrence integral equation, which is done in §5.3 using an iterative suggestion from Robin Pemantle.

5.1 Sampling Overview

Once again, we begin with a random walk S_n of Exp(1) steps and set $X_n = \lambda^{-1}(S_n)$. Throughout the argument, fix a number $0 < \rho < 1$. This number may be freely selected and merely replaces the 1/2 above, in $S_n > n/2$. Why to generalize to ρ is explained later, in Remark 5.5, but in reading, it may help simply to think $\rho = 1/2$.

Lemma 5.3. Given some starting time $v \geq 0$ and starting point S_v of the Exp(1) random walk, it is possible to jointly sample $(S_{v+1}, S_{v+2}, \ldots, S_\tau, \tau)$, where $\tau \geq v$ is the first time after which $S_n \geq \rho n$ always; that is,

$$\tau = \inf\{r \geq v \ : \ S_n \geq \rho n \text{ at every time } n \geq r\}. \tag{5.1}$$

Note τ is not a stopping time.

Proof of Proposition 5.1 using Lemma 5.3. Begin by generating $X_n = \lambda^{-1}(S_n)$ for $n \leq M = \max\{m+1, N\}$, where m is the random index that satisfies

$$X_m \leq W < X_{m+1},$$

and where $N \in \mathbb{N}$ is a deterministic index at which

$$\lambda^{-1}(\rho s) \geq sW \quad \text{when} \quad s \geq N. \tag{5.2}$$

Such an index N exists, since we assume $\lambda^{-1}(s) \gg s$. Starting from S_M, jointly sample $(S_M, \ldots, S_\tau, \tau)$, $\tau \geq M$, as in Lemma 5.3.

Now, X_1 is always marked extreme, so consider any other index $1 < j \leq m$. Deciding whether $X_j \leq W$ is extreme amounts to examining the line

$$\ell_j(n) = X_j + (n-j)s_j \qquad \text{with slope} \qquad s_j = \max_{i<j} \frac{X_j - X_i}{j - i}.$$

Here s_j is the lowest possible slope from the perspective of the LHS, so it suffices to check whether points to the right of X_j are above ℓ_j. Notice $s_j \leq X_j \leq W$, so that

$$nW = W + (n-1)W \geq \ell_j(n) \quad \text{whenever} \quad n > 1.$$

Now whenever $n \geq \tau$, use (5.1) then (5.2) to say

$$X_n = \lambda^{-1}(S_n) \geq \lambda^{-1}(\rho n) \geq nW \geq \ell_j(n).$$

It remains to check whether

$$X_n \geq \ell_j(n) \quad \text{when} \quad j < n < \tau, \tag{5.3}$$

but this is possible, since we already sampled S_n at all $n \leq \tau$ using Lemma 5.3. If (5.3) holds, then X_j is guaranteed to be extreme; if (5.3) does not hold, then X_j is not extreme. Therefore, marking points accurately amounts to checking (5.3) at each index $j = 2, 3, \ldots, m$. $\square$

5.2 Algorithmic Proof to Lemma 5.3

First, rephrase Lemma 5.3 by incorporating the drift into the walk. We may as well start the walk at time $v = 0$, also. The setup is then equivalent to a walk $\widetilde{S}_n = S_n - \rho n$ of i.i.d. $\text{Exp}(1) - \rho$ steps, starting at some $\widetilde{S}_0$, and our task is to sample

$$\left(\widetilde{S}_1, \ldots, \widetilde{S}_\tau, \tau\right) \qquad \text{so that} \qquad \widetilde{S}_n \geq 0 \text{ once } n \geq \tau.$$

As previewed, the difficulty is τ is not a stopping time. Nevertheless, there is a way around this if we know how to compute the function

$$H(h) = \mathbb{P}\left(\tau_- < \infty \mid \widetilde{S}_0 = h\right), \quad \text{where} \quad \tau_- = \inf\{n \geq 0 \ : \ \widetilde{S}_n < 0\} \cup \{+\infty\}. \tag{5.4}$$

How to compute $H(x)$ is explained later, in §5.3. For now we describe how to sample $\widetilde{S}_n$ with τ if $H(x)$ is known. First we provide the algorithmic details, then we summarize the steps with a picture. Begin with $r = 0$ and $n_0 = 0$.

1. Suppose we have $\widetilde{S}_n$ generated for $n \leq n_r$. Since $\widetilde{S}_n$ has drift $1 - \rho > 0$, we may wait until the first time $n_{r+1} \geq n_r$ that $\widetilde{S}_{n_{r+1}} \geq 0$. Call this height $h_{r+1} := \widetilde{S}_{n_{r+1}} \geq 0$. Update $r := r + 1$ and continue.
2. We have $h_r = \widetilde{S}_{n_r} \geq 0$. Compute $H(h_r)$ (as in §5.3).

3. Generate $U_r \sim \text{Unif}(0,1)$ and compare it to $H(h_r)$. By the Markov property for $\widetilde{S}_n$, this samples a hypothetical infinite continuation of the path $\widetilde{S}_n$ for all $n \geq n_r$ from $\widetilde{S}_{n_r} = h_r$, and the comparison to $H(h_r)$ determines whether the rest of the walk avoids negative values or hits them again. Here are the alternatives.

   (a) If $U_r > H(h_r)$, then by definition (5.4), the hypothetical continuation of our path ($\widetilde{S}_n$ for $n \geq n_r$, starting from $\widetilde{S}_{n_r} = h_r$) has $\tau_- = \infty$ and avoids returning to negative values. That is,

   $$\widetilde{S}_n \geq 0 \qquad \text{for every} \qquad n \geq n_r.$$

   Stop and return $\tau := n_r$ along with the path $\widetilde{S}_n$, $n \leq \tau$.

   (b) If $U_r \leq H(h_r)$, then the path from $\widetilde{S}_{n_r} = h_r$ onward is going to hit negative values again. Proceed to the next step.

4. We now generate a continuation of the path $\widetilde{S}_n$ for $n \geq n_r$ that is conditioned to hit negative values. The ordinary one-step transition kernel for $\widetilde{S}_n$ from x is

   $$p(x,y) = e^{-(y-x+\rho)},$$

   since each step $y - x \sim \text{Exp}(1) - \rho$. A step of the conditioned walk may be generated by rejection-sampling y from the Doob H-transform transition kernel

   $$p_H(x,y) = p(x,y) \cdot \frac{H(y)}{H(x)}$$

   (see [30, p. 257 and p. 381]). That is, generate y from $p(x,y)$, generate $U \sim \text{Unif}(0,1)$, and accept y if $U < p(x,y)H(y)$, otherwise reject and regenerate y and U. This works since $H(y) \leq 1$ is a probability and x is temporarily fixed, meaning we have the bound

   $$\frac{p_H(x,y)}{p(x,y)} \leq \frac{H(y)}{H(x)} \leq \frac{1}{H(x)}.$$

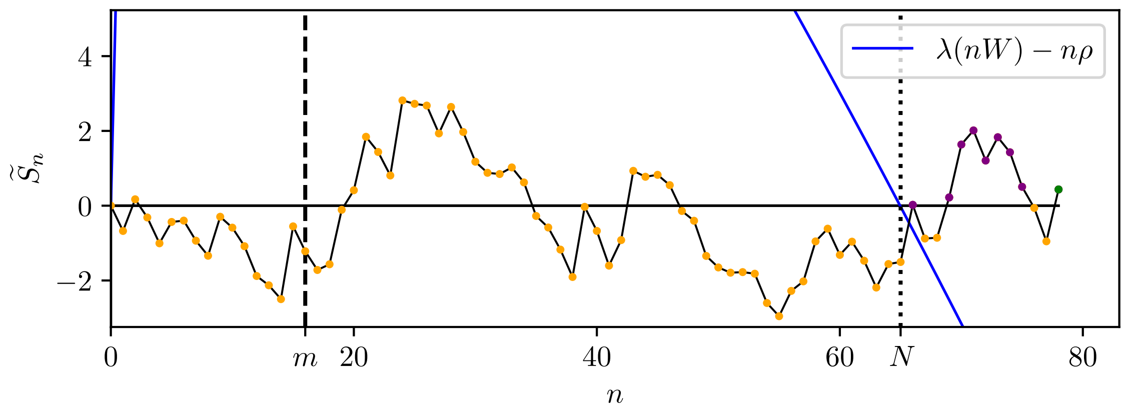

Figure 5.1: The conditioned walk $\widetilde{S}_n$ of $\text{Exp}(1) - \rho$ steps. Recall N is the index from (5.2).

> Repeatedly sampling $y = \widetilde{S}_{n+1}$ from $p_H(\widetilde{S}_n, \cdot)$ allows us to continue the walk until a time n_{r+1} when $\widetilde{S}_{n_{r+1}} < 0$. Increment $r := r + 1$ and return to the first step.

As promised, we illustrate in Figure 5.1. We generate Exp(1) steps at all orange points up to index N and further until $\widetilde{S}_n \geq 0$ at index n_r (in this picture, just one step further, $n_r = N+1$). Thereafter, at every such time n_r when the walk crosses into positive values, we flip a coin with probability $H(\widetilde{S}_{n_r})$ to decide if the walk ever returns to negative values, and if so, we take steps according to p_H at purple points until that happens. In this picture, the walk returns to negative values twice—first after one step, second after seven steps. Once a coin-flip decides $\widetilde{S}_n$ remains nonnegative, we stop at the green point. The walk transformed by λ^{-1} is shown in Figure 5.2.

5.3 Computing $H(x)$

When $x < 0$, the definition gives $H(x) = 1$. We also know $H(0) = \rho$ by Lemma 2.5. (Since steps $Z_n - \rho \sim \text{Exp}(1) - \rho$ of $\widetilde{S}_n$ are from a continuous distribution, it does not matter if

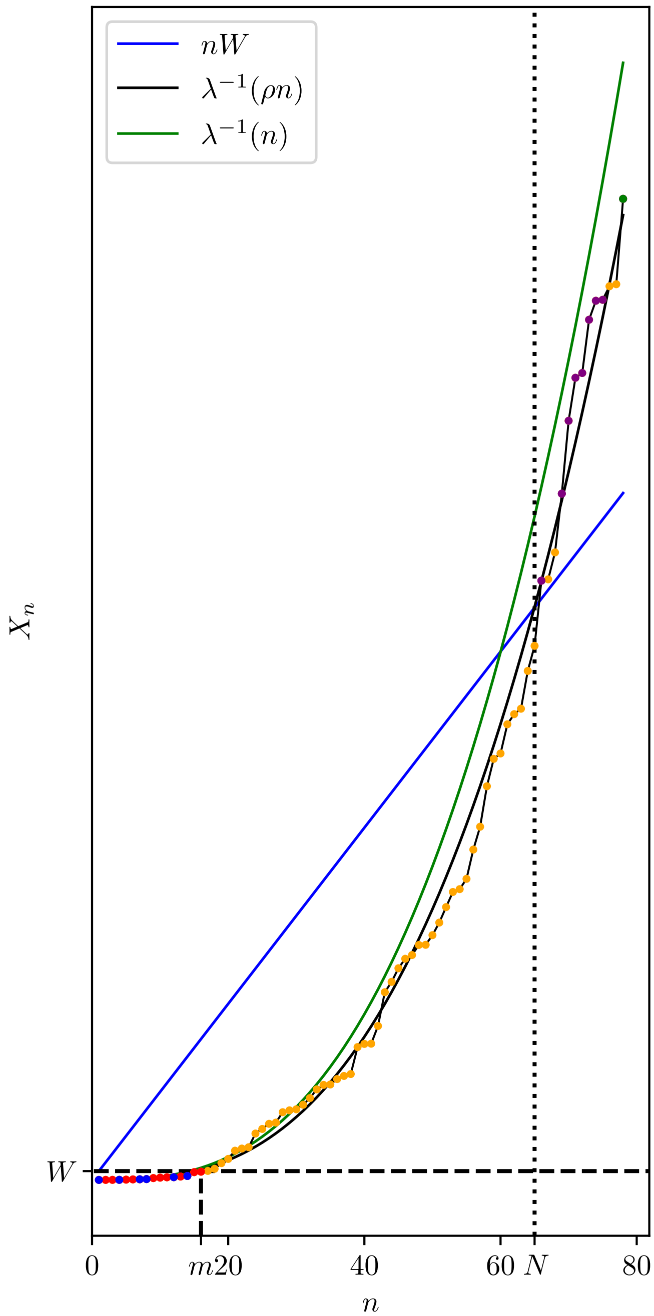

Figure 5.2: A process (corresponding to the walk in Figure 5.1) in which a small window of points near the bottom can be marked extreme or nonextreme accurately.

the inequality $\widetilde{S}_n > 0$ is strict.) To find other values, consider $x \geq 0$ and the first step:

$$\begin{aligned} H(x) &= \mathbb{E}[H(x + Z_1 - \rho)] \\ &= \int_0^\infty H(x + z - \rho) \cdot e^{-z}\, dz \\ &= e^{x-\rho} \int_{x-\rho}^\infty H(u) e^{-u}\, du \qquad (u = x + z - \rho). \end{aligned} \tag{5.5}$$

Thanks go to Robin for the suggestion of how to solve (5.5). This is done in two steps:

- prove continuity of H by a coupling method, then
- solve on intervals $i\rho \leq x \leq (i+1)\rho$ successively ($i = 1, 2, 3, \dots$), separately considering in (5.5) when the steps are too short or long enough to reach the next interval.

Lemma 5.4. $H(x)$ is decreasing, right-continuous at $x = 0$, and continuous elsewhere.

Proof. That $H(x)$ is decreasing is immediate from the definition (5.4). For the continuity, we use Robin's coupling suggestion, as just mentioned. Define two copies of our Exp(1) random walk, one started from height 0 and another from height ϵ. Write these as

$$S_{0,n} = 0 + \sum_{j=1}^n Z_{0,j} \quad \text{and} \quad S_{\epsilon,n} = \epsilon + \sum_{j=1}^n Z_{\epsilon,j}$$

and set $Z_{0,j} = Z_{\epsilon,j}$ for every $j \geq 2$. By the memoryless property, whenever $Z_{0,1} > \epsilon$ we may couple to $Z_{\epsilon,1}$ by letting $Z_{0,1} = \epsilon + Z_{\epsilon,1}$. Otherwise, draw $Z_{0,1}, Z_{\epsilon,1}$ independently. When the first steps do couple, the two walks agree at all times except $n = 0$, in which case

$$H(0) - H(\epsilon) \leq \mathbb{P}(Z_{0,1} < \epsilon) = 1 - e^{-\epsilon} \to 0 \quad \text{when } \epsilon \to 0.$$

The same proof works starting at x and $x+\epsilon$ to show $H(x) - H(x+\epsilon) \leq \mathbb{P}(Z_{x,1} < \epsilon) \to 0$. □

Now we can solve iteratively for $H(x)$. We begin with a concrete example, the interval $0 \leq x \leq \rho$, although the general case is no more difficult. We already know $H(x) = 1$ when

$x < 0$ and $H(0) = \rho$, so when $0 \le x \le \rho$, (5.5) gives

$$H(x) = e^{x-\rho}\bigg(\int_{x-\rho}^{0} + \int_{0}^{\infty} H(u)e^{-u}\,du\bigg).$$

To compute the final integral on this line, rearrange and plug in $x = 0$:

$$\begin{aligned}\int_0^\infty H(u)e^{-u}\,du &= \bigg[e^{\rho-x}H(x) - \int_{x-\rho}^{0} H(u)e^{-u}\,du\bigg]\bigg|_{x=0} \\ &= \rho e^{\rho} - \int_{-\rho}^{0} e^{-u}\,du \\ &= 1 - (1-\rho)e^{\rho}.\end{aligned}$$

Substituting this above, we find that whenever $0 \le x \le \rho$,

$$\begin{aligned}H(x) &= e^{x-\rho}\bigg(\int_{x-\rho}^{0} e^{-u}\,du + \int_0^\infty H(u)e^{-u}\,du\bigg) \\ &= e^{x-\rho}\big(e^{\rho-x} - 1 + 1 - (1-\rho)e^{\rho}\big) \\ &= 1 - (1-\rho)e^{x}.\end{aligned}$$

Suppose more generally we have an expression for $H(x)$ when $x \le i\rho$. Then when $i\rho \le x \le (i+1)\rho$, (5.5) says

$$H(x) = e^{x-\rho}\bigg(\int_{x-\rho}^{i\rho} + \int_{i\rho}^{\infty} H(u)e^{-u}\,du\bigg)$$

Plugging in $x = i\rho$ and rearranging, we learn

$$\int_{i\rho}^{\infty} H(u)e^{-u}\,du = e^{-(i-1)\rho}H(i\rho) - \int_{(i-1)\rho}^{i\rho} H(u)e^{-u}\,du.$$

The right hand side is a known quantity, so we can substitute above and find

$$\begin{aligned}H(x) &= e^{x-\rho}\bigg(\int_{x-\rho}^{i\rho} H(u)e^{-u}\,du + e^{-(i-1)\rho}H(i\rho) - \int_{(i-1)\rho}^{i\rho} H(u)e^{-u}\,du\bigg) \\ &= e^{x-\rho}\bigg(-\int_{(i-1)\rho}^{x-\rho} H(u)e^{-u}\,du + e^{-(i-1)\rho}H(i\rho)\bigg) \qquad \text{when} \qquad i\rho \le x \le (i+1)\rho.\end{aligned}$$

A computer can iterate to find exact expressions for $H(x)$ so long as $\rho \in \mathbb{Q}$, since then every term is of the form $cx^r \cdot e^{ax+b}$ with $a, b, c \in \mathbb{Q}$ and $r \in \mathbb{Z}_{\geq 0}$. We can check the computer by hand at first (see Table 5.1), but soon the calculations become tedious and a computer is helpful. Figure 5.3 shows a picture of the function H and various choices of ρ.

Remark 5.5. One may ask how to choose ρ.[7] Choosing $\rho \approx 0$ will be inefficient if N in (5.2) becomes very large. For example, $f(x) = 1/\sqrt{x}$ has $N \geq 4W/\rho$. Yet, $\rho \approx 1$ might delay τ in (5.1), as now explained. The walk $\widetilde{S}_n$ from §5.2 returns to negative values until we generate a successful coin-flip of probability $1 - H(\widetilde{S}_{n_r})$. Since the first positive value $\widetilde{S}_{n_r} \sim \text{Exp}(1)$ (memoryless), the expectation of this coin-flip probability is

$$\mathbb{E}\Big(1 - H(\widetilde{S}_{n_r})\Big) = 1 - \int_0^\infty H(s)e^{-s}\, ds = (1-\rho)e^\rho$$

as computed above. In any case, this probability is always at least $1 - H(0) = 1 - \rho$. Considering Figure 5.3, this should not be an issue unless ρ is extremely close to 1. *

[7] In Figures 5.1 and 5.2, ρ was artificially chosen with the walk.

i	$H(x)$ when $i/2 \leq x \leq (i+1)/2$ and with $\rho = 1/2$
$i < 0$	1
$i = 0$	$1 - \frac{1}{2}e^x$
$i = 1$	$1 - \frac{1}{2}e^x + \frac{1}{2}xe^{x-\frac{1}{2}} - \frac{1}{4}e^{x-\frac{1}{2}}$
$i = 2$	$1 - \frac{1}{2}e^x + e^{x-\frac{1}{2}}\left(\frac{1}{2}x - \frac{1}{4}\right) + e^{x-1}\left(-\frac{1}{4}x^2 + \frac{1}{2}x - \frac{1}{4}\right)$

Table 5.1: Expressions for $H(x)$ when $\rho = 1/2$ and $x \leq 3/2$, checked by hand and by computer.

