

EXPLICIT POSITIVE LOWER DENSITY FOR THE COLLATZ CONJECTURE

KENNY LAU

ABSTRACT. Lech Mazur proved that a positive proportion of integers n reach 1 under the Collatz map in at most $10.46 \log n$ steps. We give a quantitative refinement of his construction. For any $X \geq 2^{2^{2^{87640}}}$, at least $X/2^{2^{2^{87640}}}$ positive integers $n < X$ reach 1 within $\lfloor 10.46 \log n \rfloor$ ordinary steps, and $\lfloor 3.4881 \log n \rfloor$ accelerated steps. We also obtain results for paths that include any prescribed integer t not divisible by 3. AxiomProver was used to obtain and formalize this result in Lean.

1. INTRODUCTION

For a positive integer n , define the Collatz map $T(n)$ with the rule

$$T(n) = \begin{cases} n/2, & n \text{ even} \\ 3n + 1, & n \text{ odd} \end{cases}.$$

Conjecture 1.1 (Collatz). *Every positive integer n reaches 1 under finitely many iterations of T .*

For example, using the starting value 17 we obtain the orbit

$$17 \rightarrow 52 \rightarrow 26 \rightarrow 13 \rightarrow 40 \rightarrow 20 \rightarrow 10 \rightarrow 5 \rightarrow 16 \rightarrow 8 \rightarrow 4 \rightarrow 2 \rightarrow 1.$$

The Collatz map can be accelerated by focusing on only the odd values in an orbit. Given an odd number n , the next step $3n + 1$ is even, and then we keep dividing by 2 until we reach an odd value. In other words, the next odd number in the orbit after n is

$$S(n) := \frac{3n + 1}{2^a}$$

where $a = \text{ord}_2(3n + 1)$ is the exponent in the highest power of 2 that divides $3n + 1$. S here is called the Syracuse map, or the accelerated Collatz map. For example, using the same starting value 17, the accelerated orbit is now

$$17 \rightarrow 13 \rightarrow 5 \rightarrow 1$$

which picks out precisely the odd numbers from the original unaccelerated orbit.

Our main result establishes an *explicit positive density* of starting values that reach 1, with explicit control on number of steps:

Corollary 1.2. *Let $(X_0, c) = (2^{D_A}, D_A^{-1})$ where $D_A = 2^{2^{2^{87640}}}$. Then, for every $X \geq X_0$, at least cX odd starting values $n < X$ reach 1 within $10.46 \log(n)$ ordinary steps and $3.4881 \log(n)$ accelerated steps.*

We obtain this as a corollary of a more general theorem, which we also prove in this paper:

Theorem 1.3. *Let t be a positive integer not divisible by 3, and let $(X_0(t), c(t)) = (2^{t^3 D_A}, t^{-4} D_A^{-1})$. Then, for every $X \geq X_0(t)$, at least $c(t)X$ odd starting values $n < X$ reach t within $10.46 \log(n/t)$ ordinary steps using the Collatz map and reach the odd part of t within $3.4881 \log(n/t)$ accelerated steps using the Syracuse map.*

2020 *Mathematics Subject Classification.* Primary 11B83; Secondary 37P99, 60G50.

Key words and phrases. Collatz conjecture, arithmetic dynamics.

Remark 1.4. If t is additionally odd, i.e. $\gcd(t, 6) = 1$, then we can replace “the odd part of t ” with just t .

Remark 1.5. We require t to be not divisible by 3 because otherwise the only starting values that reach t are $2^k t$, which has density 0.

As a corollary of the main result, we obtain a conditional theorem:

Corollary 1.6. *If the Collatz conjecture holds for almost all positive integers, then it is universally true. More precisely, if we have*

$$\lim_{X \rightarrow \infty} \frac{\#\{1 \leq n \leq X \mid \text{Collatz is true for } n\}}{X} = 1,$$

then the Collatz Conjecture is true.

Proof. Assume that almost all starting values reach 1. Now let t be a starting value that does not reach 1. If t is divisible by 3, then after consecutive even steps followed by one odd step, t will not be divisible by 3. By the main theorem, a positive density of starting values reach t , and those do not reach 1, and hence it is not true that almost all starting values reach 1. $\square$

We briefly describe some historical notes about the Collatz conjecture. Krasikov and Lagarias [5] proved that at least $X^{0.84}$ integers below X reach 1, for sufficiently large X . Tao [10] proved that, for any function f tending to infinity, for almost every starting value n , the Collatz orbit of n eventually falls under $f(n)$. Here “almost every” is in the sense of logarithmic density.

From Tao’s proof it follows that if t is sufficiently large then a positive logarithmic density, namely $1 - C_T(\log t)^{-c_t}$, of starting values eventually reach some value below t . We remark that using the proofs in this paper, the constants can be determined explicitly, but this is out of the scope of this paper. Then Mazur [8] proved:

Theorem 1.7. *There exist X_0 and $c > 0$ with the property that for every $X \geq X_0$, at least cX starting values $n < X$ reach 1 within $10.46 \log n$ steps.*

The pair of constants (X_0, c) is not explicitly specified in Mazur’s proof, but Appendix B reconstructs their values to be

$$(X_0, c) = (D_M, D_M^{-1}), \quad D_M = 2^{2^{2^{2^{2^{2^{173}}}}}}. \quad (1.1)$$

Mazur also proved [7] that, more generally, for any prescribed target t with $3 \nmid t$, a positive proportion of integers eventually reach t , without a bound on the number of steps. Shaik [9], building on Mazur’s construction, obtains a positive proportion reaching 1 within $c \log n$ steps for every $c > 3/\log(4/3) \approx 10.428$.

The paper is organized as follows: Section 2 provides a summary of the main argument in Sections 3–8. Then Sections 9–14 establishes a key Fourier bound, while Sections 15 and 16 provide key estimates. Then Appendices A and B contain numerical computations, Appendix C produces the formal Lean certificate, and Appendix D declares the AI usage in this paper.

2. OUTLINE OF THE PROOF

We provide a high-level overview of the argument in this paper. Recall that we require $3 \nmid t$ (Remark 1.5). We now further assume that:

$$t \text{ is odd} \quad (\text{H1})$$

We will explain the minor modification to the argument for even t in Section 2.5.3.

2.1. Preliminaries. We outline some basic facts about the Collatz problem.

2.1.1. *The random model.* Recall that the Syracuse map S , which is the accelerated Collatz map, sends an odd number n to $\frac{3n+1}{2^a}$ where $a = \text{ord}_2(3n+1)$. Given a random even number, there is a $\frac{1}{2}$ chance that it is exactly divisible by 2 once, then a $\frac{1}{4}$ chance that it is exactly divisible by 2 twice, and so on. Therefore, a can be modeled by a geometric random variable A where $\mathbb{P}(A = a) = 2^{-a}$. The average value of a is

$$\mathbb{E}(A) = \sum_{a=1}^{\infty} a 2^{-a} = 2.$$

2.1.2. *The heuristic number of steps.* Since the expected value of a is 2, heuristically the Syracuse map is expected to send n to $\frac{3n+1}{2^2} \approx \frac{3}{4}n$. This means that n should reach 1 in about

$$\log_{4/3}(n) \approx 3.476 \log n$$

accelerated steps using the Syracuse map. Since each Syracuse application with a corresponds to $a + 1$ ordinary steps using the Collatz map, and the expected value of a is 2, this means that n should reach 1 in about

$$(2 + 1) \log_{4/3}(n) \approx 10.428 \log n$$

ordinary steps. More generally, if we want n to reach t instead, then heuristically that would take $3.476 \log \frac{n}{t}$ accelerated steps and $10.428 \log \frac{n}{t}$ ordinary steps.

2.2. First reduction: lower bound for total weight of words landing in a window. In this section we explain the backward tree model and reduce the goal to obtaining a lower bound for the total weight of the words that result in a starting value in the window $[X, 32X)$.

2.2.1. *Walking backwards.* Instead of following the Syracuse map S forwards, we analyze the preimages under S , in order to construct a large set of sources that will reach the given target t .

Suppose $S(n') = n$ for some odd n and n' . We know that for some $a \in \mathbb{N}_{>0}$ we have

$$n = \frac{3n' + 1}{2^a},$$

which is equivalent to

$$n' = \frac{2^a n - 1}{3}.$$

This parametrizes the solutions n' given n in terms of a , provided the expression gives an integer.

The set of valid solutions depends on the residue class of n modulo 6. If $n \equiv 1 \pmod{6}$, then the valid solutions are given by $a = 2, 4, 6, \dots$. For example,

$$S^{-1}(7) = \{9, 37, 149, 597, 2389, \dots\}.$$

If $n \equiv 5 \pmod{6}$, then the valid solutions are given by $a = 1, 3, 5, \dots$. For example,

$$S^{-1}(11) = \{7, 29, 117, 469, 1877, \dots\}.$$

Finally, if $n \equiv 3 \pmod{6}$ then there are no valid solutions.

2.2.2. *Letters and words.* Consider the Syracuse orbit $11 \rightarrow 17 \rightarrow 13 \rightarrow 5 \rightarrow 1$. The values of a in each step are $(1, 2, 3, 4)$. So if we want to start from the target 1 and describe the path backwards to reach the source 11, we would write $w = (4, 3, 2, 1)$, which we call a *word*. Each component is called a *letter*. The example word here has 4 letters.

As we have seen in Section 2.2.1, not all words correspond to valid integer-valued sources, but if the resulting source is an integer, then all of the intermediate steps are integers (Remark 3.1).

2.2.3. *Number of steps.* Given a word $w = (a_1, \dots, a_d)$, we define its *degree* to be

$$\deg(w) := a_1 + \dots + a_d.$$

We also denote by $|w|$ the *length* of the word, which is d . The constructed source then takes $|w|$ accelerated steps to reach the target t , and $|w| + \deg(w)$ ordinary steps (Section 3).

2.2.4. *Weight.* A backward step dictated by the letter a causes the value to be multiplied by approximately $\frac{2^a}{3}$. As a result, if we go backwards through the whole word w , then the value will be multiplied by approximately $3^{-|w|2^{\deg(w)}}$. For the example in Section 2.2.2, we verify that $11 \approx 3^{-4}2^{4+3+2+1} \approx 12.642$. In fact this will always be an overestimate.

We now define the *weight* of a word w to be

$$\text{wt}(w) := 3^{|w|}2^{-\deg(w)},$$

which is the reciprocal of the previous value, so that the source is approximately the target divided by $\text{wt}(w)$. We always have

$$\text{wt}(w) \leq \frac{t}{s}$$

where t is the target and s is the source.

2.2.5. *Window and reduction.* To obtain the bound in the main theorem, we consider the window $[X, 32X)$ and try to construct many sources that land in the window. If t is not part of a cycle, then two different words cannot produce the same source (Lemma 3.3), so the problem reduces to a count of words. We record this constraint as an explicit hypothesis:

$$t \text{ is not part of a cycle.} \tag{H2}$$

Given the target t and a valid word w with source $s \in [X, 32X)$, we have $\text{wt}(w) \leq \frac{t}{s} \leq \frac{t}{X}$, which implies

$$\frac{X}{t} \text{wt}(w) \leq 1.$$

Summing this over all valid words whose source lands in $[X, 32X)$, we obtain that the count of such words is at least X/t times the total weight $\beta_X(t)$ of these words, which is (3.3).

Therefore, the main theorem reduces to obtaining an (explicit) constant lower bound of $\beta_X(t)$, with control over the lengths and degrees of the words, in view of Section 2.2.3.

2.3. **A mixing lemma.** In this section we explain an important lemma that relates to the stability of the set of valid words across different lengths.

2.3.1. *Residue and offset.* Consider words of length 3, with target t . Say the word is (a_1, a_2, a_3) . This word is valid iff the source is an integer, where we recall the source is given by

$$\frac{2^{a_3} \frac{2^{a_2} \frac{2^{a_1} t - 1}{3} - 1}{3}}{3} = \frac{2^{a_1+a_2+a_3} t - 2^{a_2+a_3} - 2^{a_3} \cdot 3 - 9}{27}.$$

Noting that 2 is invertible modulo any power of 3, we conclude that this is an integer iff

$$t \equiv 2^{-a_1} + 2^{-a_1-a_2} \cdot 3 + 2^{-a_1-a_2-a_3} \cdot 9 \pmod{27}.$$

Generalizing this, we can define the offset $\text{off}(w)$ of a word $w = (a_1, \dots, a_d)$ to be

$$\text{off}(w) := 2^{-a_1} + 2^{-a_1-a_2} \cdot 3 + \dots + 2^{-a_1-\dots-a_d} 3^{d-1}$$

so that w is valid iff $t \equiv \text{off}(w) \pmod{3^{|w|}}$, which is Remark 3.1. Furthermore, given any $q \in \mathbb{N}$, we denote by $\text{off}_q(w)$ the residue of $\text{off}(w)$ in $\mathbb{Z}/3^q\mathbb{Z}$.

2.3.2. *Residue density.* Recall our probabilistic model where each letter is modeled by a random variable A with $\mathbb{P}(A = a) = 2^{-a}$. We form random words W by combining independent letters. The probability of a particular word being chosen is $\mathbb{P}(W = w) = 2^{-\deg(w)}$, so $\text{wt}(w) = 3^{|w|}\mathbb{P}(W = w)$.

Consider the distribution of $\text{off}_d(W) \in \mathbb{Z}/3^d\mathbb{Z}$, i.e. the residue of the offset of a random word. Define the *residue density* $\rho_d(y) := 3^d \mathbb{P}(\text{off}_d(W) = y)$ for $y \in \mathbb{Z}/3^d\mathbb{Z}$. This value coincides with the total weight of length- d words that are valid for t (see (3.5)):

$$\sum_w \text{wt}(w) \cdot 1_{\text{off}_d(w)=t} = \sum_w 3^d \mathbb{P}(W = w) \cdot 1_{\text{off}_d(w)=t} = 3^d \mathbb{P}(\text{off}_d(W) = t) = \rho_d(t).$$

2.3.3. *Mixing.* The residue density ρ_d depends on d and could differ wildly across different d . However, we have a *fine-scale mixing* lemma, in the sense of Tao [10], which states that for large enough d , the residue density ρ_d is, on average over residues, well approximated by ρ_{d_0} for much smaller d_0 . More precisely,

$$3^{-d} \sum_{y \in \mathbb{Z}/3^d \mathbb{Z}} |\rho_d(y) - \rho_{d_0}(y \bmod 3^{d_0})| \leq 2^{87533} d_0^{-3/2}$$

for any $d \geq d_0 \geq 2^{80000}$. This is Proposition 4.1, proved in Section 4.

2.4. **Constructing a large family of words.** Recall we need a constant lower bound of the total weight $\beta_X(t)$ of words whose source lands in $[X, 32X)$, so we need a controlled way to construct many such words to satisfy this constraint.

2.4.1. *Growing the family generation by generation.* We construct a family L_j of words generation by generation (Section 5). L_0 consists of only the empty word. At generation j , we take every word in the previous generation, and extend it by every element of a yet to be constructed set $\text{FC}(\delta_j, R_j)$. We then keep the words w that are valid for the target t and that additionally satisfy an unspecified filter that among other things has the effect of forcing the inequality $\frac{85}{256} \leq \text{off}(w) \leq \frac{637}{256}$, whenever $j \geq 5$.

We choose $R_j = 16 + \frac{1}{32}j$ explicitly, but leave δ_j undefined here except that we know for large j it behaves like $200 \cdot 1.01^j$.

We now describe $\text{FC}(\delta, R)$, the set of *first crossings* with average length δ and allowed overshoot R . For any word $w = (a_1, \dots, a_d)$, draw the word-graph $(i, \deg_i(w))$ where $\deg_i(w) := a_1 + \dots + a_i$ is the partial degree. Then draw a straight line passing through $(\delta, 2\delta)$ with slope $\log_2(3) \approx 1.585$. Chop each word off at the *first crossing* of the line, and keep the words with length in $(\delta - \lfloor 3\delta/5 \rfloor, \delta + \lfloor 3\delta/5 \rfloor]$ and overshoot $\leq R$. Note that by the first crossing construction, these words are prefix-disjoint, meaning no word is a prefix of any other word. Recall that the average value of a letter is 2, so almost all words eventually cross the line, and the expected place of crossing is $(\delta, 2\delta)$, so these words have expected length approximately δ and expected degree approximately 2δ .