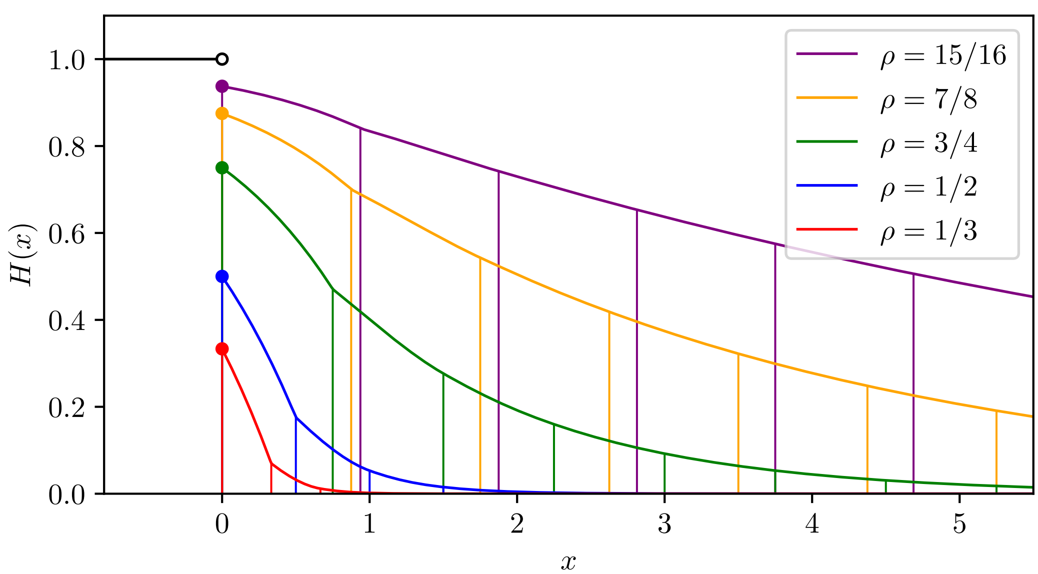

Figure 5.3: The function $H(x)$ defined in (5.4) with various choices of ρ. The initial black segment applies to all values ρ. Vertical lines separate the regions $[i\rho, (i+1)\rho]$.

Chapter 6

Expectation Bounds: Setup

This section provides setup needed in §7–§9 to justify Table 1.2.

6.1 Transformation and Reasonable Slope Choice

The introductory sketch in §1.3.4(a) and §1.3.4(c)(i)–(ii) left a few questions unanswered.

1. How do we view the process as a random walk?
2. What slope β should we choose for the lower bound?
3. What slope zones (1.4) should we choose for the upper bound?

Selecting slope zones for the upper bound will be deferred to the proof sketch in §9.1. We can answer both of the other questions now, though.

To view the process as two random walks, simply revisit Lemma 2.1. This tells us if we map the Poisson process via $X_n \mapsto \lambda(X_n) - n$, then we obtain a random walk of i.i.d. $\text{Exp}(1) - 1$ steps. We can also transform any potential support line like so, obtaining a nonlinear boundary that these $\text{Exp}(1) - 1$-step random walks must avoid. Details of the transformation are given in §6.2, but a picture was already shown in Figure 1.6.

Next, what is a reasonable choice for a slope β_* to use in a lower bound

$$\mathbb{P}(X \text{ is extreme} \mid X) \geq \mathbb{P}(X \text{ is extreme with a line of slope } \beta_* \mid X)? \tag{6.1}$$

(The conditioning will be explained later, in Remark 6.2.) From Lemma 2.3, say, we know the plot of the points (n, X_n) ought to look something like the plot of the function $s \mapsto \lambda^{-1}(s)$. If so, a reasonable choice of slope is the slope of this curve's tangent line at X—that is, at height $s = S_i$ if X happens to be $X_i = \lambda^{-1}(S_i)$—but this is

$$(\lambda^{-1})'(S_i) = \frac{1}{\lambda'(\lambda^{-1}(S_i))} = \frac{1}{f(X)},$$

and so we define

$$\beta_* \stackrel{\text{def}}{=} \frac{1}{f(X)}. \tag{6.2}$$

Another way to interpret β_* is to look back at the informal calculation in §1.3.5. There, this slope also removes the linear Taylor term in (1.11).

Remark 6.1. With $f(x) = 1/\log x$, where we have (3.3), the slope $\beta_* = \log X_n \approx \log(n \log n)$ is easily compared to the prime slopes given by McNew [31, Lemma 1]. *

6.2 Events of Interest

Using the notation of Lemma 2.1, and only temporarily including a fixed index i, the event that X_i *looks extreme to the right with slope* β is (revisit Figure 1.5)

$$\begin{aligned}
&\{X_{i+n} \geq X_i + n\beta \quad \text{for all} \quad n > 0\} \\
&\qquad = \left\{S_i + \textstyle\sum_{j=1}^{n} Z_{i+j} \geq \lambda(X_i + n\beta) \quad \text{for all} \quad n > 0\right\} \\
&\qquad = \left\{\textstyle\sum_{j=1}^{n} Z_{i+j} \geq \lambda(X_i + n\beta) - \lambda(X_i) \quad \text{for all} \quad n > 0\right\} \\
&\qquad = \left\{\mathcal{S}_n^R \geq \lambda(X_i + n\beta) - \lambda(X_i) - n \quad \text{for all} \quad n > 0\right\}
\end{aligned}$$

where on the last line we introduce $\mathcal{S}_n^R$, a mean-zero random walk started at $\mathcal{S}_0^R = 0$ with i.i.d. $\mathrm{Exp}(1) - 1$ steps (namely $\mathcal{Z}_j^R = Z_{i+j} - 1$ for $j = 1, 2, \ldots$). Similarly, the event that X_i *looks extreme to the left with slope* β is

$$\begin{aligned}\{X_{i-n} \geq X_i - n\beta \quad \text{for all} \quad 0 < n < i\} \\ = \big\{\mathcal{S}_n^L \leq \lambda(X_i) - \lambda(X_i - n\beta) - n \quad \text{for all} \quad 0 < n < i \wedge (X_i/\beta)\big\}\end{aligned} \tag{6.3}$$

where $\mathcal{S}_n^L$ is also a random walk started at $\mathcal{S}_0^L = 0$ with $\mathrm{Exp}(1)-1$ steps. The final minimum, $\wedge X_i/\beta$, is introduced when we apply λ, since a line of slope β through X crosses below the horizontal axis exactly X_i/β units to the left of X_i. (See Figure 1.5.) To the left of that crossing, all remaining points are necessarily above the line. To summarize, we define the two events

$$E_{i,R,\beta} \stackrel{\text{def}}{=} \big\{\mathcal{S}_n^R \geq \lambda(X_i + n\beta) - \lambda(X_i) - n \quad \text{for all} \quad n > 0\big\} \tag{6.4}$$

$$E_{i,L,\beta} \stackrel{\text{def}}{=} \big\{\mathcal{S}_n^L \leq \lambda(X_i) - \lambda(X_i - n\beta) - n \quad \text{for all} \quad 0 < n < i \wedge (X_i/\beta)\big\}. \tag{6.5}$$

These lead us to introduce the following two events with a fixed location X and slope β in mind, but with the index i dropped:

$$E_{R,\beta} \stackrel{\text{def}}{=} \big\{\mathcal{S}_n^R \geq \lambda(X + n\beta) - \lambda(X) - n \quad \text{for all} \quad n > 0\big\} \tag{6.6}$$

$$E_{L,\beta} \stackrel{\text{def}}{=} \big\{\mathcal{S}_n^L \leq \lambda(X) - \lambda(X - n\beta) - n \quad \text{for all} \quad 0 < n < n_X\big\}. \tag{6.7}$$

Dropping i requires a yet-undefined index n_X in (6.7). Our lower bound (6.1) in the notation just introduced becomes

$$\begin{aligned}\mathbb{P}(X \text{ is extreme} \mid X) &= \mathbb{P}\Big(\textstyle\bigcup_{\beta \geq 0} E_{L,\beta} \cap E_{R,\beta} \;\Big|\; X\Big) \\ &\geq \mathbb{P}(E_{L,\beta_*} \cap E_{R,\beta_*} \mid X) = \mathbb{P}(E_{L,\beta_*} \mid X) \cdot \mathbb{P}(E_{R,\beta_*} \mid X).\end{aligned} \tag{6.8}$$

We now discuss the conditioning (Remark 6.2) and the index n_X (Remark 6.3).

Remark 6.2 (Conditioning)**.** The conditioning on X is meant to indicate that the Poisson process has a point at X, however that point is not assumed to be the ith point for any nonrandom i. The index, if known, would be a hindrance. If we knew $X = X_i = \lambda^{-1}(S_i)$, then the steps $\mathcal{Z}_1^L, \ldots, \mathcal{Z}_i^L$ for the LHS walk would no longer be i.i.d. $\text{Exp}(1) - 1$ steps; rather these i steps would be conditioned to sum to $\lambda(X_i) - i = S_i - i$. When we do not condition on X being the ith point, we have the advantage that the distribution of the steps $\mathcal{Z}_1^L, \mathcal{Z}_2^L, \ldots$ is unchanged. Formally, that the process looks the same is a statement about the Mecke equation for a Poisson process [29, see Theorem 4.1]; more specifically, ignoring the new point at X, everything else about the process looks the same in distribution as it did before.

Of course, if there is a point at X, there is some true but random index i such that $X = X_i$, and the rest of the walk on the LHS ($\mathcal{S}_j^L$ with $j > i$) is meaningless. For this reason n_X must be defined carefully. *

Remark 6.3 (Index n_X)**.** Our choice of n_X in (6.7) will depend on what we calculate. However, there are two useful observations common to all situations.

- By Lemma 2.3, $i \sim S_i = \lambda(X_i)$ a.s., meaning $\lambda(X)$ is an asymptotically correct estimate for the index of a point at location X.
- As already remarked (see below (6.3) on p. 60), the line of slope β crosses the horizontal axis X/β units to the left of X, so we should take $n_X \leq X/\beta$.

For a lower bound on $\mathbb{P}(E_{L,\beta_*} \mid X)$ we may use $n_X = X/\beta_* = X \cdot f(X)$, even though it may not always make sense. Suppose for instance that $X = X_i$ but $n_X = Xf(X) > i$ is larger than the true index of X. This is fine for a lower bound on $\mathbb{P}(E_{L,\beta_*} \mid X)$, because whenever this happens the event (6.7) just requires the LHS walk to survive longer than necessary, which only reduces the probability. Consider the important examples:

- With $f(x) = 1/x^q$, then $i \sim X_i^p/p > X_i^p = X_i \cdot f(X_i)$, so it almost always makes sense to look $n_X = Xf(X)$ steps to the left of X.
- With $f(x) = 1/\log x$, then $i \sim X_i/\log X_i = X_i f(X_i)$ by (3.3). In this case $n_X = X/\log X$ is usually close to the true index, but it may sometimes be larger than the true index. Again, this is fine for a lower bound.

When we look at an upper bound on $\mathbb{P}(E_{L,\beta} \mid X)$ for various β, the walk is required to survive for X/β steps leftward, so we may take any n_X subject to both of the following conditions:

- n_X must be an index, so for instance $n_X < (1-\delta) \cdot \lambda(X)$ for $\delta > 0$ as $X \to \infty$, and
- n_X must be less than X/β.

For upper bounds we will actually use a small multiple of X/β, for instance with $f(x) = 1/x^q$, we will take $n_X = \epsilon X^p$. *

Note 6.4. More correctly, we might write an index like X^p or $X^{\frac{2}{3}p}$ or $X/\log X$ with floor or ceiling functions, but we will not. Either way would not matter: one step more or one step fewer would only ever lose or gain a constant factor in the probability.

Chapter 7

Intensity $1/x$

Intensity $1/x$ is simpler than $1/x^q$ and $1/\log x$, and we can find more information about it, so we treat it first and separately. That said, $1/x$ does involve a mishmash of ideas.

Partly using the random walk setup from §6, and partly using the context of slope decreases from §3, we show expectation bounds (Lemmas 7.3(c) and 7.5) that together say:

Proposition 7.1. With intensity $f(x) = \mathbf{1}[x > 1]/x$,

$$\frac{1}{4}\log t \leq \mathbb{E}\mathcal{E}_f(t) \leq \log 2 \log t + O(1) \quad \text{as } t \to \infty.$$

If we count extreme points among $X_1, X_2, \ldots, X_n$ by index n, instead of how many occur within $[0, t]$, then in §7.2 we can show another kind of result. As previewed in Remark 1.17, inspired by [19] we can use stationary sequence limits to prove:

Proposition 7.2. Let $f(x) = \mathbf{1}[x > 1]/x$ and let $\mu, \sigma^2 \geq 0$ be the constants defined in Lemma 7.8. The quantity $\dfrac{\mathcal{E}_f(X_n) - \mu n}{\sqrt{n}}$ either converges in distribution to $\mathrm{N}(0, \sigma^2)$, or if $\sigma^2 = 0$, converges in probability to zero.

Finally, §7.4 returns to the intensities $1/x \log^q x$.

7.1 Expectation Bounds

We begin with the lower bound.

Lemma 7.3. Fix intensity $f(x) = \mathbf{1}[x > 1]/x$.

(a) $\mathbb{P}(X_i \text{ is extreme}) \geq 1/4$.

(b) $\mathbb{P}(X \text{ is extreme} \mid X) \geq 1/4$.

(c) $\mathbb{E}\mathcal{E}_f(t) \geq \frac{1}{4}\log t$.

Proof. The slope choice (6.2) is $\beta_* = 1/f(X) = X$. Notice with this slope, X automatically looks extreme to the left, since a support line crosses the horizontal axis immediately ($X/\beta = 1$ in Figure 1.5). In other words, $\mathbb{P}(E_{L,\beta_*} \mid X) = 1$. Since we are free to ignore the LHS, we can consider a particular point X_i rather than just a location X, so establishing (a) will be the same as establishing (b). The relevant random walk boundary for $E_{R,X}$ is

$$\begin{aligned}\lambda(X + nX)) - \lambda(X) - n &= \log(1+n) - n \\ &\leq \log 2 - \frac{1}{2} - \frac{1}{2}n \quad \text{(tangent line from } n = 1).\end{aligned}$$

Separate out the first step and then use the Markov property to see

$$\begin{aligned}\mathbb{P}(X \text{ is extreme}) &\geq \mathbb{P}\big(\mathcal{S}_n > -\tfrac{1}{2}n + \big(\log 2 - \tfrac{1}{2}\big)\ \forall\, n \geq 1\big) \\ &\geq \mathbb{P}(\mathcal{S}_1 > \log 2 - 1) \cdot \mathbb{P}\big(\mathcal{S}_n > -\tfrac{1}{2}n\ \forall\, n \geq 1\big) \overset{*}{\geq} 1/4,\end{aligned}$$

at $\overset{*}{\geq}$ using Lemma 2.5 with $M = 1/2$. This means each of the $\log t$ points we expect within $[0, t]$ has a chance $\geq 1/4$ to be extreme, which establishes (c) as in (1.6). □

Remark 7.4. In the last proof, notice $1/x$ does not need parabolas but only a line as in Remark 1.5. The same works with $1/x \log^q x$; to keep the flow, this is delayed to §7.4. *

We now establish an upper bound using a very different strategy.

Lemma 7.5. With intensity $f(x) = \mathbf{1}[x > 1]/x$, $\mathbb{E}\mathcal{E}_f(t) \le \log 2 \log t + O(1)$.

Proof. Here we count how many slope decreases occur (as in §3). Specifically, if

$$D(t) = \sum_{j=1}^{\infty} \mathbf{1}[X_{j+2} < 2X_{j+1} - X_j \text{ and } X_{j+1} \le t] \tag{7.1}$$

is the number of slope decreases within $[0, t]$ among the $\log t$ points expected there, then

$$\mathbb{E}\mathcal{E}_f(t) \le \log t - \mathbb{E}D(t). \tag{7.2}$$

To compute $\mathbb{E}D(t)$, note that if $X_j = x \ge 1$ and $x < y < z$,

$$\mathbb{P}(x < X_{j+1} < y \mid X_j = x) = 1 - \frac{x}{y} \qquad \mathbb{P}(y < X_{j+2} < z \mid X_{j+1} = y) = 1 - \frac{y}{z},$$

so the joint density of $(X_{j+1}, X_{j+2}) = (y, z)$ given $X_j = x$ is $\big(x/y^2\big) \cdot \big(y/z^2\big) = x/yz^2$. Then

$$\begin{aligned}
&\mathbb{P}(X_{j+2} < 2X_{j+1} - X_j \text{ and } X_{j+1} \le t \mid X_j = x) \\
&= \int_x^t \int_y^\infty \frac{x}{yz^2} \mathbf{1}[z < 2y - x] \, dz \, dy \\
&= \int_x^t \frac{x}{y} \cdot \left(\frac{1}{y} - \frac{1}{2y - x}\right) dy \\
&= 1 - \frac{x}{t} - \log\left(2 - \frac{x}{t}\right).
\end{aligned} \tag{7.3}$$

Therefore, by the Mecke equation (see [29, Theorem 4.1], like in §1.3.4(d)),

$$\begin{aligned}
\mathbb{E}D(t) &= \int_1^t \frac{1}{x}\Big(1 - \frac{x}{t} - \log\left(2 - \frac{x}{t}\right)\Big) dx \\
&= \log t - \frac{t-1}{t} - \int_{1/t}^1 \frac{\log(2-u)}{u} \, du \quad \text{with } u = x/t.
\end{aligned}$$

Using the Taylor expansion $\log(2 - u) = \log 2 - \frac{1}{2}u + O(u^2)$, which is uniform over $|u| \le 1$,

$$\begin{aligned}
\mathbb{E}D(t) &= \log t - \frac{t-1}{t} - \left(\log 2 \log t - \frac{1}{2}\left(1 - \frac{1}{t}\right) + O(1)\right) \\
&= (1 - \log 2) \log t + O(1).
\end{aligned}$$

Substitute this into (7.2). ∎

7.2 Proof of Proposition 7.2

The underlying technique here, as we said above, is to use a stationary sequence with limited dependence, inspired by the argument in [19, see the proof of Lemma 2.7 and Theorem 1.3 therein]. There are two main steps:

(i) show that the events $\{X_j \text{ is extreme}\}$ (for $j = 1, 2, \ldots$) very nearly form a stationary sequence, and then

(ii) show the stationary versions can be combined despite limited dependence.

Part (i) consists of the setup through Lemma 7.7. The key result is Lemma 7.7(e), which shows $\mathcal{E}_f(X_n)$ basically agrees with a stationary version, which we will call $\mathcal{E}_{f,\text{stat}}(X_n)$. Part of this argument is somewhat related to [19, Lemma 2.7]. Part (ii) is Lemma 7.8, and that is where we really use the stationary sequence idea from [19, Lemma 2.7, Theorem 1.3]. Most of Lemma 7.8 argues about limited dependence via local quantities (compare Remark 1.17), then it calls on a standard stationary sequence limit theorem from [5].

If one reads only the conclusions to (i)–(ii)—again, Lemma 7.7(e) and Lemma 7.8—then it is easy to see these together complete the proof of Propositon 7.2. Unfortunately, both parts (i)–(ii) will involve setting up more notation.

Recall we have defined the events $E_{j,*,\beta}$ ($* = L, R$) in (6.4)–(6.5) which incorporate the index of a potential extreme point. Suppose we knew the behavior of the entire RHS walk, $\mathcal{S}_n^R$. If we did, rearranging the condition in (6.4) would say the steepest slope β with which the RHS looks extreme is

$$\beta = X_j \cdot \min_{0<n<\infty} \left\{ \frac{\exp\left(\mathcal{S}_n^R + n\right) - 1}{n} \right\}.$$

Of course, any slope $\beta > X_j$ leaves extra room on the LHS and only hurts our chances for an extreme point on the RHS (return to Figure 1.5), so it suffices to check slopes $\beta \leq X_j$. With this in mind, let

$$C_j \overset{\text{def}}{=} 1 \wedge \min_{0<n<\infty} \left\{ \frac{\exp\left(\mathcal{S}_n^R + n\right) - 1}{n} \right\},$$

and check only whether X_j looks extreme on the LHS with slope $\beta = C_j \cdot X_j$:

$$\begin{aligned}\{X_j \text{ is extreme}\} &= E_{j,L,X_j \cdot C_j} \\ &= \left\{ \mathcal{S}_n^L \leq -\log(1 - nC_j) - n \quad \text{for all} \quad 0 < n < j \wedge (1/C_j) \right\} \overset{\text{def}}{=} E(j, C_j),\end{aligned}$$

where for ease of notation we define

$$E(j, c) = \left\{ \mathcal{S}_n^L \leq -\log(1 - nc) - n \quad \text{for all} \quad 0 < n < j \wedge (1/c) \right\}. \tag{7.4}$$

In the definition, remember that $\mathcal{S}_n^L$ depends implicitly on j. Also consider the events

$$E_{\text{stat}}(j, c) = \left\{ \mathcal{S}_n^L \leq -\log(1 - nc) - n \quad \text{for all} \quad 0 < n < (1/c) \right\} \tag{7.5}$$

(which next to (7.4) lack the upper bound $n < j$). With fixed c these form a stationary sequence in j. Furthermore, the events $E_{\text{stat}}(j, C_j)$ are stationary in j, since $C_j, \mathcal{S}_n^*$ are.