2.4.2. *Weighted density.* We define the *weighted density* $Z_j(t)$ at generation j (Section 5) to be

$$Z_j(t) := \sum_{w \in L_j} \text{wt}(w) \rho_{\lfloor \frac{1}{4}\delta_j \rfloor}(s_w)$$

where s_w is the source of the word w from target t . Recall the residue density ρ represents how well words can continue from s_w .

2.4.3. *Terminal extension.* We now construct a family $L_{j,X}$ of words that will finally have sources landing near X : take every word in L_j and add every extension in $\text{FC}_X^{\text{term}}(w, \delta_j, R_j)$ so that the resulting word is valid, where FC^{term} is a slight modification of FC that among other things forces the sources to be $\geq X$.

We will use a notion of X being *reachable* from j , which we do not define here (see Section 5), but it roughly means that X is not far away from a typical source of words in L_j .

2.4.4. *Terminal weighted density.* We then define the *terminal weighted density* $Z_{j,X}^{\text{term}}(t)$ to be

$$Z_{j,X}^{\text{term}}(t) := \sum_{w \in L_{j,X}} \text{wt}(w) \rho_{\lfloor \frac{1}{4}\delta_j \rfloor}(s_w),$$

with a similar interpretation as the weighted density.

2.5. **Key estimates and the final proof.** We now state several key estimates for the quantities above that are proved in this paper (Section 6), and explain how they lead to a proof of the main theorem.

2.5.1. *Key estimates.* The weighted density enjoys a propagation error bound from one generation to the next. Precisely, for any $j \geq 6 \times 10^6$, Proposition 6.1 states that

$$|Z_{j+1}(t) - Z_j(t)| \leq 2^{87534.2} j^{\frac{5}{2}} 1.01^{-j}. \quad (2.1)$$

As a consequence, if j is sufficiently large, then the sum of all these errors beyond j is negligible, provided that the main term $Z_j(t)$ is large enough.

We indeed have a large enough main term. Precisely, let $j_0 := 6103914$ and $d_0 := 2^{87639}$. This d_0 is enough to dominate the subscript of every ρ that appears in the definition of Z_{j_0} , so that Z_{j_0} is periodic with period 3^{d_0} . Then, the estimate we need is that the average value of $Z_{j_0}(t)$ across unit residues $t \in (\mathbb{Z}/3^{d_0}\mathbb{Z})^\times$ is greater than 2^{-23} , which is supplied by Proposition 6.5 which gives us

$$\frac{1}{|(\mathbb{Z}/3^{d_0}\mathbb{Z})^\times|} \sum_{t \in (\mathbb{Z}/3^{d_0}\mathbb{Z})^\times} Z_{j_0}(t) > 2^{-23}. \quad (2.2)$$

Now we bring X into the picture, and seek a large enough bound for X such that it is reachable from some j that is large enough for the estimates below. Let $J(t) := 2 \times 10^{10} + 200 \log_2 t$. Proposition 6.7 then tells us the following:

$$\text{Let } t \geq 16^9. \text{ Then every } X \geq 2^{2^{29}} t^3 \text{ is reachable from some } j \geq J(t). \quad (2.3)$$

This requires another hypothesis on t :

$$t \geq 16^9. \quad (\text{H3})$$

We now provide an estimate to connect the weighted density to the terminal weighted density. Precisely, for $j \geq 6 \times 10^6$ and X reachable from j , according to Proposition 6.9,

$$|Z_{j,X}^{\text{term}}(t) - Z_j(t)| \leq 2^{87536} j^{\frac{9}{2}} 1.01^{-j}. \quad (2.4)$$

And finally an inequality to connect the terminal weighted density to the total weight: for any j , any X reachable from j , and any integer m with $2^{80000} \leq m \leq \frac{1}{4}\delta_j$, (6.8) provides us with the bound

$$Z_{j,X}^{\text{term}}(t) \leq \frac{32}{3} m^{\frac{4097}{2048}} \beta_X(t) + 2^{87538} m^{-\frac{3}{2}} t. \quad (2.5)$$

2.5.2. *Final proof.* We now combine the five estimates above. Firstly from the estimate on the average weighted density (2.2), we know that there is some unit residue $y_0 \in (\mathbb{Z}/3^{d_0}\mathbb{Z})^\times$ with

$$Z_{j_0}(y_0) > 2^{-23},$$

which leads to our final hypothesis on t :

$$t \equiv y_0 \pmod{3^{d_0}} \quad (\text{H4})$$

Now following the reachability estimate (2.3), we let $X \geq 2^{2^{29}} t^3$, with $j \geq J(t)$ so that X is reachable from j . This lets us use the estimate on propagation error (2.1), which tells us that the errors from generation j_0 onwards (Lemma 6.3) sum to less than $2^{-23} - 2^{-25}$, so that we can obtain

$$Z_j(t) \geq 2^{-25},$$

at which point the estimate on terminal weighted density (2.4) gives us

$$Z_{j,X}^{\text{term}}(t) \geq 2^{-26}.$$

Now the inequality linking terminal weighted density with total weight (2.5) has an additional parameter m . After numerical optimization we can choose $m = \left\lceil 2^{80000} t^{\frac{2}{3}} \right\rceil$, and this gives us

$$\beta_X(t) \geq 2^{-160071} t^{-1.34},$$

which we can substitute back into the reduction in Section 2.2.5 to obtain the count we desired, namely that

$$\#\{s \in [X, 32X) \text{ maps to } t\} \geq 2^{-160071} t^{-2.34} X,$$

subject to the hypotheses (H1), (H2), (H3), (H4). This is Proposition 7.2.

2.5.3. Discharging the hypotheses: choosing a seed. Recall the hypotheses require that $t \geq 16^9$ is odd, t is not part of a cycle, and t has a specific residue modulo 3^{d_0} . We now describe the process of removing those hypotheses, which mirrors Proposition 7.4. Starting with an arbitrary t with $3 \nmid t$, we will obtain a seed t_1 that satisfies the hypotheses above, with some constraints between the seed t_1 and the original target t .

The seed t_1 will be chosen from $\frac{2^a t - 1}{3}$, so that t_1 will reach t with only one odd step at the beginning, and the rest of the steps will be the even step where the value is divided by 2. This discharges t_1 being odd (Hypothesis (H1)). If t is odd, then t_1 will reach t in one Syracuse step; otherwise t_1 will still reach the odd part of t in one Syracuse step.

Valid choices of a that make t_1 an integer will always be either all the even numbers or all the odd numbers. By modular arithmetic, these choices of t_1 meet every residue class modulo any power of 3 (Lemma 7.3). This discharges the residue requirement (Hypothesis (H4)).

The bound requirement (Hypothesis (H3)) is discharged by simply choosing a larger a , which leaves us with the requirement that t_1 cannot be part of a cycle (Hypothesis (H2)), which is discharged by choosing *two* potential candidates for t_1 . They both map to (the odd part of) t in one step under the Syracuse map, so they cannot both be part of a cycle.

The resulting seed t_1 can be bounded in terms of t using Proposition 7.4, which says that

$$t_1 \leq D_A t.$$

2.5.4. Final proof for general t . Using Section 2.5.2 for the seed t_1 , we deduce that for $X \geq 2^{2^{29} t_1^3}$, we have

$$\#\{s \in [X, 32X) \text{ maps to } t_1\} \geq 2^{-160071} t_1^{-2.34} X.$$

Translating this back to t , using the bound $t_1 \leq D_A t$, gives us that for $X \geq 2^{t^{3D_A}}$, we have

$$\#\{s \in [X, 32X) \text{ maps to } t\} \geq t^{-2.34} D_A^{-1} X.$$

The exact version of this bound is (7.3). Note that there is some abuse of notation here considering that D_A is a triple exponential and hence absorbs many smaller operations. This also covers replacing X with $\frac{1}{32}X$ to align with the main theorem counting numbers $< X$.

2.6. The bound for the number of steps. What remains is control on the length of the words involved, as well as their degrees (Section 8). This follows from the construction since at each generation we pick the words at their first crossing with the line, which puts constraints on their expected lengths and degrees.

Recall that the heuristic number of steps is $3.476 \log \frac{n}{t}$ accelerated steps and $10.428 \log \frac{n}{t}$ ordinary steps. The extra slack in the main theorem allows the bounds on step count to hold. We also use the fact that the seed t_1 is chosen to reach (the odd part of) the target t in one accelerated step, which provides a simple bound for also the number of ordinary steps for this part.

3. WORDS AND RESIDUE DENSITIES

This section makes the objects of Section 2 precise: valid words, their weights and offsets, the reduction to a lower bound for the total weight, and the residue density with its branch transfers.

A *word* $w = (a_1, \dots, a_d)$ is a finite sequence of positive integers, of length $|w| = d$. It describes a walk backwards along the Syracuse map S : each letter a picks the preimage $(2^a n - 1)/3$ of n under S , as in Section 2.2.1. Starting from a positive odd n , the *inverse trajectory* of w is

$$x_0 = n, \quad x_i = \frac{2^{a_i} x_{i-1} - 1}{3} \quad (1 \leq i \leq |w|).$$

The word is *valid at n* if all x_i are positive odd integers. Then $3x_i + 1 = 2^{a_i} x_{i-1}$ with x_{i-1} odd, so $S(x_i) = x_{i-1}$: the Collatz map T takes x_i to x_{i-1} in one step $x \mapsto 3x + 1$ followed by a_i halvings.

Hence the *source* $\text{src}(w, n) = x_{|w|}$ reaches n in $|w|$ accelerated steps and $|w| + \deg(w)$ ordinary steps, where

$$\begin{aligned} \deg_i(w) &= a_1 + \cdots + a_i, & \deg(w) &= \deg_{|w|}(w), \\ \text{wt}(w) &= 3^{|w|} 2^{-\deg(w)}, & \text{off}(w) &= \sum_{i=1}^{|w|} 3^{i-1} 2^{-\deg_i(w)} \end{aligned}$$

are the *partial degrees*, the *degree*, the *weight* and the *offset* of w . Iterating $x_{i-1} = (3x_i + 1)/2^{a_i}$ gives the basic identity

$$n = \text{wt}(w) \text{src}(w, n) + \text{off}(w), \quad (3.1)$$

which defines $\text{src}(w, n) = (n - \text{off}(w)) / \text{wt}(w)$ for every word. For a concatenation vw ,

$$\text{wt}(vw) = \text{wt}(v) \text{wt}(w), \quad \text{off}(vw) = \text{off}(v) + \text{wt}(v) \text{off}(w). \quad (3.2)$$

For example, $w = (3, 2)$ at $n = 5$ gives $x_1 = 13$ and $x_2 = 17$, and indeed $\frac{9}{32} \cdot 17 + \frac{7}{32} = 5$.

Remark 3.1 (Validity is a congruence; cf. [8, Section 2], [10, Lemma 2.1]). A word w is valid at a positive odd n if and only if $n \equiv \text{off}(w) \pmod{3^{|w|}}$. Indeed $2^{\deg(w)} \text{off}(w)$ is an integer, so validity and (3.1) give the congruence. Conversely, if it holds then $x_{|w|}$ is an integer; the forward relations put every x_i in $\mathbb{Z}[1/2]$ and the backward relations put it in $\mathbb{Z}[1/3]$, so every x_i is an integer, odd because $2^{a_i} x_{i-1} - 1$ is odd, and positive.

3.1. Distinct sources and the reduction. To count sources through the weights of their words, distinct words must have distinct sources; this holds unless the target lies on a cycle. A positive odd integer y is *S-periodic* if $S^{\circ j}(y) = y$ for some $j \geq 1$, where $f^{\circ k}$ denotes the k -fold iterate of a map f .

Lemma 3.2 (Periodic points). *Let y, y' be positive odd integers.*

- (1) *If y and y' are S-periodic and $S(y) = S(y')$, then $y = y'$.*
- (2) *If y is not S-periodic, then for every positive odd n there is at most one $r \in \mathbb{N}$ with $S^{\circ r}(n) = y$.*

Proof. For (1), let $S^{\circ j}(y) = y$ and $S^{\circ j'}(y') = y'$ with $j, j' \geq 1$; then $r = jj'$ satisfies $S^{\circ r}(y) = y$ and $S^{\circ r}(y') = y'$, so $y = S^{\circ(r-1)}(S(y)) = S^{\circ(r-1)}(S(y')) = y'$. For (2), if $S^{\circ r}(n) = y = S^{\circ r'}(n)$ with $r < r'$, then $S^{\circ(r'-r)}(y) = y$. $\square$

Lemma 3.3 (Distinct sources). *Let n be a positive odd integer that is not S-periodic. Then distinct words valid at n have distinct sources.*

Proof. Let w be valid at n with source s . Reading its inverse trajectory backwards from s , the value before a value x' is $(3x' + 1)/2^a$, where a is the letter between them, and this is an odd integer only for $a = \text{ord}_2(3x' + 1)$, where $\text{ord}_p(y)$ is the exponent of the prime p in y . So $|w|$ and s determine w , and $S^{\circ |w|}(s) = n$. By Lemma 3.2(2), s determines $|w|$. $\square$

For real $X > 0$ and a positive odd n , let $\beta_X(n)$ be the total weight of the words valid at n whose source lies in $[X, 32X]$. If n is not S-periodic, these sources are distinct by Lemma 3.3, and each such word has $\text{wt}(w) \leq n / \text{src}(w, n) \leq n/X$ by (3.1), since $\text{off}(w) \geq 0$. Hence

$$\#\{\text{src}(w, n) : w \text{ valid at } n, \text{src}(w, n) \in [X, 32X]\} \geq \frac{X}{n} \beta_X(n). \quad (3.3)$$

3.2. The random model and the residue density. By Remark 3.1, the total weight of the words of length d valid at n depends only on n modulo 3^d ; the random model computes it as a function on residues. Write $G_d = \mathbb{Z}/3^d\mathbb{Z}$, call d its *level*, and write $[x]_d \in G_d$ for the reduction of $x \in \mathbb{Z}[1/2]$; a function on G_d is evaluated at such x through $[x]_d$. Take independent letters with $\mathbb{P}(a_i = a) = 2^{-a}$ for $a \geq 1$, and call W the resulting random word; a word w occurs as a prefix of W with probability $2^{-\deg(w)}$. Let the *offset law* μ_d be the probability mass function on G_d of the offset residue $[\text{off}(W_{\leq d})]_d$ of the first d letters, and put the *residue density*

$$\rho_d(y) = 3^d \mu_d(y). \quad (3.4)$$

Write $\langle f \rangle_d = 3^{-d} \sum_{y \in G_d} f(y)$ for $f : G_d \rightarrow \mathbb{C}$, and $\text{pr}_{d,m} : G_d \rightarrow G_m$ ($m \leq d$) for reduction; then $\langle f \circ \text{pr}_{d,m} \rangle_d = \langle f \rangle_m$ and $\langle \rho_d \rangle_d = 1$. Reduction modulo 3^m sends μ_d to μ_m (*projectivity*). For $d \geq 1$ the law lives on units, since $\text{off}(W_{\leq d}) \equiv 2^{-a_1} \pmod{3}$. By Remark 3.1,

$$\sum_{|w|=d} \text{wt}(w) \mathbf{1}[w \text{ valid at } n] = \rho_d(n), \quad (3.5)$$

where $\mathbf{1}[E]$ is 1 if the statement E holds and 0 otherwise.