One might notice a slight issue here: $\mathcal{S}_n^L$ may not exist in (7.5) if $j < n < 1/c$. (This is the same sort of issue as with n_X in Remark 6.3.) This is no trouble here, however. We may as well extend the random walk S_n from Lemma 2.1 leftwards, to values $n \leq 0$; that is, let $S_0 = 0$ and $S_n = -\sum_{j=n+1}^{0} Z_j$ for $n < 0$, where $Z_0, Z_{-1}, Z_{-2}, \ldots \overset{\text{iid}}{\sim} \text{Exp}(1)$. If $\mathcal{S}_n^L = S_n + n$ ($n \leq 0$) stands in the usual relationship to S_n, as a backwards sum of $1 - \text{Exp}(1)$ random variables, then we interpret $E_{\text{stat}}(j, c)$ as the event where a point X_j is extreme not only among the points $X_1, X_2, \ldots$ of the process but even among the *pretend points* $X_n = e^{S_n}$ with $n \leq 0$. Given that context, let

$$\mathcal{E}_{f,\text{stat}}(X_n) = \sum_{j=1}^{n} \mathbf{1}[E_{\text{stat}}(j, C_j)]$$

count the number of points X_j $(0 < j \leq n)$ that are extreme even among the pretend points X_j $(j \in \mathbb{Z})$. Lemma 7.7(e) below shows that $\mathcal{E}_f(X_n)$ will not meaningfully differ from this stationary version. To prove this, we use the following large-deviations estimate, the proof of which (omitted) is the well-known Markov-then-optimize Chernoff argument (taking $A = 1/2$ and $\rho_{1/2} = \log 2 - 1/2 \approx 0.193$ from now on for convenience).

Lemma 7.6. Let $\mathcal{S}_n$ be a sum of $\mathrm{Exp}(1) - 1$ random variables. With $0 < A < 1$ fixed, define $-\rho_A = A + \log(1 - A) < 0$. Then $\mathbb{P}(\mathcal{S}_n < -An) \leq e^{-\rho_A n}$.

As we said, the argument in the next lemma, in particular part (c), is slightly related to [19, Lemma 2.7], although the purpose is slightly different.

Lemma 7.7. The stationary events and counts are related to the originals as follows:

(a) $E_{\mathrm{stat}}(j, c) \subseteq E(j, c)$ when $c \leq 1/j$, and $E_{\mathrm{stat}}(j, c) = E(j, c)$ when $c \geq 1/j$.

(b) If $c < 1/j$, then $\mathbb{P}\Big(E(j, c)\Big) \leq e^{-\rho_{1/2}(j-1)}$ once $j \geq 8$.

(c) $\mathbb{P}\Big(E(j, c) \setminus E_{\mathrm{stat}}(j, c)\Big) \leq e^{-\rho_{1/2}(j-1)}$ once $j \geq 8$.

(d) Almost surely, $\mathbf{1}[E(j, C_j)] = \mathbf{1}[E_{\mathrm{stat}}(j, C_j)]$ for all but finitely many j.

(e) Almost surely, $\mathcal{E}_{f,\mathrm{stat}}(X_n) = \mathcal{E}_f(X_n) + O(1)$ as $n \to \infty$.

Proof. Part (a) merely distinguishes definitions (7.4) and (7.5). For part (b),

$$-\log(1 - c(j-1)) \overset{*}{<} \log j \overset{**}{<} \tfrac{1}{2}(j-1);$$

here $\overset{*}{<}$ is from $c < 1/j$, and $\overset{**}{<}$ is from

$$\frac{\log j}{j-1} \leq \frac{\log 8}{7} < \frac{1}{2} \qquad \text{when} \qquad j \geq 8. \tag{7.6}$$

Put $n = j - 1$ in definition (7.4) to see

$$\begin{aligned}\mathbb{P}(E(j,c)) &\le \mathbb{P}\big(\mathcal{S}^L_{j-1} \le -\log(1 - c(j-1)) - (j-1)\big)\\ &\le \mathbb{P}\big(\mathcal{S}^L_{j-1} < -\tfrac{1}{2}(j-1)\big)\\ &\le e^{-\rho_{1/2}(j-1)} \qquad \text{(Lemma 7.6)}.\end{aligned}$$

Part (c) follows from (a)–(b). What is more, the upper bound in part (c) does not depend on the choice of c. Therefore Borel-Cantelli with

$$\sum_{j=8}^{\infty} \mathbb{P}\Big(E(j, C_j) \setminus E_{\mathrm{stat}}(j, C_j)\Big) \le \sum_{j=8}^{\infty} e^{-\rho_{1/2}(j-1)} < \infty$$

proves (d)–(e). □

Lemma 7.8. With

$$\mu = \mathbb{P}(E_{\mathrm{stat}}(j, C_j)) \quad \text{and} \quad \sigma^2 = \mu + 2\sum_{\ell=2}^{\infty} \mathrm{Cov}\left(E_{\mathrm{stat}}(1, C_1),\ E_{\mathrm{stat}}(\ell, C_\ell)\right),$$

the quantity

$$\frac{\mathcal{E}_{f,\mathrm{stat}}(X_n) - \mu n}{\sqrt{n}}$$

either converges in distribution to $\mathrm{N}(0, \sigma^2)$, or if $\sigma^2 = 0$, converges in probability to zero.

Proof. The proof follows an idea in [19, proof of Lemma 2.7 and Theorem 1.3]—as mentioned at the start of §7.2 and Remark 1.17—to prove limited dependence between stationary events by swapping them with more local versions. More concretely, we will prove

$$\mathrm{Cov}\left(E_{\mathrm{stat}}(j, C_j),\ E_{\mathrm{stat}}(\ell, C_\ell)\right) \le O\Big(e^{-\rho_{1/2} k}\Big) \quad \text{when } j + 2k < \ell, \tag{7.7}$$

the O constant independent of j, ℓ, k. We will assume throughout that $k \ge 8$, merely to recycle (7.6). Once done, a stationary sequence central limit theorem in [5, Theorem 27.4] either yields the normal limit, or if $\sigma^2 = 0$, yields $\mathrm{Var}\, \mathcal{E}_{f,\mathrm{stat}}(X_n) = o(n)$.

Define local, k-step versions of C_j and $E_{\text{stat}}(j, c)$, in each case only considering $n < k$:

$$C_j^{\text{loc}} \stackrel{\text{def}}{=} 1 \wedge \min_{0<n<k} \left\{ \frac{\exp\left(\mathcal{S}_n^R + n\right) - 1}{n} \right\} \tag{7.8}$$

$$E_{\text{stat}}^{\text{loc}}(j, c) \stackrel{\text{def}}{=} \left\{ \mathcal{S}_n^L \leq -\log\left(1 - nc\right) - n \quad \text{for all} \quad 0 < n < (1/c) \wedge k \right\}. \tag{7.9}$$

First observe that C_j, C_j^{loc} only differ if the minimum is achieved for $n \geq k$, in which case

$$\begin{aligned}
\mathbb{P}\big(C_j \neq C_j^{\text{loc}}\big) &\leq \sum_{n=k}^{\infty} \mathbb{P}\big(\mathcal{S}_n^R < \log(1 + C_j^{\text{loc}} n) - n\big) \\
&\leq \sum_{n=k}^{\infty} \mathbb{P}\big(\mathcal{S}_n^R < \log(1 + n) - n\big) \quad (\text{since } C_j^{\text{loc}} \leq 1) \\
&\leq \sum_{n=k}^{\infty} \mathbb{P}\big(\mathcal{S}_n^R < -\tfrac{1}{2} n\big) \quad (n + 1 > k \geq 8 \text{ and (7.6) above}) \\
&\leq \sum_{n=k}^{\infty} e^{-\rho_{1/2} n} \quad (\text{Lemma 7.6}) \\
&= \frac{e^{-\rho_{1/2} k}}{1 - e^{-\rho_{1/2}}} \\
&= O\Big(e^{-\rho_{1/2} k}\Big).
\end{aligned} \tag{7.10}$$

Note the O constant here does not depend on j or k. Repeating the proof of Lemma 7.7(a), (b), (c), but substituting $n = k - 1$ instead of $j - 1$, it is possible to show with any c

$$\mathbb{P}\Big(E_{\text{stat}}^{\text{loc}}(j, c) \setminus E_{\text{stat}}(j, c)\Big) \leq e^{-\rho_{1/2} k},$$

again since $k \geq 8$. Combining the last two facts, we see

$$\mathbb{P}\Big(E_{\text{stat}}(j, C_j)\Big) = \mathbb{P}\Big(E_{\text{stat}}^{\text{loc}}(j, C_j^{\text{loc}})\Big) + O\Big(e^{-\rho_{1/2} k}\Big). \tag{7.11}$$

Now if $j + 2k < \ell$, the local events $E_{\text{stat}}^{\text{loc}}(j, C_j^{\text{loc}})$ and $E_{\text{stat}}^{\text{loc}}(\ell, C_\ell^{\text{loc}})$ are independent, so

$$\begin{aligned}
&\mathbb{P}\Big(E_{\text{stat}}(j, C_j) \cap E_{\text{stat}}(\ell, C_\ell)\Big) \\
&\quad \leq \mathbb{P}\Big(E_{\text{stat}}^{\text{loc}}(j, C_j^{\text{loc}}) \cap E_{\text{stat}}^{\text{loc}}(\ell, C_\ell^{\text{loc}})\Big) + O\Big(e^{-\rho_{1/2} k}\Big) \\
&\quad = \mathbb{P}\Big(E_{\text{stat}}^{\text{loc}}(j, C_j^{\text{loc}})\Big) \cdot \mathbb{P}\Big(E_{\text{stat}}^{\text{loc}}(\ell, C_\ell^{\text{loc}})\Big) + O\Big(e^{-\rho_{1/2} k}\Big).
\end{aligned}$$

Making the same replacements again,

$$\begin{aligned}
&\mathbb{P}\Big(E_{\text{stat}}(j,C_j)\Big)\cdot\mathbb{P}\Big(E_{\text{stat}}(\ell,C_\ell)\Big)\\
&\qquad=\Big[\mathbb{P}\Big(E^{\text{loc}}_{\text{stat}}(j,C^{\text{loc}}_j)\Big)+O\Big(e^{-\rho_{1/2}k}\Big)\Big]\Big[\mathbb{P}\Big(E^{\text{loc}}_{\text{stat}}(\ell,C^{\text{loc}}_\ell)\Big)+O\Big(e^{-\rho_{1/2}k}\Big)\Big]\\
&\qquad=\mathbb{P}\Big(E^{\text{loc}}_{\text{stat}}(j,C^{\text{loc}}_j)\Big)\cdot\mathbb{P}\Big(E^{\text{loc}}_{\text{stat}}(\ell,C^{\text{loc}}_\ell)\Big)+O\Big(e^{-\rho_{1/2}k}\Big).
\end{aligned}$$

Subtracting from the previous calculation and cancelling the product terms proves (7.7). □

7.3 Computer Simulations

Simulations estimating μ in Lemma 7.8 and Proposition 7.2 suggest neither the 1/4 nor $\log 2 \approx 0.6931$ from Proposition 7.1 is correct. This was simulated as follows, with $N = 1000$:

1. Generate the RHS walk $\mathcal{S}^R_n$ for $0 < n \le N$.
2. Compute $c = C^{\text{loc}}_j$ as in (7.8) with $k = N + 1$. This is the maximal c so that $\mathcal{S}^R_n \ge \log(1 + cn) - n$ for $0 < n \le N$.
3. Generate $\mathcal{S}^L_n$ and determine whether $\mathcal{S}^L_n \le -\log(1 - cn) - n$ for all $0 < n < 1/c$.

Repeating steps 1–3 on $W = 10^7$ walks, the success count is nearly $\text{Bin}(W, \mu)$, so approximate

$$\mu \approx \widehat{\mu} = 0.628414;$$

for reference, three standard deviations of $\widehat{\mu}$ are bounded by

$$\frac{3}{2\sqrt{W}} \approx 0.000474.$$

The true binomial success probability is not exactly μ, since on the RHS we must take $N < \infty$ to simulate. However, $N = 1000$ seems good enough, in that c was usually found within the first ten steps; in other words, the sample distribution of $\operatorname{argmin}_{1\le n\le N}\left\{\left(e^{\mathcal{S}^R_n+n} - 1\right)/n\right\}$ was focused on low values of n, as shown on the right of Figure 7.1. (Only a single walk

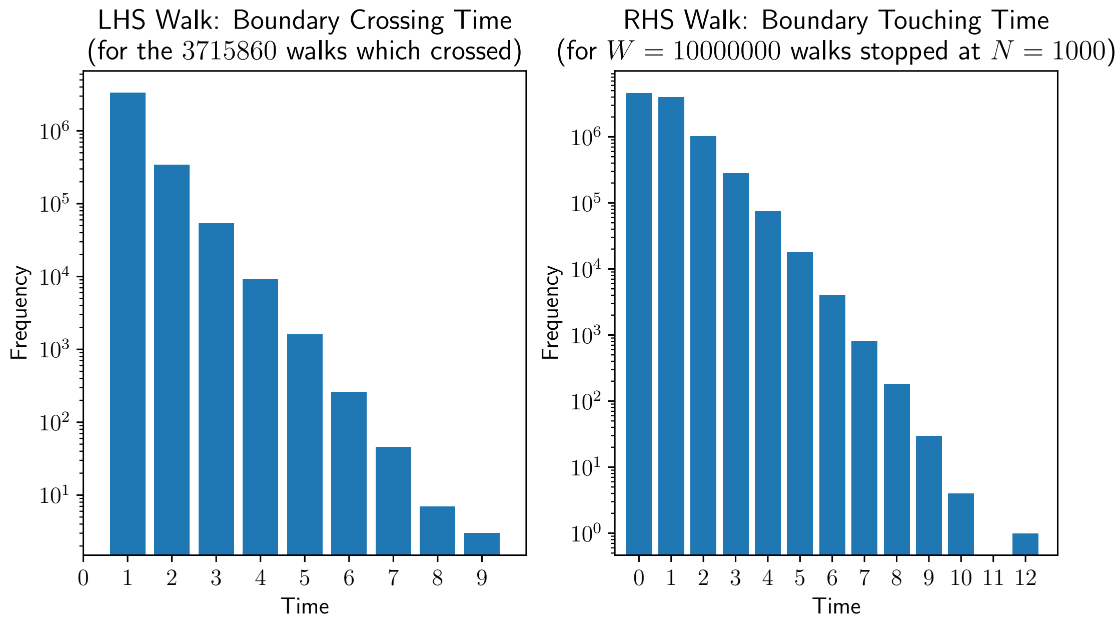

Figure 7.1: Log-scale histograms presenting relevant times from simulations. (Left) For walks that did not succeed in step 3, the first time $n < 1/c$ at which $\mathcal{S}_n^L > -\ln(1-cn)-n$. (Right) For all W walks, the (almost surely unique) time $n \leq N$ at which $\mathcal{S}_n^R = \log(1+cn) - n$.

---

determined c at the 12th step.) On a related note, we see the histograms in Figure 7.1 decay roughly linearly. Since these are log-scale, they capture the exponentials in (7.10) and (7.11).

7.4 Intensities $1/x \log^q x$ Revisited

Proposition 7.9. Fix intensity measure $f(x) = \mathbf{1}[x > e]/x \log^q x$, $0 < q < 1$, $p = 1 - q$.

(a) $\mathbb{P}(X \text{ is extreme} \mid X) \geq 1 - \dfrac{1}{\log^q X}$.

(b) $\mathbb{E}\mathcal{E}_f(t) \sim \log^p t/p$ as $t \to \infty$.

Proof. Recall from Remark 7.4 that we are essentially able to repeat the argument from

Lemma 7.3(a). For (a), since $\beta_* > X$, ignore the LHS and use $\beta = X$ to see

$$\begin{aligned}\mathbb{P}(X \text{ is extreme} \mid X) &\geq \mathbb{P}(\mathcal{S}_n > \lambda(X+nX) - \lambda(X) - n \ \ \forall n > 0 \mid X) \\ &\geq \mathbb{P}(\mathcal{S}_n > n \cdot Xf(X) - n \ \ \forall n > 0 \mid X) \quad \text{(tangent line at } n=0) \\ &= 1 - Xf(X).\end{aligned}$$

For part (b) use (1.6) to write

$$\mathcal{E}_f(t) \geq \int_e^t \left(1 - \frac{1}{\log^q x}\right) \frac{1}{x \log^q x}\, dx = \frac{\log^p t - 1}{p} - \int_e^t \frac{1}{x \log^{2q} x}\, dx,$$

and note the final integral is in any case $o(\log^p t)$:

$$\int_e^t \frac{1}{x \log^{2q} x}\, dx = \begin{cases} O(1) & \text{if } q > 1/2 \\ \log\log t & \text{if } q = 1/2 \\ \dfrac{\log^{p-q} t - 1}{p-q} & \text{if } q < 1/2. \end{cases} \qquad \square$$

Unfortunately, we still do not learn whether (almost surely) there are finitely or infinitely many nonextreme points in the case $q \leq 1/2$. However, we can say the following about the proportion—that is, the ratio with $N(t)$, the total number of points in $[0, t]$. Lemma 2.4 gives $N(t) \leq \frac{1}{p} \log^p t + o\Big(\log^{p(1/2+\delta)} t\Big)$ as $t \to \infty$, so

$$1 \geq \mathbb{E}\left(\frac{\mathcal{E}_f(t)}{N(t)}\right) \gtrsim \frac{\mathbb{E}\mathcal{E}_f(t)}{\frac{1}{p}\log^p t + \log^{p(1/2+\delta)} t} \sim 1,$$

meaning the expected proportion of extreme points converges to 1.

Chapter 8

Expectation Lower Bounds with $1/x^q$ and $1/\log x$

This section proves the following results.

Proposition 8.1. Fix intensity measure $f(x) = \mathbf{1}[x > 0]/x^q$, $0 < q < 1$, $p = 1 - q$.

(a) $\mathbb{P}(X \text{ is extreme} \mid X) \geq \Omega\Big(1/X^{\frac{2}{3}p}\Big)$ as $X \to \infty$.

(b) $\mathbb{E}\mathcal{E}_f(t) \geq \Omega\big(t^{p/3}\big)$ as $t \to \infty$.

Proposition 8.2. Fix intensity measure $f(x) = \mathbf{1}[x > e]/\log x$.

(a) $\mathbb{P}(X \text{ is extreme} \mid X) \geq \Omega\big(1/X^{2/3}\big)$ as $X \to \infty$.

(b) $\mathbb{E}\mathcal{E}_f(t) \geq \Omega\left(\dfrac{t^{1/3}}{\log t}\right)$ as $t \to \infty$.

These lower bounds on $\mathbb{E}\mathcal{E}_f(t)$ can be predicted by the informal calculation in §1.3.5, which already presented most of the technique—boosting, then avoiding a parabola. Let us recap what needs to be done here.

Remark 8.3 (Correcting the informal calculation)**.** The argument in §1.3.5 contained an obvious error. The parabolic approximation ended up on the wrong side of the actual boundary (see Figure 1.10 and Remark 1.13). While the parabola was a good approximation near the maximum at zero, far away from the maximum it overshot the boundary. The way to fix this was already mentioned in Remark 1.5; it can be compared with techniques explicit in [4, Lemmas 1(i), 5, 8] or implicit in [7]. All we need to do is show the walk is unlikely to hit the true boundary far from the maximum, which can be done by swapping out the parabolic approximation for a linear one. In other words, the fix here is twofold: (i) near the maximum, ensure the parabola begins on the correct side of the true boundary, and (ii) far from the maximum, avoid where the parabola overshoots the true boundary by including a third stage of time with a linear boundary. *

We begin with $1/x^q$ in §8.1. After focusing on the RHS of the walk, the LHS is straightforward. The $1/\log x$ case in §8.2 is basically the same, actually a bit easier.

8.1 Lower Bound: $1/x^q$

8.1.1 On the RHS

The event (6.6), that X looks extreme to the right with a line of slope $\beta_* = 1/f(X) = X^q$, is in this case

$$E_{R,X^q} = \left\{ \mathcal{S}_n^R \geq \frac{(X + nX^q)^p}{p} - \frac{X^p}{p} - n \quad \text{for all} \quad n > 0 \right\}.$$

Claim 8.4. $\mathbb{P}(E_{R,X^q} \mid X) \geq \Omega\big(1/X^{p/3}\big)$ as $X \to \infty$.

We will show $\mathcal{S}_n$ can avoid the boundary curve by succeeding in three stages—the first two just like in §1.3.5, but with the corrections in Remark 8.3. The three-stage success is described below and a sample is shown in Figure 8.1. For this description, fix any $0 < \epsilon < 1$.

Stage I. Two things must happen: $\mathcal{S}_n$ must stay nonnegative during $0 \leq n \leq X^{\frac{2}{3}p}$, and the final position must not be too low, specifically $\mathcal{S}_{X^{\frac{2}{3}p}} \geq X^{p/3}$.

Stage II. With $\mathcal{S}_n$ starting at height $X^{p/3}$, again two things must happen: the walk must remain above the red parabola when $X^{\frac{2}{3}p} \leq n \leq \epsilon X^p$, and the final height must be nonnegative, $\mathcal{S}_{\epsilon X^p} \geq 0$.

Stage III. Finally, starting from height $\mathcal{S}_{\epsilon X^p} = 0$, require that $\mathcal{S}_n$ remain above the green line for all $n \geq \epsilon X^p$.

Lower bounds on the chance each stage succeeds are derived below in (8.4), (8.5), and (8.6). Multiplying these three yields Claim 8.4. We first justify that these bounds are on the correct side of the boundary curve and therefore really fix the issue. Note

$$0 < \epsilon < 1 < \frac{3}{2-p} = \frac{-\binom{p}{2}}{\binom{p}{3}}. \tag{8.1}$$

Lemma 8.5. Let $V_U = -\frac{1}{p}\left(\binom{p}{2} + \epsilon\binom{p}{3}\right) > 0$. If $0 \leq n \leq \epsilon X^p$, then

$$\lambda(X + n\beta_*) - \lambda(X) - n \leq -\frac{V_U}{X^p} n^2. \tag{8.2}$$

Proof. First, $V_U > 0$ by (8.1). Let $0 \leq n \leq \epsilon X^p$ and expand the binomial term:

$$\frac{(X + nX^q)^p}{p} - \frac{X^p}{p} - n = \frac{X^p}{p} \sum_{j \geq 2} \binom{p}{j} \left(\frac{n}{X^p}\right)^j.$$

When $j \geq 3$, $\left|\binom{p}{j+1}/\binom{p}{j}\right| = 1/(j+1)(j-1-p) < 1$, so the terms of the series decrease in magnitude for $j \geq 3$. Since the terms alternate and the $j = 4$ term is negative,

$$\frac{(X + nX^q)^p}{p} - \frac{X^p}{p} - n \leq \frac{\binom{p}{2} n^2}{pX^p} + \frac{\binom{p}{3} n^3}{pX^{2p}} = \frac{n^2}{pX^p}\left[\binom{p}{2} + \binom{p}{3}\underbrace{\frac{n}{X^p}}_{\leq \epsilon}\right] \leq \frac{-V_U n^2}{X^p}. \qquad \square$$

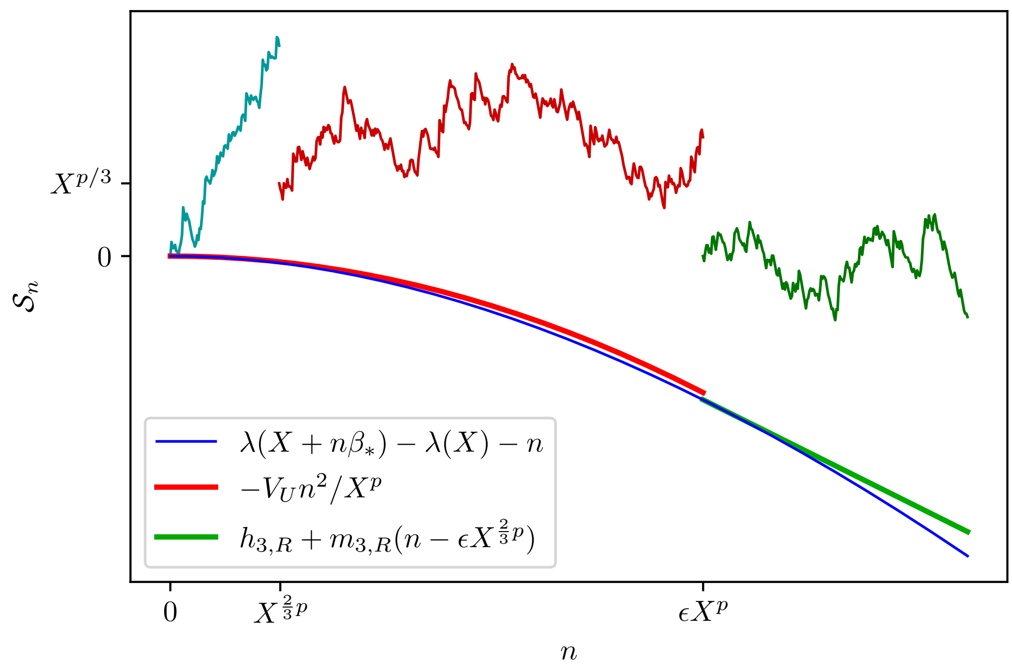

Figure 8.1: Bounds (8.2) and (8.3) used to establish Claim 8.4, as well as three paths that succeed during Stages I, II, III.

Lemma 8.6. When $n \geq \epsilon X^p$,

$$\lambda(X + n\beta_*) - \lambda(X) - n \leq h_{3,R} + m_{3,R}(n - \epsilon X^p), \tag{8.3}$$

where

$$m_{3,R} = \frac{1}{(1+\epsilon)^q} - 1 \quad \text{and} \quad h_{3,R} = \lambda(X + \epsilon X) - \lambda(X) - \epsilon X^p.$$

Proof. At time $n = \epsilon X^p$, the true boundary (in blue) is at height $h_{3,R}$, and it has slope

$$\beta_* f(X + n\beta_*) - 1 = \frac{X^q}{(X + \epsilon X)^q} - 1 = m_{3,R}.$$

After this point one can also check the second derivative, $\beta_*^2 f'(X + n\beta_*)$, is negative. $\square$

We now compute the chance of success during each stage.

Stage I. Just as in the informal calculation, revisit Lemma 2.6 to see

$$\begin{aligned}\mathbb{P}(\text{Stage I succeeds}) &= \mathbb{P}\Big(\mathcal{S}_n \geq 0 \text{ when } 0 \leq n \leq X^{\frac{2}{3}p} \text{ and } \mathcal{S}_{X^{\frac{2}{3}p}} \geq X^{p/3}\Big) \\ &\geq \Omega\Big(1/X^{p/3}\Big).\end{aligned} \tag{8.4}$$

Stage II. This is the stage with a parabolic rescaling (Remark 1.5). Setting

$$N = \left(\frac{X^p}{2V_U}\right)^{2/3}$$

like in (1.12), only with a different constant, gives

$$\begin{aligned}&\mathbb{P}\Big(\mathcal{S}_n \geq -\frac{V_U}{X^p}n^2 \quad \forall\, X^{\frac{2}{3}p} \leq n \leq \epsilon X^p \;\Big|\; \mathcal{S}_{X^{\frac{2}{3}p}} = X^{p/3}\Big) \\ &= \mathbb{P}\left(\frac{\mathcal{S}_{Nt}}{\sqrt{N}} + \frac{X^{p/3}}{\sqrt{N}} \geq -\frac{V_U}{X^p}N^{3/2}t^2 \quad \forall\, X^{\frac{2}{3}p} \leq n = Nt \leq \epsilon X^p \;\Big|\; \mathcal{S}_{X^{\frac{2}{3}p}} = 0\right) \\ &\gtrsim \mathbb{P}\Big(B_t + (2V_U)^{1/3} > -\frac{1}{2}t^2 \quad \forall\, t > 0 \;\Big|\; B_0 = 0\Big) \quad \text{(Donsker, and requiring all } t > 0) \\ &\geq \Omega(1),\end{aligned}$$

on the last line arguing like in (1.14) and referring to Groeneboom [18, Corollary 2.1 and Figure 2] or to Janson, Louchard, and Martin-Löf [25, Theorem 2.4 and Figure 2]. Stage II also requires a nonnegative ending height. By the central limit theorem,

$$\mathbb{P}\Big(\mathcal{S}_{\epsilon X^p} \geq 0 \;\Big|\; \mathcal{S}_{X^{\frac{2}{3}p}} = X^{p/3}\Big) \gtrsim \frac{1}{2}.$$

Multiplying the last two probabilities with Lemma 2.7 gives

$$\mathbb{P}\Big(\text{Stage II succeeds} \;\Big|\; \text{Stage I succeeds}\Big) \geq \Omega(1). \tag{8.5}$$

Stage III. By (8.3),

$$\begin{aligned}&\mathbb{P}\Big(\text{Stage III succeeds} \;\Big|\; \text{Stages I, II both succeed}\Big) \\ &\quad\geq \mathbb{P}\Big(\mathcal{S}_n \geq h_{3,R} + m_{3,R}(n - \epsilon X^p) \quad \forall\, n \geq \epsilon X^p \;\Big|\; S_{\epsilon X^p} = 0\Big) \to 1,\end{aligned} \tag{8.6}$$

the convergence to 1 by Lemma 2.8 with $a = -h_{3,R} = \Theta(X^p)$ and $b = -m_{3,R} = \Theta(1)$.

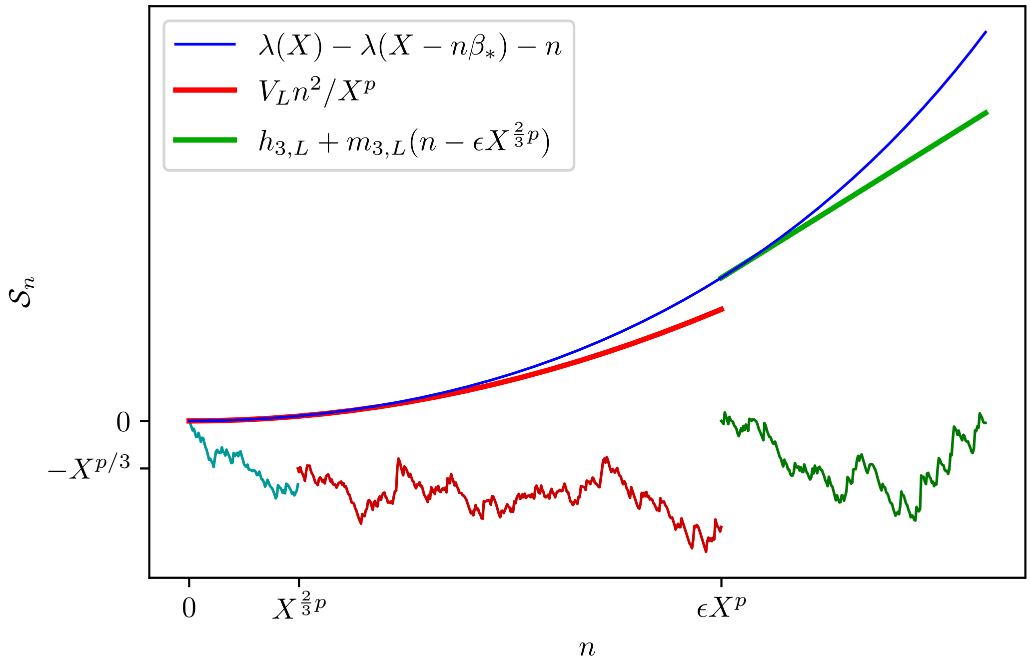

Figure 8.2: Bounds used to establish Claim 8.9.

8.1.2 On the LHS

We now switch to the event (6.7), but apart from dealing with the index $n_X = Xf(X) = X^p$ (Remark 6.3), very little is different. The three-stage procedure is exactly the same as on the RHS, only the inequalities are flipped positive to negative (Figure 8.2).

Stage I. Require $\mathcal{S}_n \le 0$ during $0 \le n \le X^{\frac{2}{3}p}$ with final height $\mathcal{S}_{X^{\frac{2}{3}p}} \le -X^{p/3}$.

Stage II. Continuing from height $-X^{p/3}$, require that $\mathcal{S}_n$ remain below the red parabola while $X^{\frac{2}{3}p} \le n \le \epsilon X^p$ and end at a height $\mathcal{S}_{\epsilon X^p} \le 0$.

Stage III. Beginning at height zero, require that $\mathcal{S}_n$ remain below the green line when $n > \epsilon X^p$. (We only need $n < n_X$, but for a lower bound this is unimportant.)

The bounds in Figure 8.2 are stated below, but the proofs are omitted; they are similar to those of (8.2) and (8.3).

Lemma 8.7. Let $V_L = q/2$. If $0 \le n \le X^p$,

$$\lambda(X) - \lambda(X - n\beta_*) - n \ge \frac{V_L}{X^p} n^2.$$

Lemma 8.8. If $n \ge \epsilon X^p$, then

$$\lambda(X) - \lambda(X - n\beta_*) - n \ge h_{3,L} + m_{3,L}(n - \epsilon X^p),$$

where

$$m_{3,L} = \frac{1}{(1-\epsilon)^q} - 1 \qquad \text{and} \qquad h_{3,L} = \lambda(X) - \lambda(X - \epsilon X) - \epsilon X^p.$$

Calculations just like (8.4), (8.5), and (8.6) above establish:

Claim 8.9. $\mathbb{P}(E_{L,X^q} \mid X) = \Omega\big(1/X^{p/3}\big)$ as $X \to \infty$.

Proof of Proposition 8.1. For part (a), use (6.8) and multiply the probabilities in Claims 8.4 and 8.9. Part (b) is then from (a) and (1.6). □

8.2 Lower Bound: $1/\log x$

The strategy for this case is very similar to the one for $1/x^q$. We consider (6.6)–(6.7) with $\beta_* = \log X$ and $n_X = X/\log X$ as in Remark 6.3. Here are the relevant bounds for the RHS. (Proofs are delayed briefly.)

Lemma 8.10. Let $V_U = \frac{\log 2}{2\log(2e)}$. When $0 < n \le \frac{X}{\log X}$,

$$\lambda(X + n\log X) - \lambda(X) - n \le -\frac{V_U n^2}{X}.$$

Lemma 8.11. When $n \ge \frac{X}{\log X}$,

$$\lambda(X + n\log X) - \lambda(X) - n \le h_{3,R} + m_{3,R}\left(n - \frac{X}{\log X}\right),$$

where

$$m_{3,R} = \frac{\log X}{\log(2X)} - 1 \quad \text{and} \quad h_{3,R} = \lambda(2X) - \lambda(X) - \frac{X}{\log X}.$$

Note here that $m_{3,R} = -\Theta(1/\log X) \to 0$ is not order 1, as it was with intensity $1/x^q$. However, by Lemma 8.10 we also have $h_{3,R} \leq -\frac{V_U X}{\log^2 X}$. This means $|h_{3,R} \cdot m_{3,R}| \to \infty$, so we can still apply Lemma 2.8. Repeating what we did with $1/x^q$:

Stage I. Require $\mathcal{S}_n \geq 0$ during $0 \leq n \leq X^{2/3}$ with ending height $\mathcal{S}_{X^{2/3}} \geq X^{1/3}$.

Stage II. Continuing from height $X^{1/3}$, require that $\mathcal{S}_n$ remain above the parabola given in Lemma 8.10 during $X^{2/3} \leq n \leq \frac{X}{\log X}$. Also require $\mathcal{S}_{X/\log X} \geq 0$.

Stage III. Starting from height zero, require that $\mathcal{S}_n$ remain above the line in Lemma 8.11 for $n \geq \frac{X}{\log X}$.

The LHS is nearly the same as above with the expected reflection. However, since $n_X = X/\beta_* = X/\log X$ steps are already covered by Stage II, the walk may end there. The only bound needed for the LHS is as follows.

Lemma 8.12. Let $V_L = 1/2$. If $0 < n \leq \frac{X}{\log X}$,

$$\lambda(X) - \lambda(X - n\log X) - n \geq \frac{V_L}{X} n^2$$

Proof of Proposition 8.2. Part (a) is as above for $1/x^q$, computing the chance the three stages succeed on each side. For part (b), use (1.6) to say

$$\mathbb{E}\mathcal{E}_f(t) \geq \int_e^t \frac{1}{\log x} \cdot \Omega\left(\frac{1}{x^{2/3}}\right) dx \geq \frac{1}{\log t} \int_e^t \Omega\left(\frac{1}{x^{2/3}}\right) dx \geq \Omega\left(\frac{t^{1/3}}{\log t}\right). \qquad \square$$

We now return to the boundary calculations above. Lemma 8.11 may be checked like (8.3) and so its proof is omitted. The remaining two are triangular area calculations.

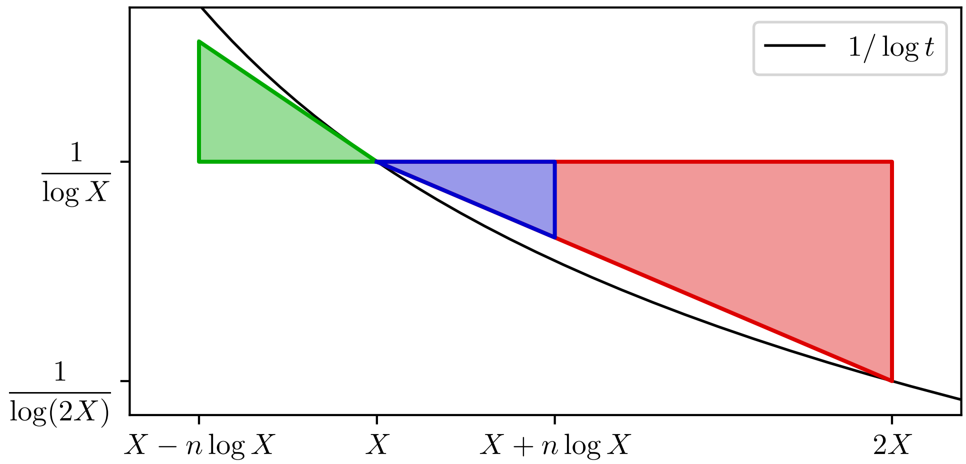

Figure 8.3: Triangular area approximations. The blue region is scaled down by a factor $n \log X/X$ from the red triangle, which has two corners along the curve $1/\log t$. The upper boundary to the green triangle is the tangent line to $1/\log t$ at $t = X$.

Proof of Lemma 8.10. Computing the area of the blue triangle in Figure 8.3 at $\overset{\triangle}{<}$,

$$\lambda(X + n\log X) - \lambda(X) - n = \int_X^{X+n\log X} \frac{1}{\log t} - \frac{1}{\log X}\, dt$$

$$\overset{\triangle}{<} -\frac{1}{2}\overbrace{(n\log X)}^{\text{width}}\overbrace{\left(\frac{n\log X}{X}\right)\cdot\left(\frac{1}{\log X} - \frac{1}{\log(2X)}\right)}^{\text{height}} = -\overbrace{\left(\frac{\log X \cdot \log 2}{2\log(2X)}\right)}^{(*)}\cdot\frac{n^2}{X}.$$

The coefficient $(*)$ is increasing, because $\frac{d}{dx}\left[\frac{\log x}{\log(2x)}\right] = \frac{\log 2}{x\log^2(2x)} > 0$, so we can replace it by $V_U = \frac{\log e \cdot \log 2}{2\log(2e)}$. □

Proof of Lemma 8.12. The green triangular area in Figure 8.3 gives

$$\lambda(X) - \lambda(X - n\log X) - n$$
$$= \int_{X-n\log X}^{X} \frac{1}{\log t} - \frac{1}{\log X}\, dt \overset{\triangle}{>} \frac{1}{2}(n\log X)^2\left(\frac{1}{X\log^2 X}\right) = \frac{n^2}{2X}.$$ □

Chapter 9

Expectation Upper Bounds with $1/x^q$ and $1/\log x$

This section proves the remaining upper bounds for Table 1.2. As usual, in each of the next results, (b) follows from (a) via (1.6). Note the upper bounds match the lower bounds in §8 apart from hidden factors of $\log X$ in (a) or $\log t$ in (b).