To average quantities such as $\rho_k(\text{src}(w, n))$ over the residue of n , we transport functions along words as in [8, Section 2]. For a word w , the map $\iota_w : G_k \rightarrow G_{k+|w|}$, $\iota_w(y) = \text{wt}(w)y + \text{off}(w)$, is well-defined and injective, since its linear coefficient is $3^{|w|}$ times a unit. For f on G_k , the *branch transfer* $\mathcal{T}_w f$ on $G_{k+|w|}$ equals $\text{wt}(w)f(y)$ at $\iota_w(y)$ and zero off the image. By [8, Lemma 2.1],

$$\langle \mathcal{T}_w f \rangle_{k+|w|} = 2^{-\deg(w)} \langle f \rangle_k, \quad \langle |\mathcal{T}_w f| \rangle_{k+|w|} = 2^{-\deg(w)} \langle |f| \rangle_k. \quad (3.6)$$

For positive odd n , Remark 3.1 and injectivity of ι_w give

$$(\mathcal{T}_w f)(n) = \begin{cases} \text{wt}(w)f(\text{src}(w, n)), & w \text{ valid at } n, \\ 0, & \text{otherwise.} \end{cases} \quad (3.7)$$

By (3.2), $\mathcal{T}_w \rho_k$ is $3^{k+|w|}$ times the law of the offset residue of $W_{\leq k+|w|}$ restricted to the event that W begins with w .

4. MIXING AND A BOUNDED SUBDENSITY

This section proves the mixing lemma of Section 2 and the bounded subdensity used by the census (Proposition 4.1), assuming the Fourier bound (4.2) below, and derives the transfer estimate (Lemma 4.8). Here n denotes a level. Set $n_0 = 2^{80000}$ and

$$A_* = \frac{10241}{4096}, \quad \alpha = \frac{4097}{2048}, \quad A_* - \frac{\alpha}{2} = \frac{3}{2}. \quad (4.1)$$

The Fourier bound below decays with exponent A_* and the bounded subdensity is at most $8m^\alpha$; the relation $A_* - \alpha/2 = 3/2$ gives the exponent $3/2$ of mixing in (4.13). Theorem 14.6, an explicit form of [10, Proposition 1.17] proved in Sections 9–14, gives for every $n \geq 1$ and every frequency $\xi \in G_n$ prime to 3 the *Fourier bound*

$$|\hat{\mu}_n(\xi)| \leq (160D_*)^{A_*} n^{-A_*}, \quad \hat{\mu}_n(\xi) = \sum_{y \in G_n} \mu_n(y) e^{-2\pi i \xi y / 3^n}, \quad (4.2)$$

where D_* is defined in Appendix A.1. Put

$$C = \frac{477}{20} (3200D_*)^{A_*} + \frac{159}{10}. \quad (4.3)$$

Proposition 4.1 (Fine-scale mixing and subdensity). *The mixing constant C of (4.3) satisfies*

$$87532.002162790 < \log_2 C < \log_2(C+1) < 87532.002162791. \quad (4.4)$$

For integers $n_0 \leq m \leq d$,

$$\langle |\rho_d - \rho_m \circ \text{pr}_{d,m}| \rangle_d \leq C m^{-3/2}. \quad (4.5)$$

For every integer $m \geq n_0$ there is a function $\underline{\rho}_m : G_m \rightarrow \mathbb{R}$, the bounded subdensity, with

$$0 \leq \underline{\rho}_m \leq \rho_m, \quad \underline{\rho}_m \leq 8m^\alpha, \quad \langle \rho_m - \underline{\rho}_m \rangle_m \leq m^{-3/2}. \quad (4.6)$$

4.1. Restricting at the cut. Following [10, Section 6], we *cut* a random word of length n at the first letter where its partial degree exceeds a *cut height* V with 2^V slightly below 3^n : the position of this letter is the *cut point*, the prefix ending at it is the *head*, and the rest is the *tail word*. A head w' with degree l is kept only if $2^l \text{off}(w') < 3^n$; then it has probability $2^{-l} \leq 2^{-V}$, which is 3^{-n} up to a polynomial factor, and it is determined by its offset residue (Lemma 4.6), which bounds the subdensity pointwise. The tail word is fresh, and its offset law inherits the Fourier bound, which gives mixing.

For $n \geq n_0$ set

$$V = n \log_2 3 - \frac{6145}{4096} \log_2 n - 2, \\ k^{\text{lo}} = \left\lfloor V/2 - \frac{1}{4} n^{2049/4096} \right\rfloor, \quad k^{\text{hi}} = \left\lceil V/2 + \frac{1}{4} n^{2049/4096} \right\rceil.$$

For integers $1 \leq k \leq n$ and $l \geq 1$, let the *gate* $\mathcal{G}_{k,l}$ consist of the heads w' of length k satisfying

$$k^{\text{lo}} < k \leq k^{\text{hi}}, \quad \deg_{k-1}(w') \leq V < \deg(w') = l, \quad 2^l \text{off}(w') < 3^n. \quad (4.7)$$

The corresponding *slice* consists of a head followed by any word of length $n - k$. These slices are disjoint, since the partial degrees increase strictly and so pass V only once. Let $\mathcal{R}_n$ be their union.

Since letters have mean 2, the cut point lies near $V/2$, and Lemma 4.5 below shows that the slices carry almost all the probability. The cut point leaves $(k^{\text{lo}}, k^{\text{hi}}]$ only if a partial degree deviates from twice its length by about $n^{2049/4096}$, which the following Chernoff bound controls.

Lemma 4.2 (Geometric tail). *For real $\mathbf{t} < \log 2$ put $\mathbf{M}_{\text{geo}}(\mathbf{t}) = \mathbb{E}e^{\mathbf{t}(a_1-2)} = e^{-\mathbf{t}}/(2 - e^{\mathbf{t}})$. Then $0 \leq \log \mathbf{M}_{\text{geo}}(\mathbf{t}) \leq 8\mathbf{t}^2$ for $|\mathbf{t}| \leq 1/4$, and $\log \mathbf{M}_{\text{geo}}(\mathbf{t}) \leq 9\mathbf{t}^2/8$ for $|\mathbf{t}| \leq 1/32$. For integers $d \geq 1$ and real $x \geq 0$,*

$$\mathbb{P}(|\deg_d(W) - 2d| \geq x) \leq 2 \exp\left(-\min\{x^2/(32d), x/8\}\right). \quad (4.8)$$

Proof. The formula for $\mathbf{M}_{\text{geo}}$ is a geometric series. Its logarithm has value and derivative zero at 0, and its second derivative $2e^{\mathbf{t}}/(2 - e^{\mathbf{t}})^2$ is nonnegative and increasing, at most 6 on $|\mathbf{t}| \leq 1/4$ (as $e^{1/4} \leq 4/3$) and at most $\frac{496}{225} < \frac{9}{4}$ on $|\mathbf{t}| \leq 1/32$ (as $e^{1/32} \leq 32/31$); Taylor's formula gives the bounds. Since $\mathbb{E}e^{\mathbf{t}(\deg_d(W)-2d)} = \mathbf{M}_{\text{geo}}(\mathbf{t})^d$, the Cramér–Chernoff method [1, Section 2.2] for $\pm(\deg_d(W) - 2d)$ with $\mathbf{t} = \min(x/(16d), 1/4)$ gives (4.8), which [8, Section 3.1] states with strict inequality. $\square$

The product condition in (4.7) fails only if the offset of the word before the cut point, scaled by the overshoot, is large. The offset has a $1/x$ tail, and the overshoot is geometric and independent of what precedes the crossing.

Lemma 4.3 (Offset tail). *For every $n \in \mathbb{N}$ and real $x \geq 0$, $\mathbb{P}(\text{off}(W_{\leq n}) \geq x) \leq 7/(x+2)$.*

Proof. First, for real $x \geq 5$,

$$3 \sum_{a \geq 1} \frac{2^{-a}}{2^a x + 5} \leq \frac{1}{x+2}. \quad (4.9)$$

Indeed $\frac{3 \cdot 2^{-a}}{2^a x + 5} = \frac{3 \cdot 4^{-a}}{x} - \frac{15 \cdot 4^{-a}}{x(2^a x + 5)}$; summing with $\sum_a 4^{-a} = \frac{1}{3}$ and keeping the terms $a = 1, 2$ of the subtracted series, it suffices that $\frac{2}{x+2} \leq \frac{15(18x+25)}{16(2x+5)(4x+5)}$, that is $15(18x+25)(x+2) - 32(2x+5)(4x+5) = 14x^2 - 45x - 50 \geq 0$, which holds for $x \geq 5$.