Proposition 9.1. Fix intensity measure $f(x) = \mathbf{1}[x > 0]/x^q$ with $0 < q < 1$, $p = 1 - q$.

(a) $\mathbb{P}(X \text{ is extreme} \mid X) \leq \widetilde{O}\Big(1/X^{\frac{2}{3}p}\Big)$ as $X \to \infty$.

(b) $\mathbb{E}\mathcal{E}_f(t) \leq \widetilde{O}\big(t^{p/3}\big)$ as $t \to \infty$.

Proposition 9.2. With the intensity measure $f(x) = \mathbf{1}[x > e]/\log x$,

(a) $\mathbb{P}(X \text{ is extreme} \mid X) \leq \widetilde{O}\big(1/X^{2/3}\big)$ as $X \to \infty$.

(b) $\mathbb{E}\mathcal{E}_f(t) \leq \widetilde{O}\big(t^{1/3}\big)$ as $t \to \infty$.

9.1 Proof of Proposition 9.1(a)

The upper bound, though more involved than the lower bound, follows the relatively straightforward idea suggested by Robin Pemantle, outlined briefly in §1.3.4(c)(ii). Recall this argument came down to selecting zones (1.4) to evaluate (1.5). To avoid losing the key idea amid technicalities, this section gives a detailed proof sketch using certain unjustified equations, lemma statements, and incomplete definitions. For a complete argument, there are mid-proof pointers to the omitted details.

Here is how to prove the proposition. Since $\beta_* = X^q$ was a convenient slope choice, parametrize our slope zones as $\beta_j = \alpha_j X^q$, where

$$0 = \alpha_{I_0} < \cdots < \alpha_{-2} < \alpha_{-1} < \alpha_0 < \alpha_1 < \alpha_2 < \cdots < \alpha_{I_1} = \infty, \tag{9.1}$$

is some sequence of coefficients we will choose later. Our first step is to discard the ends of (9.1). Specifically, if either $|\alpha_j - 1| > 1/2$ or $|\alpha_{j+1} - 1| > 1/2$, then one of the two probabilities shown in (1.5) is small. More formally:

Lemma 9.3. There are constants $c, \nu > 0$ so that

$$\mathbb{P}(X \text{ is extreme} \mid X) \leq \mathbb{P}\Big(\bigcup\nolimits_{|\alpha-1|\leq 1/2} E_{L,\alpha X^q} \cap E_{R,\alpha X^q} \;\Big|\; X\Big) + O\big(e^{-cX^\nu}\big). \tag{9.2}$$

Remark 1.16 mentioned that Lemma 9.3 relates to Theorem 1.15 by Prakasa Rao [34, Lemma 4.1]; the proofs both rely on the same underlying idea—that a distant point is unlikely to overcome the wrong slope. With that said, here is a two-sentence proof of Lemma 9.3. Figure 1.6 shows $\alpha = 1$ in blue and shows how mild adjustments to a support line slope will, after transformation, shift and raise the boundary curve. This effect becomes stronger as $X \to \infty$, so if $\alpha \not\approx 1$, then either on the left or on the right our mean-zero walk must climb over a steep hill, and this is unlikely. The full proof (in §9.2) is only more involved because we need careful estimates for the true boundary curves.

We now aim to determine a sequence

$$\frac{1}{2} = \alpha_0 < \alpha_1 < \alpha_2 < \cdots < \alpha_{J-1} < \alpha_J = \frac{3}{2}. \tag{9.3}$$

That is, in (9.1) it will suffice to set $\alpha_{-1} = 0$, $\alpha_{J+1} = \infty$ and $I_0 = -1$, $I_1 = J+1$. Although parametrizing slopes as αX^q is initially convenient, as we approach $\alpha = 1$ there is a better scale. We will describe our choice (9.3) by determining a sequence

$$-\frac{X^{p/3}}{(2q)^{1/3}\log X} = \kappa_0 < \kappa_1 < \cdots < \kappa_{J-1} < \kappa_J = \frac{X^{p/3}}{(18q)^{1/3}\log X} \tag{9.4}$$

and declaring

$$\kappa_j = \left(\frac{\alpha_j - 1}{\alpha_j^{2/3}}\right) \cdot \frac{X^{p/3}}{q^{1/3}\log X}. \tag{9.5}$$

Note this is indeed a reparametrization: $(\alpha - 1)/\alpha^{2/3}$ strictly increases when $0 < \alpha < \infty$. (Again, this is the slope rescaling we compared with [19, Lemma 3.1] in Remark 1.16.)

The choice (9.5) will seem more natural and convenient once we make (1.5) precise and set up the Brownian motion approximation. Unlike the lower bound, where we used Donsker's Theorem, here we need a Brownian motion coupling derived from Komlós, Major, and Tusnády [27]. Extending (9.2) with the idea in (1.5) and this coupling will lead to

$$\begin{aligned} &\mathbb{P}(X \text{ is extreme} \mid X) \\ &\quad \le O\big(e^{-cX^\nu}\big) + O\left(\frac{1}{X^{\frac{2}{3}p}}\right) + \sum_{j=0}^{J-1} D(-\kappa_j, H_R, T) D(\zeta\kappa_{j+1}, H_L, T), \end{aligned} \tag{9.6}$$

which requires significant explanation.

- The initial O probability loss is from (9.2), to discard $|\alpha - 1| > 1/2$.

- We incur the second O because we use the Komlós, Major, and Tusnády [27] coupling. Loosely stated, their result couples a walk and Brownian motion so that the worst-case vertical difference, $\max_t \left| \frac{\mathcal{S}_{Nt}}{\sqrt{N}} - B_t \right| \leq H$, is very likely to be small. For the RHS and LHS we will call these worst-case distances H_R and H_L; both are $\Theta(\log X / X^{p/3})$. The coupling will be run during times $0 \leq t \leq T = \log X$.
- The function $D(\kappa, H, T)$ will be a probability defined in terms of the coupled Brownian motion. When this Brownian motion coupling is good,
  - the event defining $D(-\kappa_j, H_R, T)$ will cover the RHS event $E_{R,\alpha_j X^q}$, and
  - the event defining $D(\zeta\kappa_{j+1}, H_L, T)$ will cover the LHS event $E_{L,\alpha_{j+1} X^q}$.

  More precisely, the function $D(\kappa, H, T)$ represents the probability that Brownian motion remains for time $T = \log X$ above a parabola if started H units above it. Here κ controls where along the parabola the Brownian motion starts. (For intuition, suppose Figure 1.6 showed parabolas with different κ values.) The probabilities represented by $D(\kappa, H, T)$ will be estimated using techniques from Groeneboom's paper [17].
- The factor $\zeta \stackrel{\text{def}}{=} 1 - \frac{1}{\log^4 X} = 1 - \frac{1}{T^4}$ allows the LHS and RHS to use different parabolic bounds for the curves in Figure 1.6. (The reason for these to differ is the same as in §8, Remark 1.13, and Remark 8.3, namely, to keep the parabolic bounds on the appropriate side of the boundary.)

See §9.3 for all the missing definitions, as well as a proof of (9.6).

We will prove (in §9.4) the following bound on $D(\kappa, H, T)$ by repeating detailed calculations in Groeneboom's paper [17, §5] with only minor adjustments:

Lemma 9.4. Let $T = \log X$ and $H = O\big(\log X / X^{\Xi}\big)$ for some fixed exponent $0 < \Xi \leq 1/3$.

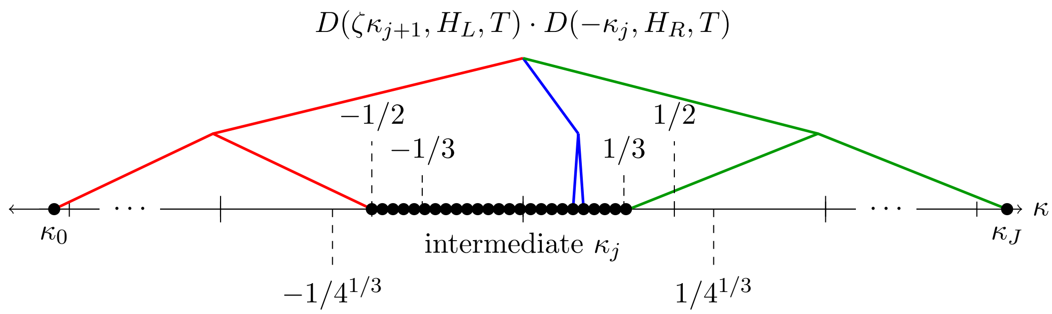

Figure 9.1: The sequence of κ_j values.

As $X \to \infty$,

$$D(\kappa, H, T) \leq \begin{cases} \widetilde{O}\Big(He^{\kappa^3 T^3/6}\Big) & \text{if } |\kappa| \leq 1/4^{1/3} \\ \\ 1 & \text{always.} \end{cases} \tag{9.7}$$

Note 9.5. We use $\Xi = p/3$ now, with $f(x) = 1/x^q$, and $\Xi = 1/3$ later, with $f(x) = 1/\log x$.

The benefit of Lemma 9.4 is the factor H in (9.7). If we pop factors $H_R \cdot H_L$ out of the sum (9.6), and if we can control the remaining sum, we will be done: remember part (a) is to show $\mathbb{P}(X \text{ is extreme} \mid X) \leq \widetilde{O}(X^{-\frac{2}{3}p}) = \widetilde{O}(H_R \cdot H_L)$, and the first two O terms in (9.6) are small enough already. To control the sum and only gain factors of $T = \log X$ (which absorb into the $\widetilde{O}$), we choose the slope zones carefully. The basic idea of these zones is sketched in Figure 9.1, which has dots along the κ axis representing values κ_j we should select for our sequence. More formally (in §9.5) we will show:

Lemma 9.6. With $\zeta = 1 - \frac{1}{T^4}$, we can define the sequence $\kappa_1, \ldots, \kappa_{J-1}$ so that:

(a) κ_j is a strictly increasing sequence,

(b) $-1/2 = \kappa_1 < \zeta\kappa_1 < -1/3$,

(c) $1/3 < \kappa_{J-1} \leq 1/2$,

(d) the intermediate terms are close enough that

$$\zeta\kappa_{j+1} - \kappa_j \le \frac{1}{T^3} \qquad \text{when} \qquad 1 \le j \le J-2, \tag{9.8}$$

(e) the number of terms in the sequence satisfies $J \le O(T^4)$.

Using these conditions we now check (9.6) has the desired order by looking at the sum term by term. First, we show the $j = 0$ and $j = J - 1$ terms are both smaller than required.

- In the term $j = 0$ (red pair in Figure 9.1), look at the factor $D(\zeta\kappa_1, H_L, T)$. Since $-1/4^{1/3} < -1/2 < \zeta\kappa_1 < -1/3$, by (9.7),

  $$D(\zeta\kappa_1, H_L, T) \le \widetilde{O}\Big(H_L e^{-T^3/6\cdot 3^3}\Big),$$

  which is exponentially small. We may simply bound $D(-\kappa_0, H_R, T) \le 1$.

- When $j = J - 1$ (in green), then we do the same, vice versa. We can bound the LHS factor by $D(\zeta\kappa_J, H_L, T) \le 1$ and focus on the RHS: since $-1/2 \le -\kappa_{J-1} < -1/3$,

  $$D(\kappa_{J-1}, H_R, T) \le \widetilde{O}\Big(H_R e^{-T^3/6\cdot 3^3}\Big).$$

Again, these estimates show the $j = 0$ and $j = J - 1$ terms are exponentially smaller than required.

For the intermediate terms, $1 \le j \le J - 2$ (in blue), remember that (9.6) has the RHS and LHS κ values move across the diagram in tandem, always in pairs κ_j, κ_{j+1}. These pairs of κ are close enough to satisfy (9.8), so by the identity $a^3 - b^3 = (a-b)(a^2 + ab + b^2)$,

$$(\zeta\kappa_{j+1})^3 - \kappa_j^3 = (\zeta\kappa_{j+1} - \kappa_j) \cdot \underbrace{\big((\zeta\kappa_{j+1})^2 + \zeta\kappa_j\kappa_{j+1} + \kappa_j^2\big)}_{O(1)} \le O\big(1/T^3\big). \tag{9.9}$$

Here one can check both $|\zeta\kappa_{j+1}| \le 1/4^{1/3}$ and $|-\kappa_j| \le 1/4^{1/3}$, so we can apply (9.7) to both sides. Then using (9.9) gives

$$\begin{aligned}
&D(-\kappa_j, H_R, T) \cdot D(\zeta\kappa_{j+1}, H_L, T) \\
&\qquad \le \widetilde{O}\left[H_R \cdot H_L \cdot \exp\left(\frac{1}{6}T^3\left[(\zeta\kappa_{j+1})^3 - \kappa_j^3\right]\right)\right] \\
&\qquad \le \widetilde{O}\left(H_R \cdot H_L \cdot e^{O(1)}\right) \\
&\qquad = \widetilde{O}\left(X^{-\frac{2}{3}p}\right).
\end{aligned}$$

To conclude, Lemma 9.6 guarantees the total number of these intermediate terms in the sum (9.6) is fewer than order T^4, which is a factor absorbed by the $\widetilde{O}$. This proves the proposition (after filling the gaps).

9.2 Restriction to $\alpha \approx 1$

Recall (6.6)–(6.7), the events that X looks extreme to the right or to the left with a given slope β. In this case

$$\begin{aligned}
E_{R,\beta} &\stackrel{\text{def}}{=} \left\{\mathcal{S}_n^R \ge \frac{(X+n\beta)^p}{p} - \frac{X^p}{p} - n \quad \text{when} \quad n > 0\right\} \\
E_{L,\beta} &\stackrel{\text{def}}{=} \left\{\mathcal{S}_n^L \le \frac{X^p}{p} - \frac{(X-n\beta)^p}{p} - n \quad \text{when} \quad 0 < n < n_X\right\}.
\end{aligned}$$

Here we take $\beta = \alpha X^q$ and $n_X = \epsilon X^p$ with ϵ small. Remembering Remark 6.3 and calculating $X/\beta = X^p/\alpha$ and $\lambda(X) = X^p/p$, this is valid so long as

$$\epsilon < \min\{1/\alpha,\ 1/p\}. \tag{9.10}$$

We begin by stating more parabolic bounds. Proofs of (9.11) and (9.12) below are omitted, since they are just like for (8.2), with binomial expansion.

Lemma 9.7. If $0 \le n \le X^p$,

$$\frac{(X+n\alpha X^q)^p}{p} - \frac{X^p}{p} - n \ge -\frac{q\alpha^2}{2X^p}n^2 + (\alpha-1)n. \tag{9.11}$$

Lemma 9.8. If $0 \leq n \leq \epsilon X^p$,

$$\frac{X^p}{p} - \frac{(X - n\alpha X^q)^p}{p} - n \leq \frac{\alpha^2 C_{\epsilon,\alpha}}{X^p} n^2 + (\alpha - 1)n \tag{9.12}$$

where

$$C_{\epsilon,\alpha} = \frac{1}{p} \sum_{j=2}^{\infty} \left| \binom{p}{j} \right| (\epsilon\alpha)^{j-2}. \tag{9.13}$$

Proof of Lemma 9.3. We are going to use the parabolic bounds to show that if the two-sided random walk avoids a line whose slope has $|\alpha - 1| > \eta$, then on one of the two sides, an event of exponentially small probability has occurred. The lemma is just the case $\eta \stackrel{\text{def}}{=} 1/2$. Specifically, we use (2.2) with a few choices of $a, b > 0$ to say

$$\mathbb{P}(|\mathcal{S}_{aX^p}| > bX^p) = \mathbb{P}\left(\left| \frac{\mathcal{S}_{aX^p}}{\sqrt{aX^p}} \right| > \frac{bX^{p/2}}{\sqrt{a}} \gg X^{p/6} \right) \leq O\big(e^{-c_1 X^{c_2}}\big).$$

First suppose $\alpha \geq 1 + \eta$ and set $n = \delta X^p$ with $\delta \leq 1$ to be chosen. By (9.11), the walk on the RHS is required to satisfy

$$\mathcal{S}_n^R \geq -\frac{q\alpha^2}{2X^p} n^2 + (\alpha - 1)n = \left(-\frac{q\alpha^2\delta^2}{2} + (\alpha - 1)\delta \right) X^p. \tag{9.14}$$

Now optimize in $0 \leq \delta \leq 1$:

- If $\frac{\alpha-1}{q\alpha^2} \geq 1$, set $\delta = 1$ so that (9.14) becomes

$$\mathcal{S}_n^R \geq \left(-\frac{q\alpha^2}{2} + (\alpha - 1) \right) X^p \geq \frac{\alpha - 1}{2} X^p \geq \frac{\eta}{2} X^p = \frac{1}{4} X^p. \tag{9.15}$$

- If $\frac{\alpha-1}{q\alpha^2} \leq 1$, set $\delta = \frac{\alpha-1}{q\alpha^2}$ so that (9.14) requires

$$\mathcal{S}_n^R \geq \frac{(\alpha - 1)^2}{2q\alpha^2} X^p.$$

Since $(\alpha - 1)^2/\alpha^2$ is minimized among $\alpha \geq 1 + \eta$ when $\alpha = 1 + \eta$,

$$\mathcal{S}_n^R \geq \frac{(\alpha - 1)^2}{2q\alpha^2} X^p \geq \frac{\eta^2}{2q(1 + \eta)^2} X^p = \frac{1}{18q} X^p. \tag{9.16}$$

One of these two cases holds, so either (9.15) or (9.16) says $\mathcal{S}_n^R \geq \Omega(X^p)$ in $n = \delta X^p \leq X^p$ steps. The initial discussion now applies.

If $\alpha \leq 1 - \eta$, the argument is similar, but we use the LHS. We must be careful with ϵ and $C_{\epsilon,\alpha}$, so we set this up. Since $1/\alpha > 1$, we may temporarily choose $\epsilon \stackrel{\text{def}}{=} 1/2 < 1$ to satisfy (9.10). Substituting $n = \delta X^p$ in (9.12) with $0 < \delta \leq \epsilon$, we require

$$\mathcal{S}_n^L \leq \big(\alpha^2 C_{1/2,\,1}\delta^2 - (1-\alpha)\delta\big)X^p$$

by using $C_{\epsilon,\alpha} < C_{1/2,\,1}$. The rest of the argument is similar to the one above on the RHS. $\square$

9.3 Technical Setup, Definitions, & Proof of (9.6)

9.3.1 Technical Setup

In light of Lemma 9.3 we now consider only $1/2 \leq \alpha \leq 3/2$. For the remainder of the upper bound proof, when we use the LHS bound (9.12) we let $\epsilon \to 0$ in a way depending on X. Notice (9.10) is now automatically valid, though this is not the reason to have $\epsilon \to 0$. The choice of exactly how ϵ varies will seem unusual until we construct the sequence of κ values at the end, where we will want ϵ, ζ as in (9.18) below.

In (9.13), use $\alpha \leq \frac{3}{2}$, $\epsilon \to 0$, and $\frac{1}{p}\left|\binom{p}{j}\right| = \frac{(1-p)(2-p)\cdots(j-1-p)}{j!} \leq \frac{q}{2}$ when $j \geq 2$ to say

$$C_{\epsilon,\alpha} \leq \frac{q}{2}\sum_{j=2}^{\infty}\left(\frac{3}{2}\cdot\epsilon\right)^{j-2} = \frac{q/2}{1-\frac{3}{2}\epsilon} \stackrel{\text{def}}{=} \frac{1}{2}C_\epsilon. \tag{9.17}$$

Here C_ϵ depends on ϵ, hence on X, but not dramatically: $C_\epsilon \to q$. From this definition,

$$1 \geq \frac{q}{C_\epsilon} = 1 - \frac{3}{2}\epsilon.$$

Now suppose $3\delta \sim \frac{3}{2}\epsilon$ with $\delta = 1/\log^4 X$. Comparing to the expansion $(1-\delta)^3 = 1 - 3\delta +$

$O(\delta^2)$ as $\delta \to 0$, we may define ϵ in such a way to have the key equality

$$\zeta \stackrel{\text{def}}{=} \left(\frac{q}{C_\epsilon}\right)^{1/3} = 1 - \frac{1}{\log^4 X}, \quad \text{meaning} \quad \epsilon \sim \frac{2}{\log^4 X}. \tag{9.18}$$

Before moving on, we rescale (9.11) and (9.12) the same way we did (1.12). Notice (9.20) below merely restates (9.12) including $C_{\epsilon,\alpha} \le \frac{1}{2}C_\epsilon$. To obtain (9.19) and (9.21), complete the square and change notation.

Lemma 9.9. If $0 \le n \le X^p$,

$$\frac{1}{\sqrt{N}}\left(\frac{(X + n\alpha X^q)^p}{p} - \frac{X^p}{p} - n\right) \ge -\frac{1}{2}(t-\theta)^2 + \frac{\theta^2}{2} \tag{9.19}$$

where $N = (X^p/q\alpha^2)^{2/3}$, $n = Nt$, and $\theta = X^{\frac{1}{3}p}(\alpha - 1)/(q\alpha^2)^{1/3}$.

Lemma 9.10. If $|\alpha - 1| \le 1/2$, ϵ is as in (9.18), C_ϵ is as in (9.17), and $0 \le n \le \epsilon X^p$, then

$$\frac{X^p}{p} - \frac{(X - n\alpha X^q)^p}{p} - n \le \frac{\alpha^2 C_\epsilon}{2X^p}n^2 + (\alpha - 1)n. \tag{9.20}$$

Also in this case,

$$\frac{1}{\sqrt{N}}\left(\frac{X^p}{p} - \frac{(X - n\alpha X^q)^p}{p} - n\right) \le \frac{1}{2}(t+\theta)^2 - \frac{\theta^2}{2} \tag{9.21}$$

where $N = (X^p/\alpha^2 C_\epsilon)^{2/3}$, $n = Nt$, and $\theta = X^{\frac{1}{3}p}(\alpha - 1)/(\alpha^2 C_\epsilon)^{1/3}$.

9.3.2 Coupling and Proof of (9.6)

Recall from above $T = \log X$ and the values N_*, θ_* used with the RHS in (9.19) and with the LHS in (9.21):

$$N_L = \left(\frac{X^p}{C_\epsilon \alpha_{j+1}^2}\right)^{2/3} \qquad N_R = \left(\frac{X^p}{q\alpha_j^2}\right)^{2/3}$$

$$\theta_L = \frac{X^{p/3}(\alpha_{j+1} - 1)}{(\alpha_{j+1}^2 C_\epsilon)^{1/3}} \qquad \theta_R = \frac{X^{p/3}(\alpha_j - 1)}{(\alpha_j^2 q)^{1/3}}.$$

The indices of α on each side differ because the LHS and RHS walk through the list of adjacent α_j in tandem. It is now convenient to recall (9.5) and the definition of ζ from (9.18), to say

$$\theta_L = \zeta\kappa_{j+1}T \qquad\qquad \theta_R = \kappa_j T.$$

Note 9.11. The constants distinguishing N_L, N_R are unimportant within a O: remember that $C_\epsilon \to q$ and $1/2 \leq \alpha_j \leq 3/2$, so each $N_* = \Theta(X^{\frac{2}{3}p})$.

We now need to couple the random walk to Brownian motion. To do so, we quote the following theorem, which is a special case of the Komlós, Major, Tusnády coupling [27, Theorem 1]:

Theorem 9.12. Suppose $\mathcal{S}_n$ is a random walk of i.i.d. $\mathrm{Exp}(1)-1$ steps and B_t is a standard Brownian motion. There are constants $C_1, C_2, C_3 > 0$ so that for any N, T, x,

$$\mathbb{P}\left(\max_{n=0,1,\ldots,NT}\left|\frac{\mathcal{S}_n}{\sqrt{N}} - B_{n/N}\right| \geq \frac{C_1\log(NT)+x}{\sqrt{N}}\right) \leq C_2\exp(-C_3 x). \tag{9.22}$$

We need two minor modifications: (i) double the result to a two-sided walk and (ii) keep the walk and Brownain motion close at times $t \neq n/N$, besides just the integral times. The adaptation of (9.22) that we use is this:

Lemma 9.13. Suppose for $* = L, R$, each of $\mathcal{S}^*_n$ is a random walk of i.i.d. $\mathrm{Exp}(1) - 1$ steps and each of B^*_t is a standard Brownian motion. View $\mathcal{S}^*_n$ as a polygonal path by setting

$$\mathcal{S}^*_{n+t} = (1-t)\mathcal{S}^*_n + t\mathcal{S}^*_{n+1} \qquad \text{with} \qquad t \in [0,1].$$

With $C_0 = 2(C_1 + 1/C_3)$—the latter constants from Theorem 9.12—let

$$F_* = \left\{\max_{0\leq t\leq T}\left|\frac{\mathcal{S}^*_{N_*t}}{\sqrt{N_*}} - B^*_t\right| \geq \frac{C_0\log(N_*T)}{\sqrt{N_*}}\right\} \qquad (* = L, R) \tag{9.23}$$

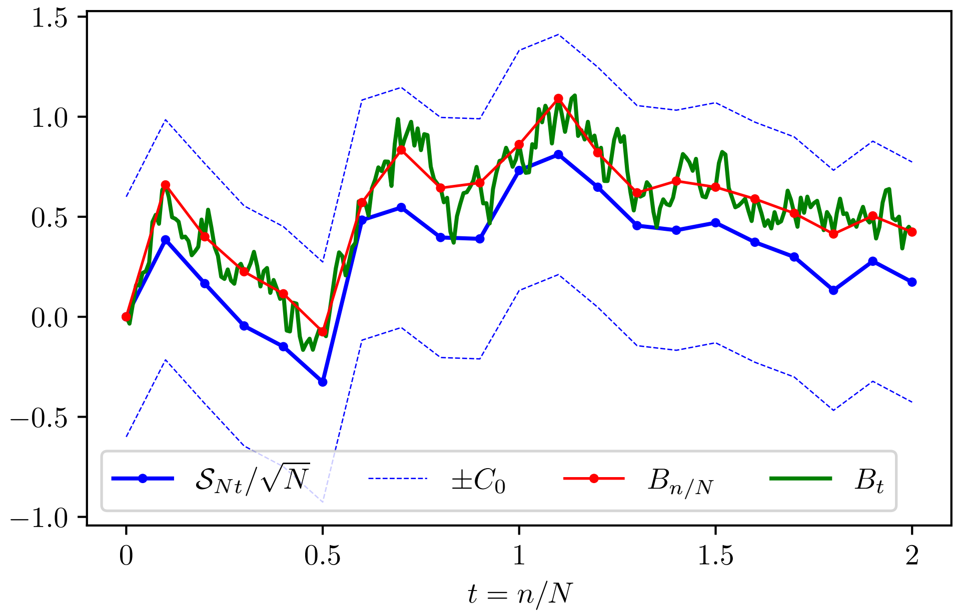

Figure 9.2: Relevant distances between paths. The green path is within $\frac{1}{2}C_0$ of the red path, which is within another $\frac{1}{2}C_0$ of the blue path. (These paths are not a true sample from the coupling in [27, Theorem 1], but the picture is sufficient for intuition.)

---

denote the events where, on one particular side, the polygonal walk moves far apart from its Brownian motion at any time. Then $\mathbb{P}(F_L \cup F_R) \le O\left(\frac{1}{N_L T} + \frac{1}{N_R T}\right)$.

Proof. It is enough to look at one side and drop the $*$. Setting $x = \frac{1}{C_3}\log(NT)$, by (9.22),

$$\mathbb{P}\left(\max_{n=0,1,\dots,NT}\left|\frac{\mathcal{S}_n}{\sqrt{N}} - B_{n/N}\right| \ge \frac{\frac{1}{2}C_0\log(NT)}{\sqrt{N}}\right) \le O\left(\frac{1}{NT}\right). \tag{9.24}$$

We now fill in the intermediate times. To say what this means, consider Figure 9.2 (not a true sample, but sufficient for intuition). By (9.24), the red path remains within $\frac{1}{2}C_0$ of the blue exponential steps at times n/N. On any of the NT intervals $t \in \left(\frac{n}{N}, \frac{n+1}{N}\right)$, the green Brownian motion behaves as a bridge of length $\ell = 1/N$ relative to the red line segment connecting $B_{n/N}$ to $B_{(n+1)/N}$. We can check as follows that no green bridge of length ℓ deviates by more than another $\frac{1}{2}C_0$ from any red line segment. Rescale and use the Kolmogorov-Smirnov statistic [10, equation (8.4.12)] to bound

$$\mathbb{P}\left(\max_{0\le t\le\ell}\left|B_t - \frac{t}{\ell}B_\ell\right| < \frac{C_0}{2}\right) = \mathbb{P}\left(\max_{0\le s\le 1}|B_s - sB_1| < \frac{C_0}{2\sqrt{\ell}}\right) \qquad (\text{set } t = \ell s)$$
$$= \sum_{j\in\mathbb{Z}}(-1)^j e^{-j^2C_0^2/2\ell} \ge 1 - \sum_{j\neq 0} e^{-|j|C_0^2/2\ell} = 1 - \frac{2}{e^{C_0^2N/2}-1}.$$

To guarantee all NT bridges are small, note

$$\left[\mathbb{P}\left(\max_{0\le t\le\ell}\left|B_t - \frac{t}{\ell}B_\ell\right| < C_0\right)\right]^{NT} \ge \left[1 - \frac{2}{e^{-C_0^2N/2}-1}\right]^{NT} = 1 - O\left(\frac{2NT}{e^{2C_0^2N}-1}\right),$$

and the last error is again $O(1/NT)$, since $N \gg T$. □

With N_*, T as above, this error probability is $O(1/N_*T) = O\left(1/X^{\frac{2}{3}p}\log X\right)$ and the maximum error size in (9.23) will be denoted by

$$H_* \stackrel{\text{def}}{=} \frac{C_0\log(N_*T)}{\sqrt{N_*}} = \Theta\left(\frac{\log X}{X^{p/3}}\right) \qquad (* = L, R).$$

To cover the random walk events $E_{*,\alpha X^q}$ by events about the coupled Brownian motions, we only need to allow some extra room for this error of size H_*. Since the boundaries defining events $E_{*,\alpha X^q}$ rescale to (9.19) and (9.21), we define

$$E_{R,\alpha_j,\text{BM}} = \left\{B_t^R > -\frac{1}{2}(t-\theta_R)^2 + \frac{\theta_R^2}{2} - H_R \ \ \forall\, 0 < t < T\right\} \tag{9.25}$$
$$E_{L,\alpha_{j+1},\text{BM}} = \left\{B_t^L < \frac{1}{2}(t+\theta_L)^2 - \frac{\theta_L^2}{2} + H_L \ \ \forall\, 0 < t < T\right\}. \tag{9.26}$$

Now, if the random walk remains close to the Brownian motion, and if the walk avoids its own boundary during the time interval $0 \le n \le N_*T$, then the Brownian motion must avoid the looser-by-H_* boundary during the time interval $0 \le t \le T$. Formally,

$$E_{*,\alpha X^q} \cap F_*^c \subseteq E_{*,\alpha,\text{BM}} \qquad \text{with both} \qquad * = L, R. \tag{9.27}$$

The probabilities of the events $E_{*,\alpha,\text{BM}}$ can both be written in terms of

$$D(\kappa, H, T) \stackrel{\text{def}}{=} \mathbb{P}\Big(\max_{\kappa T \le t \le (\kappa+1)T} \big(B_t - \tfrac{1}{2}t^2\big) < H \;\Big|\; B_{\kappa T} = \tfrac{1}{2}(\kappa T)^2 \Big) \tag{9.28}$$

which will be evaluated later using Groeneboom's paper [17]. For instance, on the RHS, shifting and reflecting shows

$$\begin{aligned}
\mathbb{P}(E_{R,\alpha_j,\text{BM}} \mid X) &= \mathbb{P}\Big(\min_{0<t<T} \big(B_t^R + \tfrac{1}{2}(t-\theta_R)^2 - \tfrac{1}{2}\theta_R^2\big) > -H_R \;\Big|\; B_0^R = 0,\ X \Big) \\
&= \mathbb{P}\Big(\max_{-\theta_R \le t \le -\theta_R + T} \big(B_t^R - \tfrac{1}{2}t^2\big) < H_R \;\Big|\; B_{-\theta}^R = \frac{\theta_R^2}{2},\ X \Big) \\
&= D(-\kappa_j, H_R, T).
\end{aligned}$$

Likewise on the LHS, $\mathbb{P}(E_{L,\alpha_{j+1},\text{BM}} \mid X) = D(\zeta\kappa_{j+1}, H_L, T)$.

Proof of (9.6). Continuing from (9.2), condition on keeping both sides of the walk close to their Brownian motions (the event $F_L^c \cap F_R^c$). Write

$$\begin{aligned}
&\mathbb{P}(X \text{ is extreme} \mid X) \\
&\quad \le \mathbb{P}\Big(\textstyle\bigcup_{\alpha\,:\,|\alpha-1|\le 1/2} \Big[(F_R^c \cap E_{R,\alpha X^q}) \cap (F_L^c \cap E_{L,\alpha X^q}) \Big] \mid X \Big) \\
&\qquad\qquad + \mathbb{P}(F_L \cup F_R \;\big|\; X) + O\big(e^{-cX^p}\big) \\
&\quad \le \mathbb{P}\Big(\textstyle\bigcup_{\alpha\,:\,|\alpha-1|\le 1/2} E_{R,\alpha,\text{BM}} \cap E_{L,\alpha,\text{BM}} \;\big|\; X \Big) + O\Big(\frac{1}{X^{\frac{2}{3}p}} \Big) \quad \text{(Lemma 9.13 and (9.27))} \\
&\quad \le O\Big(\frac{1}{X^{\frac{2}{3}p}} \Big) + \sum_{j=0}^{J-1} \mathbb{P}\big(E_{R,\alpha_j,\text{BM}} \mid X\big) \mathbb{P}\big(E_{L,\alpha_{j+1},\text{BM}} \;\big|\; X\big) \quad \text{(as in (1.5))} \\
&\quad = O\Big(\frac{1}{X^{\frac{2}{3}p}} \Big) + \sum_{j=0}^{J-1} D(-\kappa_j, H_R, T) D(\zeta\kappa_{j+1}, H_L, T).
\end{aligned}$$

□

9.4 Proof of Lemma 9.4

After the definition (9.28), we said the function $D(\kappa, H, T)$ would be evaluated using results by Groeneboom [17], and now we say how to do this.[8] From Groeneboom's work [17, see Corollary 2.1, equations (2.14), (2.15), and (2.24), and integrate the transition density as just above equation (5.1), all using $c = 1/2$]—reusing much of the notation there for clarity—we can write:

$$D(\kappa, H, T) = e^{-T^3\left(\frac{1}{2}\kappa^2+\frac{1}{2}\kappa+\frac{1}{6}\right)-\kappa TH} \int_0^\infty e^{(\kappa+1)Ty} r(T, H, y)\, dy, \tag{9.29}$$

where

$$r(T, H, y) = \frac{1}{\pi} \int_{-\infty}^{\infty} e^{i\lambda T} g_{i\lambda}(\min\{H, y\}) h_{i\lambda}(\max\{H, y\})\, d\lambda, \tag{9.30}$$

and where the functions $g_z(t), h_z(t)$ are defined using Airy functions and $\sigma = 2^{1/3}$ as

$$g_z(t) = \frac{\pi}{\sigma} \cdot \frac{\mathrm{Ai}(\sigma z)\,\mathrm{Bi}(\sigma z + \sigma t) - \mathrm{Ai}(\sigma z + \sigma t)\,\mathrm{Bi}(\sigma z)}{\mathrm{Ai}(\sigma z)}, \tag{9.31}$$

$$h_z(t) = \mathrm{Ai}(\sigma z + \sigma t). \tag{9.32}$$

Note 9.14. Only in §9.4, λ refers to a real number in (9.30) and *not* the function (1.3). The latter function is not needed here, so there should not be any confusion. We use λ to agree as much as possible with the notation in [17]. Also, to be clear, at times below we evaluate $g_z(h)$ at a numerical value $h \geq 0$. This is not the function h_z.

Once again, the goal of this section is to bound $D(\kappa, H, T)$ when $|\kappa| \leq 1/4^{1/3}$. We bound this by essentially replicating intricate calculations done by Groeneboom in [17,

[8] The quantity $D(\kappa, H, T)$ here, if rewritten in terms of κ, T, H, and the notation in the paper by Groeneboom [17], would be the quantity $Q_{1/2}^{(\kappa T, 0)}\left(\max_{\kappa T \leq t \leq (\kappa+1)T} X_t < H\right)$.

§5. Appendix] which will give asymptotic estimates for the integral in (9.29) with (9.30) substituted for r. There are a few very minor changes here to asymptotics (our $H \to 0$ takes the place of Groeneboom's fixed value x, but that is all) and to some constant factors, however the setup is almost exactly like Groeneboom's, and as a result, his technique can carry over to this scenario. Nevertheless, verifying certain portions of this remain quite technical, so while a few calculations are merely sketched out below, a few will be worked out again here in full. There are essentially three parts as we recycle the argument in [17, §5. Appendix].

(a) Our part (a) corresponds to the argument near the bottom of p. 104 in [17] to eliminate the $y \approx 0$ region of the integral (9.29). In particular, Groeneboom notes there an alternative expression for r (otherwise not needed here) gives the useful inequality

$$r(T, H, y) \leq 1 \quad \text{if } T \text{ is large.} \tag{9.33}$$

Set $m = 3$ and use (9.33) on the short interval $y \in [0, mH]$ of (9.29) to find

$$\begin{aligned} &D(\kappa, H, T) \\ &\leq e^{-T^3\left(\frac{1}{2}\kappa^2+\frac{1}{2}\kappa+\frac{1}{6}\right)-\kappa TH}\left[\int_0^{mH} e^{(\kappa+1)Ty} \cdot 1\, dy + \int_{mH}^{\infty} e^{(\kappa+1)Ty} r(T, H, y)\, dy\right] \\ &\leq O\Big(He^{-T^3/24}\Big) + e^{-T^3\left(\frac{1}{2}\kappa^2+\frac{1}{2}\kappa+\frac{1}{6}\right)-\kappa TH} \int_{mH}^{\infty} e^{(\kappa+1)Ty} r(T, H, y)\, dy. \end{aligned} \tag{9.34}$$

To explain the first exponent, note $\min_\kappa \left(\frac{1}{2}\kappa^2 + \frac{1}{2}\kappa + \frac{1}{6}\right) = \frac{1}{24}$ and $TH \to 0$. For later, we record the selection

$$m = 3, \quad \text{so that} \quad m > 1 \quad \text{and} \quad 1 - m = -2. \tag{9.35}$$

(b) Our part (b) corresponds to Groeneboom's estimates on pp. 105–106 in [17]. We need to establish bounds on $g_{i\lambda}(h)$ and its derivative $g'_{i\lambda}(h)$ (with respect to h) that we can use in part (c). In particular, we will use that as $\lambda \to \pm\infty$ with $0 \le h \le 1/\sigma$,

$$|g_{i\lambda}(h)| \sim \frac{1}{\sqrt{\pi}|\sigma\lambda|^{1/4}} \exp\left[h|\lambda|^{1/2} + O\left(\frac{h^2}{|\lambda|^{1/2}}\right) - \frac{2}{3}|\lambda|^{3/2}\right] \tag{9.36}$$

$$\text{and} \quad \left|g'_{i\lambda}(h)\right| \sim \frac{\sigma^{5/4}|\lambda|^{1/4}}{\sqrt{\pi}} \exp\left[h|\lambda|^{1/2} + O\left(\frac{h^2}{|\lambda|^{1/2}}\right) - \frac{2}{3}|\lambda|^{3/2}\right], \tag{9.37}$$

hence

$$|g_{i\lambda}(h)| \le O\left[\exp\left(h|\lambda|^{1/2} - \frac{2}{3}|\lambda|^{3/2}\right)\right] \tag{9.38}$$

$$\text{and} \quad \left|g'_{i\lambda}(h)\right| \le O\left[|\lambda|^{1/4}\exp\left(\frac{1}{\sigma}|\lambda|^{1/2} - \frac{2}{3}|\lambda|^{3/2}\right)\right]. \tag{9.39}$$

We will skip the details of how to prove these, since the details are tedious and not important later, and moreover since the calculation closely follows Groeneboom's estimates in [17, pp. 105–106]:

- one first replaces each Bi in the numerator of (9.31) with Ai using identities given in [17, eq. (5.5) and just below (5.11)], [1, eq. (10.4.9)];
- one next applies (the latter for $g'_{i\lambda}(h)$ only)

$$\mathrm{Ai}(z) \sim \frac{1}{2\sqrt{\pi}z^{1/4}} \exp\left(-\frac{2}{3}z^{3/2}\right) \quad \text{and} \quad \mathrm{Ai}'(z) \sim \frac{z^{1/4}}{2\sqrt{\pi}} \exp\left(-\frac{2}{3}z^{3/2}\right) \tag{9.40}$$

  as $|z| \to \infty$, $|\arg z| < \pi$ from [17, eq. (5.7)] or [25, eqs. (A.1)–(A.2)], from [1, eqs. (10.4.59), (10.4.61)]; and finally
- one analyzes the real parts of the exponents $z^{3/2}$ using a binomial expansion.