Now induct on n . For $n = 0$ the offset is 0, and $1 \leq 7/2$. For $x < 5$ the bound exceeds 1. For $x \geq 5$, write $W_{\leq n} = (a)v$; then $\text{off}(W_{\leq n}) = 2^{-a}(1 + 3\text{off}(v))$, so the event is $\text{off}(v) \geq x_a = (2^a x - 1)/3 \geq 0$, where v is a random word of length $n - 1$ independent of a . By the induction hypothesis and (4.9), the probability is at most $\sum_a 2^{-a} \frac{7}{x_a + 2} = 21 \sum_a \frac{2^{-a}}{2^a x + 5} \leq \frac{7}{x+2}$. $\square$

Lemma 4.4 (Product tail). *Let $n \geq n_0$. The probability that the cut point i of $W_{\leq n}$ exists and $2^{\deg_i(W)} \text{off}(W_{\leq i}) \geq 3^n$ is at most $\frac{4}{5}n^{-3/2}$.*

Proof. Put $c = \frac{6145}{4096}$, $R = n^c$, $L = \log_2 n$ and $H = \lfloor V \rfloor + 1$, so $2^V = 3^n/(4R)$ and $2^H \leq 3^n/(2R)$. Since $\text{off}(W_{\leq i}) = \text{off}(W_{\leq i-1}) + 3^{i-1}2^{-\deg_i(W)}$ and $i \leq n$, the event forces $2^{\deg_i(W)} \text{off}(W_{\leq i-1}) \geq \frac{2}{3} \cdot 3^n \geq R2^H$, that is, with the overshoot $r = \deg_i(W) - H \geq 0$, $\text{off}(W_{\leq i-1}) \geq R2^{-r}$.

Fix i and r . Given the first $i - 1$ letters, with $\deg_{i-1}(W) \leq V$ and $\gamma = H - \deg_{i-1}(W) \geq 1$, the letter a_i equals $\gamma + r$ with probability $2^{-\gamma-r} = 2^{-r-1}\mathbb{P}(a_i \geq \gamma)$. So the probability that i is the cut point with overshoot r and $\text{off}(W_{\leq i-1}) \geq R2^{-r}$ is 2^{-r-1} times the probability that i is the cut point and $\text{off}(W_{\leq i-1}) \geq R2^{-r}$. The latter events are disjoint in i and imply $\text{off}(W_{\leq n}) \geq R2^{-r}$, so by Lemma 4.3 the probability in question is at most

$$\sum_{r \geq 0} 2^{-r-1} \frac{7}{R2^{-r} + 2} = \sum_{r \geq 0} \frac{7}{2R + 2^{r+2}} \leq \frac{7}{2R}(\log_2 R + 2).$$

For the last step let $r_0 = \lfloor \log_2 R \rfloor \geq 1$: the $r_0 + 1$ terms with $r \leq r_0$ are each below $\frac{7}{2R}$, and the terms with $r > r_0$ are below $7 \cdot 2^{-r-2}$ and sum to $7 \cdot 2^{-r_0-2} < \frac{7}{2R}$.

It remains to show $\frac{7}{2}(cL + 2)n^{-c} \leq \frac{4}{5}n^{-3/2}$, that is $\frac{35}{8}(cL + 2) \leq 2^{L/4096}$, as $c - \frac{3}{2} = \frac{1}{4096}$. Put $s = L/80000 \geq 1$, $K = \frac{35}{8}(80000c + 3) = \frac{35 \cdot 491612288}{32768}$ and $\beta_0 = 2^{625/32}$, so $2^{L/4096} = \beta_0^s$. Then $\frac{35}{8}(cL + 2) \leq sK$, $K < \beta_0$ by the integer comparison $(35 \cdot 491612288)^{32} < 2^{1105}$, and $s\beta_0 \leq \beta_0^s$ since $\log \beta_0 > 1$. $\square$

Lemma 4.5 (Mass retained by the gates). *The probability of $\mathcal{R}_n$ is at least $1 - n^{-3/2}$.*

Proof. The elementary bounds

$$0.33n < k^{\text{lo}} < k^{\text{hi}} < 13n/16, \quad k^{\text{hi}} - k^{\text{lo}} \leq n^{2049/4096} \quad (4.10)$$

follow from $79/50 < \log_2 3 < 317/200$, $\frac{1}{4}n^{2049/4096} \leq n/100$, and $(6145/4096)\log_2 n + 2 < 5\log n \leq 0.3n$. At $d = k^{\text{lo}}$ and at $d = k^{\text{hi}}$, with $x = \frac{1}{2}n^{2049/4096}$, the exponent in (4.8) is at least $n^{1/2048}/128 > 10\log n$; the latter inequality holds at n_0 and persists since $n^{1/2048}/\log n$ increases in n thereon. Thus the two events $\deg_{k^{\text{lo}}}(W) > V$ and $\deg_{k^{\text{hi}}}(W) \leq V$ have total probability at most $4n^{-10}$, and outside them the cut point lies in the required range. By Lemma 4.4 the product condition fails with probability at most $\frac{4}{5}n^{-3/2}$, and $\frac{4}{5} + 4n^{-17/2} \leq 1$. $\square$

For each gate, let $\Gamma_{k,l}(y)$ be the probability mass of its heads with offset residue $y \in G_n$; sums over gates include only nonempty gates. For functions on G_n , $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are the unnormalized ℓ^1 and supremum norms, and $*$ is convolution in counting measure.

Lemma 4.6 (Head bounds). *The head masses satisfy*

$$\|\Gamma_{k,l}\|_\infty \leq 2^{-l}, \quad \sum_y \Gamma_{k,l}(y)^2 \leq 2^{-l}\|\Gamma_{k,l}\|_1, \quad \sum_{k,l} \|\Gamma_{k,l}\|_1 \leq 1, \quad (4.11)$$

and

$$\sum_{k,l} 3^n 2^{-l} < 8n^\alpha. \quad (4.12)$$

Proof. By (4.7), $2^l \text{off}(w')$ is an integer in $[0, 3^n)$ for every head w' in $\mathcal{G}_{k,l}$, so two heads in the same gate with equal offset residues have equal offsets, and offsets determine words of a fixed length [10, Lemma 6.2]. As in [10, Corollary 6.3], there is therefore at most one head per residue, and its

probability is 2^{-l} . This proves the first two bounds. Each slice has probability $\|\Gamma_{k,l}\|_1$, since its tail word is unrestricted, so disjointness proves the third. There are at most $n^{2049/4096}$ head lengths by (4.10). For each length, sum the geometric series over $l \geq \lfloor V \rfloor + 1 > V$:

$$\sum_{k,l} 3^n 2^{-l} \leq n^{2049/4096} 3^n 2^{1-(\lfloor V \rfloor + 1)} < 8n^{(2049+6145)/4096} = 8n^\alpha.$$

□

The *transported tail law* $\mathcal{L}_{k,l}$ is the law on G_n obtained by sending μ_{n-k} , the offset law of the tail word, through $y \mapsto 3^k 2^{-l} y$. By (3.2) and independence of head and tail word, the offset law restricted to the slice of $\mathcal{G}_{k,l}$ is $\Gamma_{k,l} * \mathcal{L}_{k,l}$. Define

$$\underline{\rho}_n = 3^n \sum_{k,l} \Gamma_{k,l} * \mathcal{L}_{k,l}.$$

This is a restriction of the law, so $0 \leq \underline{\rho}_n \leq \rho_n$, and Lemma 4.5 gives $\langle \rho_n - \underline{\rho}_n \rangle_n \leq n^{-3/2}$. Since convolution with a probability measure cannot increase the supremum norm, (4.11) and (4.12) give $\underline{\rho}_n \leq 8n^\alpha$. This proves (4.6) at every level $n \geq n_0$.

4.2. From the Fourier bound to mixing. For $m \leq n$, let $\text{Av}_{n,m}$ average a function on each fiber of $\text{pr}_{n,m}$, and put $\text{Osc}_{n,m}(f) = \sum_y |f(y) - \text{Av}_{n,m} f(y)|$. A frequency $\xi \in G_n$ *factors through* G_m if $3^{n-m} \mid \xi$. By projectivity, $\text{Av}_{n,m} \mu_n = 3^{-n} \rho_m \circ \text{pr}_{n,m}$, so $\text{Osc}_{n,m}(\mu_n) = \langle |\rho_n - \rho_m \circ \text{pr}_{n,m}| \rangle_n$. We bound it first for $n \leq \frac{10}{9}m$ (Tao's regime [10, (6.1)], which leaves the tail word a fixed fraction of the level), and then along a chain of levels.