The only differences between establishing (9.36) and (9.37) are one first differentiates the numerator of (9.31) with respect to h, and one uses the second estimate in (9.40).

(c) Our part (c) corresponds to the calculations and justification of Groeneboom's equations (5.12) and (5.14) on pp. 106–108 in [17]. This part will check (see below)

$$\int_{-\infty}^{\infty}\int_{mH}^{\infty}\left|e^{(\kappa+1)Ty}g_{i\lambda}(H)h_{i\lambda}(y)\right| dy\, d\lambda \le \widetilde{O}\left(He^{(\kappa+1)^3T^3/6}\right). \tag{9.41}$$

Notice that parts (a) and (c) together justify Lemma 9.4. The O term of (9.34) in (a) is smaller than required, since $-1/24 \le \kappa^3/6$ when $|\kappa| \le 1/4^{1/3}$. In the second integral term of (9.34), substitute (9.30) then use part (c):

$$\begin{aligned} &e^{-T^3\left(\frac{1}{2}\kappa^2+\frac{1}{2}\kappa+\frac{1}{6}\right)-\kappa TH} \int_{mH}^{\infty} e^{(\kappa+1)Ty} r(T,H,y)\,dy \\ &\quad\le e^{-T^3\left(\frac{1}{2}\kappa^2+\frac{1}{2}\kappa+\frac{1}{6}\right)-\kappa TH} \cdot \widetilde{O}\Big(He^{(\kappa+1)^3T^3/6}\Big) \\ &\quad= \widetilde{O}\Big(He^{\kappa^3T^3/6}\Big). \end{aligned}$$

(Again, $TH \to 0$ in the exponent.) This gives the conclusion required for Lemma 9.4.

The remainder of this section provides more detail about part (c). Again, here we closely follow Groeneboom's arguments in [17, §5. Appendix, pp. 106–108], especially the justification of Groeneboom's equations (5.12) and (5.14). Groeneboom's argument applies Laplace's method to estimate an integral like the one in part (c), up to a few minor changes. (For background about Laplace's method, see especially [9, §4.2–4.3], or [11].) Applying these arguments involves rescaling the integration variables y, λ. In parallel to Groeneboom's substitutions, it will be convenient to introduce u, v defined by

$$y = (\rho T)^2 u, \quad \lambda = (\rho T)^2 v, \quad \text{where} \quad \rho = \kappa + 1, \tag{9.42}$$

which differ only by constant multiples from those in [17, middle of p. 107]. Since $|\kappa| \le 1/4^{1/3} \approx 0.62996$, it will be useful to keep in mind throughout that

$$\frac{1}{3} < \rho = \kappa + 1 < \frac{5}{3} \tag{9.43}$$

is positive, bounded above, and bounded away from zero.

Before beginning, let us mention again the few minor changes between (9.41) and the setup of integrals in Groeneboom's paper [17, §5. Appendix]:

- The asymptotics here are slighlty different. We consider $g_{i\lambda}(H)$ with $H \to 0$ as $X, T \to \infty$, whereas when Groeneboom considers $g_{i\lambda}(x)$, his x is fixed. To side-step this issue below, the rough idea is to write $g_{i\lambda}(H) \approx H g'_{i\lambda}(0)$, or more formally to use the derivative bound in (b). This explains the factor H in (9.41) above.
- As already mentioned around (9.42), some of the constants here are slightly different. Letting Groeneboom's $c = 1/2$ and replacing his t with $(\kappa+1)T = \rho T$ gives other changes. These do not much matter within the argument.

Again, these changes are small enough that Groeneboom's argument via Laplace's method still works, so we will use essentially the same technique here. There are a few minor tweaks below, however. For one, Groeneboom splits into cases $|v| < 2$ and $|v| > 2$, which after the change-of-constants would correspond to splitting around 1 here. Instead, we will split around $(7/6)^2$; this is somewhat arbitrary, but since $7/6 > 1$, it helps slightly with the exponents where $|v| > (7/6)^2$.

To begin, notice $\exp\left(-\frac{2}{3}\operatorname{Re}\left[z^{3/2}\right]\right)$ does not vanish for any small z, so use (9.40) within a O to bound (9.32), then subsitute for both u, v as in (9.42) to say

$$\begin{aligned}
&\int_{-\infty}^{\infty} |g_{i\lambda}(H)| \int_{mH}^{\infty} e^{\rho T y} |h_{i\lambda}(y)|\, dy\, d\lambda \\
&\quad \le \int_{-\infty}^{\infty} |g_{i\lambda}(H)| \int_{mH}^{\infty} O\left(\exp\left[\rho T y - \frac{2}{3}\operatorname{Re}\left[\sigma^{3/2}\cdot (y+i\lambda)^{3/2}\right]\right]\right) dy\, d\lambda \\
&\quad = (\rho T)^4 \int_{-\infty}^{\infty} |g_{i\lambda}(H)| \int_{mH/(\rho T)^2}^{\infty} O\left(\exp\left[(\rho T)^3 \underbrace{\left(u - \tfrac{2\sqrt{2}}{3}\operatorname{Re}\left[(u+iv)^{3/2}\right]\right)}_{w_v(u) \stackrel{\text{def}}{=}}\right]\right) du\, dv.
\end{aligned}$$

Divide the v integral into two regions,

$$R_1 = \left\{0 \le |v| \le \left(\frac{7}{6}\right)^2\right\} \quad \text{and} \quad R_2 = \left\{|v| \ge \left(\frac{7}{6}\right)^2\right\},$$

and focus first on R_1. To begin Laplace's method for the inner integral, we look at

$$\begin{aligned} w_v(u) &= u - \frac{2\sqrt{2}}{3}\operatorname{Re}\Big[(u+iv)^{3/2}\Big], \\ w_v'(u) &= 1 - \sqrt{2}\operatorname{Re}\Big[(u+iv)^{1/2}\Big], \\ w_v''(u) &= -\frac{\sqrt{2}}{2}\operatorname{Re}\left[\frac{1}{(u+iv)^{1/2}}\right]. \end{aligned}$$

To understand solutions u of $w_v'(u) = 0$, Groeneboom checks desired properties using a polar-coordinate representation [17, see p. 107 bottom to p. 108 top]. A slightly more geometric approach is this: since the real part of $(u+iv)^{1/2}$ must be $1/\sqrt{2}$, each $u+iv$ with $w_v'(u) = 0$ sits along the parabola

$$t \mapsto \left(\frac{1+it}{\sqrt{2}}\right)^2 = \frac{1-t^2}{2} + it \qquad (t \in \mathbb{R}), \tag{9.44}$$

so $w_v'(u) = 0$ tells us $u = \frac{1}{2}(1-v^2)$ and $(u+iv)^{1/2} = \frac{1}{\sqrt{2}}(1+iv)$; then

$$w_v''\left(\frac{1-v^2}{2}\right) = -\frac{\sqrt{2}}{2}\operatorname{Re}\left[\frac{\sqrt{2}}{1+iv}\right] = -\frac{1}{1+v^2}$$

and

$$w_v\left(\frac{1-v^2}{2}\right) = \frac{1-v^2}{2} - \frac{2\sqrt{2}}{3}\operatorname{Re}\left[\frac{1}{2^{3/2}}(1+iv)^3\right] = \frac{1}{6} + \frac{1}{2}v^2.$$

Notice $\frac{1}{2}(1-v^2)$ is sometimes positive and sometimes negative when $v \in R_1$, and so sometimes it is less than the lower integral bound, $mH/(\rho T)^2$; in any case we can bound

$$\begin{aligned} \int_{mH/(\rho T)^2}^{\infty} e^{(\rho T)^3 w_v(u)}\,du &\le \int_{\min\left\{0,\ \frac{1}{2}(1-v^2)\right\}}^{\infty} e^{(\rho T)^3 w_v(u)}\,du \\ &\overset{*}{\le} O\left(\sqrt{\frac{1+v^2}{(\rho T)^3}} \cdot e^{(T\rho)^3\left(\frac{1}{6}+\frac{1}{2}v^2\right)}\right) = O\Big((\rho T)^{-3/2} e^{(\rho T)^3\left(\frac{1}{6}+\frac{1}{2}v^2\right)}\Big) \end{aligned}$$

using Laplace's method at $\overset{*}{\le}$ and then $|v| \le O(1)$. (Again, see [9, §4.2].) We now use this in the double integral with $v \in R_1$. Simultaneously, write $|g_{i\lambda}(H)| \le H \cdot \max_{0\le h\le 1/\sigma} |g_{i\lambda}'(h)|$

(since $g_{i\lambda}(0) = 0$) and use (9.39) to bound this maximum. Then

$$\int_{R_1}\int_{mH/(\rho T)^2}^{\infty} e^{\rho T y}|g_{i\lambda}(H)h_{i\lambda}(y)|\,du\,dv$$
$$\le \int_{R_1} O\Bigg(H\cdot|\lambda|^{1/4}\exp\left[\frac{1}{\sigma}|\lambda|^{1/2} - \overbrace{\frac{2}{3}|\lambda|^{3/2}}^{(*)}\right]\cdot(\rho T)^{-3/2}\exp\left[(\rho T)^3\left(\frac{1}{6}+\frac{1}{2}v^2\right)\right]\Bigg)\,dv$$
$$\le O\Bigg(H\int_{R_1}|\lambda|^{1/4}\exp\left[\frac{1}{\sigma}|\lambda|^{1/2} - \frac{1}{12}|\lambda|^{3/2}\right]\cdot\exp\left[(\rho T)^3\left(\frac{1}{6}+\frac{1}{2}v^2-\frac{7}{12}|v|^{3/2}\right)\right]dv\Bigg),$$

in the final line separating $(*)$ as $\frac{2}{3} = \frac{1}{12} + \frac{7}{12}$, keeping 1/12 in the first exponent and moving 7/12 into the second. Within the second exponent,

$$\max_{|v|\le(7/6)^2}\left(\frac{1}{6}+\frac{1}{2}v^2-\frac{7}{12}|v|^{3/2}\right) = \frac{1}{6},$$

so we have

$$\int_{v\in R_1}\int_{mH/(\rho T)^2}^{\infty} e^{\rho T y}|g_{i\lambda}(H)h_{i\lambda}(y)|\,dy\,d\lambda$$
$$\le O\Bigg(He^{(\rho T)^3/6}\cdot(\rho T)^2\int_{-\infty}^{\infty}|\lambda|^{1/4}\exp\left[\frac{1}{\sigma}|\lambda|^{1/2}-\frac{1}{12}|\lambda|^{3/2}\right]d\lambda\Bigg)$$
$$= \widetilde{O}\Big(He^{(\rho T)^3/6}\Big).$$

This is the desired scale over R_1.

On R_2, where $|v| > (7/6)^2$, we use the boundary version of Laplace's method (see [9, §4.3]). From

$$w_v(0) = \frac{2}{3}|v|^{3/2}, \qquad w_v'(0) = 1 - |v|^{1/2}, \qquad \text{and} \qquad w_v''(u) < 0 \text{ when } u \ge 0,$$

we have

$$w_v(u) \le w_v(0) + w_v'(0)u = \frac{2}{3}|v|^{3/2} - \Big(|v|^{1/2} - 1\Big)u \qquad \text{when } u \ge 0,$$

so that with $v \in R_2$ and $|v|^{1/2} - 1 \geq 1/6 > 0$,

$$\int_{mH/(\rho T)^2}^{\infty} e^{(\rho T)^3 w_v(u)}\, du \leq O\Big((\rho T)^{-3} \exp\Big[\frac{2}{3}|v|^{3/2}(\rho T)^3 - mH \cdot \rho T\Big(|v|^{1/2} - 1\Big)\Big]\Big).$$

From this and (9.38),

$$\begin{aligned}
&\int_{R_2}\int_{mH/(T\rho)^2}^{\infty} e^{\rho T}|g_{i\lambda}(H)h_{i\lambda}(y)|\, dy\, d\lambda \\
&\quad \leq (\rho T)^4 \int_{R_2} O\Bigg[\exp\Bigg(H|\lambda|^{1/2} \underbrace{-\frac{2}{3}|\lambda|^{3/2}}_{\text{cancels with}}\Bigg) \\
&\qquad\qquad \cdot (\rho T)^{-3}\exp\Bigg(\overbrace{\frac{2}{3}|v|^{3/2}(\rho T)^3} - mH \cdot \rho T\Big(|v|^{1/2} - 1\Big)\Bigg)\Bigg]\, dv \\
&\quad = \rho T \int_{R_2} O\Big[\exp\Big(\rho T H(1-m)|v|^{1/2}\Big)\Big]\, dv \\
&\quad \leq O\Big(\frac{1}{\rho T H^2}\Big)
\end{aligned}$$

since $1 - m = -2$ by (9.35) and $\int_{-\infty}^{\infty} e^{-C|v|^{1/2}}\, dv = 4/C^2$. Finally, rewrite in terms of X. Remember $0 < \Xi \leq 1/3$ and $\rho > 0$ from (9.43) to see

$$\frac{1}{\rho T H^2} \lesssim X^{2/3} \ll \frac{e^{(\rho \log X)^3/6}}{X^{1/3}} \lesssim H e^{(\rho T)^3/6},$$

so the final double integral over R_2 was also $\widetilde{O}\Big(H e^{(\rho T)^3/6}\Big)$.

9.5 Constructing κ_j, Proof of Lemma 9.6

Values κ_0, κ_J were defined in (9.4). We are interested in the intermediate terms, $\kappa_1, \ldots, \kappa_{J-1}$. Build the sequence inductively so that (9.8) holds with equality. That is, define

$$\kappa_1 = -\frac{1}{2} \qquad \text{and} \qquad \kappa_j = \frac{1}{\zeta}\Big(\kappa_{j-1} + \frac{1}{T^3}\Big) \quad \text{when } j > 1, \tag{9.45}$$

stopping at an index where $1/3 < \kappa_{J-1} \leq 1/2$. That this happens is checked below. Recalling $\zeta \stackrel{\text{def}}{=} 1 - T^{-4} \nearrow 1$, parts (b) and (d) of the lemma are automatic.

We now prove (a), that the sequence is increasing. From (9.45), notice

$$\kappa_j - \kappa_{j-1} = \frac{1}{\zeta}\left(\kappa_{j-1}(1-\zeta) + \frac{1}{T^3}\right).$$

- If $\kappa_{j-1} \geq 0$, this is clearly positive, since $1 - \zeta = 1/T^4 > 0$.
- Alternatively, if $\kappa_{j-1} < 0$, induct backwards to say $\kappa_{j-1} \geq \kappa_1 = -1/2 > -1$. Then $\kappa_j - \kappa_{j-1} > 0$ once $T > 1$, since then

$$\kappa_{j-1}(1-\zeta) + \frac{1}{T^3} > \zeta - 1 + \frac{1}{T^3} = -\frac{1}{T^4} + \frac{1}{T^3} > 0.$$

This proves (a), except perhaps at $j = 0, J$, which can be checked separately.

Now we prove (c) and (e) together, that eventually $1/3 < \kappa_{J-1} \leq 1/2$ after $J \leq O(T^4)$ terms. The closed form of (9.45) is

$$\kappa_j = \frac{-1}{2\zeta^{j-1}} + \frac{1}{T^3} \cdot \frac{1/\zeta - 1/\zeta^j}{1 - 1/\zeta} \qquad (j \geq 1), \tag{9.46}$$

so $\kappa_j > 1/3$ if

$$\frac{1}{T^3} \cdot \frac{1/\zeta - 1/\zeta^j}{1 - 1/\zeta} > \frac{1}{3} + \frac{1}{2\zeta^{j-1}}.$$

Substituting $\zeta = 1 - 1/T^4$, choosing $j = T^4$, and then letting $T \to \infty$, the LHS of this equation diverges to $+\infty$, whereas the RHS clearly converges to $1/3 + 1/2$. This shows we can find an index $J - 1 \leq O(T^4)$ where $\kappa_{J-1} > 1/3$. If $J - 1$ is the first such index, then by (9.45),

$$\kappa_{J-1} \leq \frac{1}{\zeta}\left(\frac{1}{3} + \frac{1}{T^3}\right) \leq \frac{1}{2}$$

if T is large. This completes the construction.

9.6 Proof of Proposition 9.2(a)

Most of the work and all of the ideas needed for $1/\log x$ carry over from $1/x^q$. Really, the only changes are in the rescaling steps, so we only present a skeleton outline of these. In this case, we use slopes parameterized by α on the order

$$\beta = \alpha/f(X) = \alpha \log X.$$

Given any such β, Remark 6.3 requires that n_X be asymptotically much smaller than both $\lambda(X) \sim X/\log X$ (as in (3.3)) and $X/\beta = X/\alpha \log X$, so if we select $n_X = \epsilon X/\log X$, we need

$$\epsilon < \min\{1,\ 1/\alpha\}. \tag{9.47}$$

To discard $\alpha \not\approx 1$, we need bounds that play the roles of (9.11) and (9.12). These are: on the RHS, when $n > 0$,

$$\lambda(X + n\beta) - \lambda(X) - n \geq (\alpha - 1)n - \frac{\alpha^2 n^2}{2X}, \tag{9.48}$$

and on the LHS, when $0 \leq n \leq \epsilon X/\log X$,

$$\lambda(X) - \lambda(X - n\beta) - n \leq (\alpha - 1)n + \frac{\alpha^2 n^2}{2X} \cdot \gamma_{\epsilon,\alpha} \quad \text{with} \quad \gamma_{\epsilon,\alpha} = \frac{1}{1 - 2\epsilon\alpha}. \tag{9.49}$$

These will be checked below, but both are similar to triangle estimates used before. Proceeding as in Lemma 9.3 establishes constants $c, \nu > 0$ so that

$$\mathbb{P}(X \text{ is extreme} \mid X) \leq \mathbb{P}\Big(\bigcup_{\alpha\,:\,|\alpha-1|\leq 1/2} E_{L,\beta} \cap E_{R,\beta} \;\Big|\; X\Big) + O\Big(e^{-c(X/\log X)^{\nu}}\Big).$$

Now with $|\alpha - 1| \leq 1/2$, define

$$\gamma_\epsilon = \gamma_{\epsilon,3/2} = \frac{1}{1 - 3\epsilon}$$

so that $\gamma_{\epsilon,\alpha} < \gamma_\epsilon$. Complete squares in (9.48) and (9.49) (using $\gamma_{\epsilon,\alpha} < \gamma_\epsilon$ in the latter), and let $n = N_* t$ in each case: we find that

$$\begin{aligned} &\frac{1}{\sqrt{N_R}}(\lambda(X + n\beta) - \lambda(X) - n) \\ &\qquad \geq -\frac{1}{2}(t - \theta_R)^2 + \frac{\theta_R^2}{2} \quad \text{with} \quad N_R = \left(\frac{X}{\alpha^2}\right)^{2/3} \quad \text{and} \quad \theta_R = \frac{(\alpha - 1)X^{1/3}}{\alpha^{2/3}}, \end{aligned} \tag{9.50}$$

and when $0 \le n = N_L t \le \frac{\epsilon X}{\log X} = n_X$,

$$\frac{1}{\sqrt{N_L}}(\lambda(X) - \lambda(X - n\beta) - n) \\ \le \frac{1}{2}(t + \theta_L)^2 - \frac{\theta_L^2}{2} \quad \text{with} \quad N_L = \left(\frac{X}{\alpha^2 \gamma_\epsilon}\right)^{2/3} \quad \text{and} \quad \theta_L = \frac{(\alpha - 1)X^{1/3}}{\alpha^{2/3}\gamma_\epsilon^{1/3}}. \tag{9.51}$$

The reparameterization here in place of (9.5) is

$$\kappa_j = \left(\frac{\alpha_j - 1}{\alpha_j^{2/3}}\right) \cdot \frac{X^{1/3}}{\log X},$$

the only differences being the exponent $1/3$ instead of $p/3$ and no factor of q. We now define $\zeta \to 1$ and $\epsilon \to 0$ depending on X in such a way that

$$\zeta \overset{\text{def}}{=} \frac{1}{\gamma_\epsilon^{1/3}} = (1 - 3\epsilon)^{1/3} \overset{\text{def}}{=} 1 - \frac{1}{\log^4 X}, \qquad \text{which means} \qquad \epsilon \sim \frac{1}{\log^4 X}. \tag{9.52}$$

Then as before

$$\theta_L = \zeta\kappa_{j+1}T, \qquad \theta_R = \kappa_j T, \qquad T = \log X,$$

and the rest of the proof is more or less the same as the $1/x^q$ case: define events $E_{*,\alpha,\mathrm{BM}}$ as in (9.25)–(9.26)—but now using new values of θ_* and H_*, the allowed error size in Lemma 9.13—then argue about how to bound (9.6) using (9.7).