Lemma 4.7 (Convolution oscillation). *Let $0 \leq m \leq n$ be integers and $\Lambda \geq 0$. Suppose f, ψ are complex-valued functions on G_n and $|\hat{\psi}(\xi)| \leq \Lambda$ for all frequencies that do not factor through G_m . Then*

$$\text{Osc}_{n,m}(f * \psi)^2 \leq \Lambda^2 3^n \sum_y |f(y)|^2.$$

Proof. This is the computation of [10, (6.10)–(6.11)]. Fiber averaging keeps exactly the Fourier coefficients at frequencies factoring through G_m , so Cauchy–Schwarz and Parseval give

$$\|f * \psi - \text{Av}_{n,m}(f * \psi)\|_1^2 \leq \sum_{\xi \text{ not factoring through } G_m} |\hat{f}(\xi) \hat{\psi}(\xi)|^2 \leq \Lambda^2 3^n \sum_y |f(y)|^2.$$

□

Suppose $n_0 \leq m \leq n$ and $9n \leq 10m$. Let $\xi \in G_n$ not factor through G_m , and let 3^r be its order in G_n , so $r \geq m + 1$. By (4.10), $20(k-1) \leq 17n$, and consequently $r - k \geq m + 1 - k \geq n/20$. Since $\mathcal{L}_{k,l}$ is the image of μ_{n-k} under $y \mapsto 3^k 2^{-l} y$, $\widehat{\mathcal{L}_{k,l}}(\xi) = \widehat{\mu}_{n-k}(2^{-l}\xi)$, where $2^{-l}\xi$ is reduced into G_{n-k} and has order 3^{r-k} there; by projectivity this is $\widehat{\mu}_{r-k}(\xi')$ for some $\xi' \in G_{r-k}$ prime to 3. Since the bound (4.2) decreases with the level, it gives the hypothesis of Lemma 4.7 for $\psi = \mathcal{L}_{k,l}$ with $\Lambda = (160D_*)^{A_*} (n/20)^{-A_*} = (3200D_*)^{A_*} n^{-A_*}$. Summing Lemma 4.7 over the gates and using (4.11) and (4.12) gives

$$\sum_{k,l} \text{Osc}_{n,m}(\Gamma_{k,l} * \mathcal{L}_{k,l}) \leq \Lambda \left(\sum_{k,l} 3^n 2^{-l} \right)^{1/2} \left(\sum_{k,l} \|\Gamma_{k,l}\|_1 \right)^{1/2} \leq 3(3200D_*)^{A_*} n^{-3/2}.$$

Fiber averaging is an ℓ^1 contraction, so the part of μ_n outside the slices adds at most twice its probability. Thus

$$\text{Osc}_{n,m}(\mu_n) \leq B_{\text{osc}} n^{-3/2}, \quad B_{\text{osc}} = 3(3200D_*)^{A_*} + 2. \quad (4.13)$$

Telescoping, as in [10, Section 6], reaches all levels; by (4.3), $C = \frac{159}{20} B_{\text{osc}}$. We show $\text{Osc}_{n,m}(\mu_n) \leq C m^{-3/2}$ for $n_0 \leq m \leq n$, which is (4.5), by induction on $n - m$. For $m \geq n_0$, $\bar{n} = \lfloor 10m/9 \rfloor \geq$

$\frac{99999}{90000}m > m$, and the integer inequality $90000^3 179^2 < 159^2 99999^3$ gives $(\bar{n}/m)^{-3/2} < 159/179$. If $n \leq \bar{n}$, use (4.13). Otherwise $\text{Av}_{n,m} = \text{Av}_{n,m} \text{Av}_{n,\bar{n}}$, and projectivity identifies $\text{Osc}_{n,m}(\text{Av}_{n,\bar{n}} \mu_n)$ with $\text{Osc}_{\bar{n},m}(\mu_{\bar{n}})$. The triangle inequality, the induction hypothesis for $(n, \bar{n})$ and (4.13) for $(\bar{n}, m)$ give

$$\text{Osc}_{n,m}(\mu_n) \leq \left(\frac{159}{20} + 1\right) B_{\text{osc}} \bar{n}^{-3/2} < \frac{159}{20} B_{\text{osc}} m^{-3/2} = C m^{-3/2}.$$

The finite arithmetic in Appendix A gives (4.4), completing the proof of Proposition 4.1 from the Fourier bound.

4.3. Transfer through a prefix-disjoint family. Each generation of Section 5 appends one word from a family and then continues randomly; if the family captures almost all of the probability, the residue density after the appended word is close to the residue density of the unconditioned random word. For a finite family $\mathcal{V}$ of words write $\text{mass}(\mathcal{V}) = \sum_{w \in \mathcal{V}} 2^{-\deg(w)}$ for its *mass* and call $1 - \text{mass}(\mathcal{V})$ its *deficit*. The family is *prefix-disjoint* if no word is a proper prefix of another; then the events that W begins with w ($w \in \mathcal{V}$) are disjoint and $\text{mass}(\mathcal{V}) \leq 1$. For $k \in \mathbb{N}$ and $Q \geq k + |w|$, write $\mathcal{T}_{w,Q} \rho_k = \mathcal{T}_w(\rho_k \circ \text{pr}_{Q-|w|,k})$.

Lemma 4.8 (Finite transfer). *Let $d, k, m \in \mathbb{N}$, put $Q = d + k$, and suppose $k \geq n_0$ and $n_0 \leq m \leq Q$. If $\mathcal{V}$ is a finite prefix-disjoint family whose words have length at most d , then*

$$\left\langle \left| \sum_{w \in \mathcal{V}} \mathcal{T}_{w,Q} \rho_k - \rho_m \circ \text{pr}_{Q,m} \right| \right\rangle_Q \leq \frac{C}{k^{3/2}} + \frac{C}{m^{3/2}} + 1 - \text{mass}(\mathcal{V}).$$

Proof. This is [8, (3.7)] with $h = d$, $\ell = m$ and exponent $3/2$, in our normalization of ρ , and his proof applies verbatim: besides (3.6) and the disjointness of the events, it uses mixing only from the levels $Q - |w| \geq k$ and Q down to the levels k and m , where (4.5) holds since $k, m \geq n_0$. $\square$

5. THE FAMILY OF WORDS

This section constructs the families L_j and $L_{j,X}$ of Section 2 and the weighted densities Z_j and $Z_{j,X}^{\text{term}}$. The estimates on scales, first-crossing families and the histogram are proved in Sections 15–16.

5.1. Scales and first crossings. Generation j appends to each word a *block*, a word of length about the *scale* δ_j whose overshoot (defined below) is at most the *cap* R_j . The scales are $\delta_0 = 9$ and $\delta_{j+1} = \delta_j + e_{\delta_j}$, where

$$e_{\delta} = \begin{cases} 1, & 9 \leq \delta \leq 12, \\ \max\{i : \lceil \delta/100 \rceil \leq i \leq \delta, 9 \cdot 3^{\delta} \cdot 16^i < 4^{\delta+1}\}, & 13 \leq \delta \leq 255, \\ \lceil \delta/100 \rceil, & \delta \geq 256. \end{cases}$$

The increments make the sources of generation j exceed 16^{δ_j} (Lemma 5.4). The finite maxima are nonempty, and with $g = 101/100$ one has (Lemma 15.1)

$$g^j \delta_i \leq \delta_{i+j}, \quad \delta_j < 201g^j. \quad (5.1)$$

The integer comparisons $101^{70} > 2 \cdot 100^{70}$ and $101^{69} < 2 \cdot 100^{69}$ give $1/70 < \log_2 g < 1/69$. The caps are $R_j = 16 + \lfloor j/32 \rfloor$; they grow with j so that the losses $2^{-(R_j+1)}$ decay geometrically (Lemma 15.11).