To wrap up, we check (9.48)–(9.49) using the triangles in Figure 9.3.

Proof of (9.48). Since $\beta = \alpha \log X$,

$$\lambda(X + n\beta) - \lambda(X) - n = (\alpha - 1)n - \int_X^{X+n\beta} \frac{1}{\log X} - \frac{1}{\log t}\, dt,$$

but we have an upper bound from the green triangle:

$$\int_X^{X+n\beta} \frac{1}{\log X} - \frac{1}{\log t}\, dt \le \frac{1}{2}(n\beta) \cdot \left(n\beta \cdot \frac{1}{X \log^2 X}\right) = \frac{\alpha^2 n^2}{2X}.$$ □

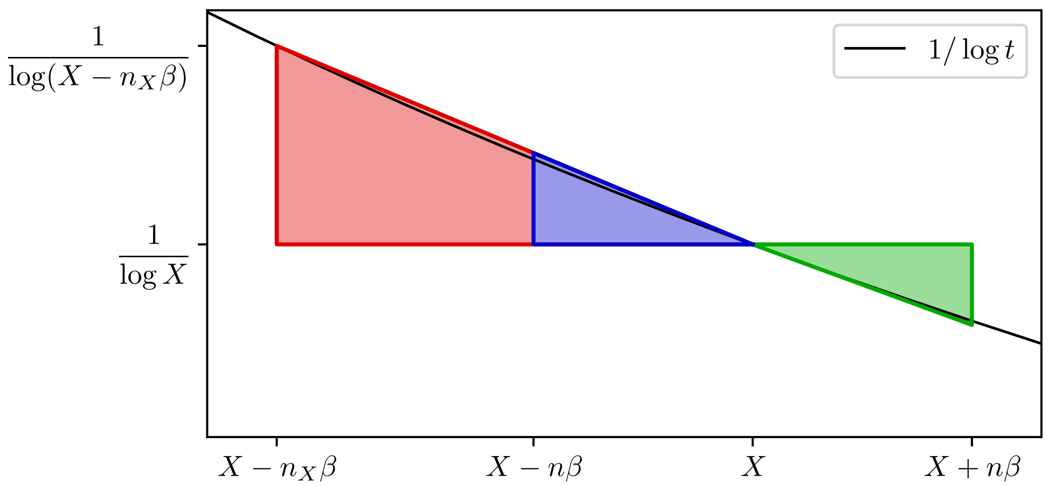

Figure 9.3: More triangular area approximations. The blue triangle is scaled down from the red one, which has two points on the curve $1/\log t$. The green triangle's lower boundary is the tangent line to $1/\log t$ at $t = X$.

Proof of (9.49). We have

$$\lambda(X) - \lambda(X - n\beta) - n = (\alpha - 1)n + \int_{X-n\beta}^{X} \frac{1}{\log t} - \frac{1}{\log X}\, dt.$$

From the blue area,

$$\begin{aligned}
\int_{X-n\beta}^{X} \frac{1}{\log t} - \frac{1}{\log X}\, dt &\le \frac{1}{2}(n\beta) \cdot \left(\frac{n\beta}{n_X\beta}\right) \cdot \left(\frac{1}{\log(X - n_X\beta)} - \frac{1}{\log X}\right) \\
&= \frac{1}{2}(n\alpha)^2 \cdot \frac{\log^2 X}{\epsilon\alpha X} \cdot \left(\frac{\log\left(\frac{1}{1-\epsilon\alpha}\right)}{\log X \log(X - \epsilon\alpha X)}\right) \\
&= \frac{1}{2}(n\alpha)^2 \cdot \frac{1}{\epsilon\alpha X} \cdot \left(\frac{\log\left(\frac{1}{1-\epsilon\alpha}\right)}{1 + \frac{\log(1-\epsilon\alpha)}{\log X}}\right)
\end{aligned}$$

Now Taylor expand to see $-\log(1 - \epsilon\alpha) < \epsilon\alpha + (\epsilon\alpha)^2 + (\epsilon\alpha)^3 + \cdots = \frac{\epsilon\alpha}{1-\epsilon\alpha}$. Use this in the numerator. In the denominator use $\frac{1}{1 + \frac{\log(1-\epsilon\alpha)}{\log X}} < \frac{1}{1 - \frac{\epsilon\alpha}{(1-\epsilon\alpha)\log X}} < \frac{1}{1 - \frac{\epsilon\alpha}{1-\epsilon\alpha}} = \frac{1 - \epsilon\alpha}{1 - 2\epsilon\alpha}$, then cancel the $1 - \epsilon\alpha$. □

Chapter 10

Conjectures Built from Similar Results of Piet Groeneboom

Here we merely rescale the intensity $f(x)$ in a way Robin Pemantle suggested (described in §10.1) to compare a squished Poisson process to points from a density $g(x)$ on $[0, 1]$. There are technical differences, but the rescaling leads one to each of the guesses §1.3.6(a) and §1.3.6(b). After the rescaling, we see how for the example $1/\log x$ in §10.2.

10.1 Rescaling

This section describes for any intensity $f(x)$ the rescaling that Robin Pemantle suggested. We first need some notation. For a given intensity $f(x)$, define

$$\Lambda_s \stackrel{\text{def}}{=} \lambda^{-1}(s).$$

Notice Λ_n may be regarded as the typical position for the nth point X_n (see Lemma 2.1 with $S_n \approx n$). Temporarily fix a length $\ell > 0$. From an intensity $f(x) = f_1(x)$, consider the squished and scaled intensity

$$f_\ell(x) = \ell f_1(\ell x). \tag{10.1}$$

The expected number of points with this squished intensity is $\lambda_\ell(t) = \int_0^t f_\ell(x)\,dx = \lambda_1(\ell t)$, which satisfies

$$\lambda_\ell^{-1}(s) \overset{*}{=} \frac{1}{\ell}\lambda_1^{-1}(s).$$

Notice $\overset{*}{=}$ tells us that if we couple Poisson processes $X_n^{(1)}$ and $X_n^{(\ell)}$ for f_1 and f_ℓ via Lemma 2.1, then

$$X_n^{(\ell)} \overset{\text{def}}{=} \lambda_\ell^{-1}(S_n) \overset{*}{=} \frac{1}{\ell}\lambda_1^{-1}(S_n) \overset{\text{def}}{=} \frac{1}{\ell}X_n^{(1)};$$

in particular, points from intensity $f_\ell(x)$ are a linear $\frac{1}{\ell}$-scaling of points from $f_1(x)$.

As mentioned above, Robin Pemantle's suggestion is to rescale so $\approx n$ points are within $[0,1]$, meaning we set $\ell = \Lambda_n$ and renormalize to define the density

$$g_n(x) \overset{\text{def}}{=} \frac{\Lambda_n f(\Lambda_n x)}{\int_0^1 \Lambda_n f(\Lambda_n y)\,dy} = \frac{\Lambda_n f(\Lambda_n x)}{n} \qquad \text{on } 0 < x < 1. \tag{10.2}$$

As Remark 1.10 noted, there is a chance this rescaling could resemble one in forthcoming work mentioned in [19, Lemma 3.1].

10.2 Explanation of the conjecture §1.3.6(b) about $1/\log x$

We focus on the guess §1.3.6(b) here, but first, let us say that in the case of $1/x^q$, where the rescaling (10.2) makes $g_n(x) = g(x) = p\,\mathbf{1}[0 < x < 1]/x^q$ independent of n, a similar calculation to the one below leads to the guess §1.3.6(a). Now consider $1/\log x$.

We begin with some comments about the rescaling (10.2), defined here for each n by

$$g_n(x) = \frac{\Lambda_n\,\mathbf{1}[e/\Lambda_n < x < 1]}{n \cdot \log(\Lambda_n x)} \quad \text{which has} \quad g_n'(x) = -\frac{\Lambda_n\,\mathbf{1}[e/\Lambda_n < x < 1]}{n \cdot x \cdot \log^2(\Lambda_n x)}. \tag{10.3}$$

In this case $\Lambda_n \sim n \log n$ from (3.3). Unlike with $1/x^q$, these densities change with n. Related observations were given in §1.3.6 and Remark 1.19, where we promised to verify the following: if $0 < x < 1$ is fixed as $n \to \infty$, then (10.3) gives

$$\begin{aligned} g_n(x) &\sim \frac{n \log n}{n(\log n + \log\log n + \log x)} \sim 1 \\ \text{and} \quad g_n'(x) &\sim -\frac{n \log n}{nx\big(\log n + \log\log n + \log x\big)^2} \sim 0, \end{aligned} \tag{10.4}$$

whereas if $x \sim e/\Lambda_n \to 0$ as $n \to \infty$ then $g_n(x) \sim \log n$.

That g_n changes with n means we cannot genuinely apply (1.7/P.G.1) or (1.8/P.G.2), because Groeneboom's context in [19, §3] considered a single fixed density, g. Nonetheless, the idea here is to compute $\mathbb{E}\mathcal{N}_n$ with the formula (1.7/P.G.1) and guess that is the answer. In particular, conjecture that

$$\begin{aligned} \mathbb{E}\mathcal{E}_{1/\log x}(\Lambda_n) &\approx \mathbb{E}\Big[\mathcal{N}_{N(\Lambda_n)}\Big] \\ &\approx k_1 n^{1/3} \int_{e/\Lambda_n}^{1} \left(\frac{g_n'(x)^2}{4 g_n(x)}\right)^{1/3} dx \quad \text{(as in (1.7/P.G.1) above)} \\ &= k_1 n^{1/3} \left(\frac{\Lambda_n}{4n}\right)^{1/3} \int_{e/\Lambda_n}^{1} \frac{1}{x^{2/3}\log(\Lambda_n x)}\, dx \quad \text{(using definition (10.3) above)} \\ &= \frac{k_1}{4^{1/3}} \int_{e}^{\Lambda_n} \frac{1}{u^{2/3}\log u}\, du \quad \text{(setting } u = \Lambda_n x\text{).} \end{aligned}$$

In a moment, we will calculate:

Claim 10.1. As $n \to \infty$, $\displaystyle\int_{e}^{\Lambda_n} \frac{1}{u^{2/3}\log u}\, du \sim \frac{3n^{1/3}}{(\log n)^{2/3}}$.

First, to complete the calculation above, let $t = \Lambda_n$ and $n \sim \dfrac{t}{\log t}$, so that we guess

$$\mathbb{E}\mathcal{E}_{1/\log x}(t) \approx \frac{k_1}{4^{1/3}} \cdot \frac{3n^{1/3}}{(\log n)^{2/3}} \sim k_1 \cdot \frac{3}{4^{1/3}} \cdot \frac{t^{1/3}}{\log t}.$$

This should explain the mean and variance in §1.3.6(b). Again, the normality conjecture is simply based on Groeneboom's result (1.8/P.G.2) from [19, Theorem 3.1].

Proof of Claim 10.1. We use the methods of [36, Ch 3.4]. An immediate lower bound is

$$\int_e^{\Lambda_n} \frac{1}{u^{2/3}\log u}\,du > \int_e^{\Lambda_n} \frac{1}{u^{2/3}\log \Lambda_n}\,du = \frac{3\left(\Lambda_n^{1/3} - e^{1/3}\right)}{\log \Lambda_n} \sim \frac{3n^{1/3}}{(\log n)^{2/3}}.$$

For the other direction, split the integral around an interior bound $b \in (e, \Lambda_n)$:

$$\begin{aligned}\int_e^{\Lambda_n} \frac{1}{u^{2/3}\log u}\,du &< \int_e^{b} \frac{1}{u^{2/3}}\,du + \int_b^{\Lambda_n} \frac{1}{u^{2/3}\log b}\,du \\ &= 3\left(b^{1/3} - e^{1/3} + \frac{\Lambda_n^{1/3} - b^{1/3}}{\log b}\right).\end{aligned} \tag{10.5}$$

Set $b = \dfrac{n}{(\log n)^3}$. Then since $\log b \sim \log n$ and $\Lambda_n \sim n\log n$,

$$b^{1/3} = \frac{n^{1/3}}{\log n} \ll \frac{n^{1/3}}{(\log n)^{2/3}} \sim \frac{\Lambda_n^{1/3}}{\log n} \sim \frac{\Lambda_n^{1/3}}{\log b}.$$

Using this last observation in (10.5) gives

$$\int_e^{\Lambda_n} \frac{1}{u^{2/3}\log u}\,du \lesssim 3\left(\frac{\Lambda_n^{1/3}}{\log b}\right) \sim \frac{3n^{1/3}}{(\log n)^{2/3}}. \qquad \square$$

Appendix A

Computation and Figures

Although §5 outlines a strategy to guarantee extreme points are correct, most pictures were not generated this way. Most were generated more simply and so could be inaccurate (in the sense of §5) though accurate enough for our intuition. Figure 5.2 was, however, generated using the strategy in §5.

The computations related to H in §5.3, the simulations related to g_r in §1.4 and the simulations for the constant μ in §7.3 used Python (`https://www.python.org/`). The figures were made with Python, matplotlib (v. 3.4.2, see `https://matplotlib.org/` and [23]), and TikZ.

Bibliography

[1] Milton Abramowitz and Irene A. Stegun (eds.), *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, U.S. Government Printing Office, Washington, D.C., 1964.

[2] Imre Bárány, *Random points and lattice points in convex bodies*, Bulletin of the American Mathematical Society **45** (2008), no. 3, 339–365.

[3] Andrew D. Barbour, *A note on the maximum size of a closed epidemic*, Journal of the Royal Statistical Society. Series B (Methodological) **37** (1975), no. 3, 459–460.

[4] A. D. Barbour, *Brownian motion and a sharply curved boundary*, Advances in Applied Probability **13** (1981), no. 4, 736–750.

[5] Patrick Billingsley, *Probability and measure*, anniversary ed., John Wiley and Sons, Inc., New Jersey, 2012.

[6] Richard Crandall and Carl Pomerance, *Prime numbers: A computational perspective*, second ed., Springer, New York, 2005.

[7] H. E. Daniels, *The maximum size of a closed epidemic*, Advances in Applied Probability **6** (1974), no. 4, 607–621.

[8] H. E. Daniels and T. H. R. Skyrme, *The maximum of a random walk whose mean path has a maximum*, Advances in Applied Probability **17** (1985), no. 1, 85–99.

[9] N. G. de Bruijn, *Asymptotic methods in analysis*, Dover Publications, Inc., New York, 1981.

[10] Rick Durrett, *Probability: Theory and examples*, fifth ed., Cambridge University Press, Cambridge, 2019.

[11] A. Erdélyi, *Asymptotic expansions*, Dover Publications, Inc., New York, 1956.

[12] William Feller, *An introduction to probability theory and its applications*, third ed., vol. I, John Wiley & Sons, Inc., 1968.

[13] William Feller, *An introduction to probability theory and its applications*, second ed., vol. II, John Wiley & Sons, Inc., 1971.

[14] Piet Groeneboom, *The concave majorant of Brownian motion*, The Annals of Probability **11** (1983), no. 4, 1016–1027.

[15] P. Groeneboom, *Estimating a monotone density*, Proceedings of the Berkeley Conference In Honor of Jerzy Neyman and Jack Kiefer, Vol. II (Berkeley, Calif., 1983) (Lucien M. Le Cam and Richard A. Olshen, eds.), Wadsworth Statistics/Probability Series, Wadsworth, Belmont, California, 1985, pp. 539–555.

[16] Piet Groeneboom, *Limit theorems for convex hulls*, Probability Theory and Related Fields **79** (1988), no. 3, 327–368.

[17] Piet Groeneboom, *Brownian motion with a parabolic drift and Airy functions*, Probability Theory and Related Fields **81** (1989), no. 1, 79–109.

[18] Piet Groeneboom, *The maximum of Brownian motion minus a parabola*, Electronic Journal of Probability **15** (2010), paper no. 62, pages 1930–1937.

[19] Piet Groeneboom, *Vertices of the least concave majorant of Brownian motion with parabolic drift*, Electronic Journal of Probability **16** (2011), paper no. 84, pages 2234–2258.

[20] P. Groeneboom and H. P. Lopuhaä, *Isotonic estimators of monotone densities and distribution functions: basic facts*, Statistica Neerlandica **47** (1993), no. 3, 175–183.

[21] Olle Häggström, *Problem solving is often a matter of cooking up an appropriate Markov chain*, Scandinavian Journal of Statistics **34** (2007), no. 4, 768–780.

[22] T. E. Harris, *A lower bound for the critical probability in a certain percolation process*, Proceedings of the Cambridge Philosophical Society **56** (1960), 13–20.

[23] J. D. Hunter, *Matplotlib: A 2D graphics environment*, Computing in Science & Engineering **9** (2007), no. 3, 90–95.

[24] Svante Janson, *Moments of the location of the maximum of Brownian motion with parabolic drift*, Electronic Communications in Probability **18** (2013), no. 15, 1–8.

[25] Svante Janson, Guy Louchard, and Anders Martin-Löf, *The maximum of Brownian motion with parabolic drift*, Electronic Journal of Probability **15** (2010), paper no. 61, pages 1893–1929.

[26] Samuel Karlin and Yosef Rinott, *Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions*, Journal of Multivariate Analysis **10** (1980), no. 4, 467–498.

[27] J. Komlós, P. Major, and G. Tusnády, *An approximation of partial sums of independent RV'-s, and the sample DF. I*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **32** (1975), 111–131.

[28] Thomas G. Kurtz, *Strong approximation theorems for density dependent Markov chains*, Stochastic Processes and their Applications **6** (1978), no. 3, 223–240.

[29] Günter Last and Mathew Penrose, *Lectures on the Poisson process*, Cambridge University Press, Cambridge, 2018.

[30] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer, *Markov chains and mixing times*, second ed., American Mathematical Society, Providence, Rhode Island, 2017.

[31] Nathan McNew, *The convex hull of the prime number graph*, Irregularities in the Distribution of Prime Numbers (J. Pintz and M. Th. Rassias, eds.), Springer, Cham, 2018, pp. 125–141.

[32] Peter Mörters and Yuval Peres, *Brownian motion*, Cambridge University Press, Cambridge, 2010.

[33] Carl Pomerance, *The prime number graph*, Mathematics of Computation **33** (1979), no. 145, 399–408.

[34] B. L. S. Prakasa Rao, *Estimation of a unimodal density*, Sankhyā: The Indian Journal of Statistics, Series A **31** (1969), no. 1, 23–36.

[35] Erik Sparre Andersen, *On the fluctuations of sums of random variables II*, Mathematica Scandinavica **2** (1954), no. 2, 195–223.

[36] Joel Spencer and Laura Florescu, *Aysmptopia*, American Mathematical Society, Providence, Rhode Island, 2014.

[37] Gérald Tenenbaum and Michel Mendès France, translated by Philip G. Spain, *The prime numbers and their distribution*, Student Mathematical Library, vol. 6, American Mathematical Society, Providence, RI, 2000.

[38] Edward Tutaj, *Prime numbers with a certain extremal type property*, Annales Universitatis Paedagogicae Cracoviensis. Studia Mathematica **17** (2018), 127–151.