For an integer scale $\delta \geq 9$, put $\ell_{\delta} = \lfloor 3\delta/5 \rfloor$, $d_{\delta}^{\text{lo}} = \delta - \ell_{\delta}$, $d_{\delta}^{\text{hi}} = \delta + \ell_{\delta}$, and, with $B_0(n) = \lceil n \log_2 3 \rceil$, the *shift radius* $r_{\delta} = 2\ell_{\delta} - B_0(\ell_{\delta}) - 2e_{\delta}$. For an integer shift u define the *line*

$$H_{\delta,u}(d) = 2\delta + B_0((d - \delta)_+) - B_0((\delta - d)_+) + u - r_{\delta},$$

where $x_+ = \max(x, 0)$. For $n \in \mathbb{N}$, $n \log_2 3 \leq B_0(n) < n \log_2 3 + 1$, and $B_0(n) = \text{bitlen}(3^n)$ for $n \geq 1$, where the *bit length* $\text{bitlen}(x)$ of $x \in \mathbb{N}$ is the least $k \in \mathbb{N}$ with $x < 2^k$. Hence, for all integers $\delta \geq 9$, u and $d \geq 0$,

$$|H_{\delta,u}(d) - (2\delta + (d - \delta) \log_2 3 + u - r_\delta)| < 1, \quad (5.2)$$

so H_{δ,r_δ} is the line of Section 2 through $(\delta, 2\delta)$ of slope $\log_2 3$, rounded to integers, and $H_{\delta,u}$ is that line raised by $u - r_\delta$. A word w of length d and degree $H_{\delta,u}(d) + o$ has $2^{\deg(w)} 3^{-d} \approx (4/3)^\delta 2^{u-r_\delta+o}$: by (3.1) it enlarges the source by a factor that does not depend on its length. The shift radius is chosen so that, for every shift $0 \leq u \leq 2r_\delta$, most random words cross the line at a length in $(d_\delta^{\text{lo}}, d_\delta^{\text{hi}}]$ (Lemma 15.10); blocks use $u = r_\delta$, the terminal extension all $u \in [0, 2r_\delta]$.

Lemma 5.1 (Shift radius). *For $\delta \geq 9$, $0 \leq r_\delta \leq \frac{23}{100}\delta < \frac{3}{10}\delta$, with $r_\delta = 0$ exactly for $9 \leq \delta \leq 13$. For $\delta \geq 256$, $r_\delta > \frac{229}{1000}\delta - 4$.*

Proof. Since $B_0(\ell_\delta) \geq \ell_\delta \log_2 3 \geq \frac{19}{12}\ell_\delta$ (as $3^{12} > 2^{19}$), $\ell_\delta \leq 3\delta/5$ and $e_\delta \geq \delta/100$ (Lemma 15.1), $r_\delta \leq \frac{5}{12} \cdot \frac{3\delta}{5} - \frac{\delta}{50} = \frac{23\delta}{100}$. For $\delta \geq 256$, $B_0(\ell_\delta) < \ell_\delta \log_2 3 + 1$, $2 - \log_2 3 > \frac{83}{200}$ (Lemma A.1(1)), $\ell_\delta \geq 3\delta/5 - 1$ and $e_\delta \geq \delta/100 + 1$ give

$$r_\delta > \frac{83}{200} \left(\frac{3\delta}{5} - 1 \right) - 1 - \frac{\delta}{50} - 2 = \frac{229}{1000}\delta - \frac{683}{200},$$

which is positive. For the 247 integers $9 \leq \delta \leq 255$, r_δ is an exact integer computation from the definitions; it is nonnegative, and zero exactly for $9 \leq \delta \leq 13$. $\square$

A word w belongs to the *first-crossing family* $\text{FC}_u(\delta, R)$, where $R \in \mathbb{N}$, if $d_\delta^{\text{lo}} < |w| \leq d_\delta^{\text{hi}}$, its partial degrees stay strictly below the line at lengths $d_\delta^{\text{lo}}, \dots, |w| - 1$ (so $|w|$ is the *first crossing*), and $H_{\delta,u}(|w|) \leq \deg(w) \leq H_{\delta,u}(|w|) + R$. Its *overshoot* is $\deg(w) - H_{\delta,u}(|w|) \in [0, R]$. Such a family is finite and prefix-disjoint (Lemma 15.2). The *block family* $\text{FC}(\delta, R)$ of Section 2, from which the blocks are drawn, consists of the words of $\text{FC}_{r_\delta}(\delta, R)$ that satisfy two further restrictions. For $\delta \geq 256$, the *length restriction* puts the length in $[\delta - \ell_\delta^c, \delta + \ell_\delta^c]$, where

$$\ell_\delta^c = \min \left(\left\lceil \sqrt{32\delta \lceil \log_2 \delta \rceil} \right\rceil, \ell_\delta - 1 \right),$$

and $\ell_\delta^c = \ell_\delta$ for $\delta < 256$; it narrows the length spread for the histogram estimate (5.6) and Lemma 8.1 at a deficit polynomially small in δ (Lemma 15.7). At $\delta = 9, 10, 11$, the overshoot is at least $o_\delta(d)$, the least nonnegative integer o with

$$16 \cdot 3^d \cdot 2^{3\delta} < (2^{3\delta} - 1) \cdot 2^{H_{\delta,r_\delta}(d)+o},$$

and $o_\delta(d) = 0$ for $\delta \geq 12$; with the increments e_δ , this makes sources grow fast enough at the first scales.

Lemma 5.2 (Weight, offset and source of a first-crossing word). *Let $\delta \geq 9$, u and $R \geq 0$ be integers, and let $w \in \text{FC}_u(\delta, R)$. Then*

- (1) $2^{-R-1+r_\delta-u}(3/4)^\delta < \text{wt}(w) < 2^{1+r_\delta-u}(3/4)^\delta$;
- (2) $0 < \text{off}(w) < 2^\delta$;
- (3) if $w \in \text{FC}(\delta, R)$, then $16^{e_\delta} \text{wt}(w) < 1 - 2^{-3\delta}$;
- (4) for every odd $n \geq 16^\delta$,

$$2^{u-r_\delta-2}(4/3)^\delta n < \text{src}(w, n) < 2^{R+1+u-r_\delta}(4/3)^\delta n. \quad (5.3)$$

Proof. For (1), $\text{wt}(w) = 3^{|w|} 2^{-\deg(w)}$ and $H_{\delta,u}(|w|) \leq \deg(w) \leq H_{\delta,u}(|w|) + R$; insert (5.2). For (2), as in [8, proof of Proposition 3.1]: every summand of $\text{off}(w)$ is positive and, since the letters are at least 1, the i th summand is at most $3^{i-1} 2^{-i}$; as $|w| \leq d_\delta^{\text{hi}} \leq 8\delta/5$ and $(3/2)^8 < 2^5$, $0 < \text{off}(w) < (3/2)^{|w|} < 2^\delta$.

For (3), $u = r_\delta$, and we distinguish four ranges. If $\delta \in \{9, 10, 11\}$, then $e_\delta = 1$ and $\deg(w) \geq H_{\delta, r_\delta}(|w|) + o_\delta(|w|)$, so the definition of $o_\delta(|w|)$ gives $16 \text{ wt}(w) \leq 16 \cdot 3^{|w|} \cdot 2^{-H_{\delta, r_\delta}(|w|) - o_\delta(|w|)} < 1 - 2^{-3\delta}$. If $\delta = 12$, then $e_\delta = 1$, $d_{12}^{\text{lo}} = 5$, $d_{12}^{\text{hi}} = 19$, and for $d = 6, \dots, 19$ the values $H_{12, r_{12}}(d)$ are

$$14, 16, 17, 19, 20, 22, 24, 26, 28, 29, 31, 32, 34, 36.$$

The fourteen exact integer comparisons $9 \cdot 3^d \cdot 4^{12} \leq 16 \cdot 3^{12} \cdot 2^{H_{12, r_{12}}(d)}$ hold (with equality at $d = 10$). Hence $16 \text{ wt}(w) \leq \frac{256}{9}(3/4)^{12} = 3^{10}/2^{16} < 1 - 2^{-36}$. If $13 \leq \delta \leq 255$, then e_δ belongs to the set defining it, so $16^{e_\delta} < \frac{4}{9}(4/3)^\delta$, and (1) gives $16^{e_\delta} \text{ wt}(w) < \frac{8}{9} < 1 - 2^{-3\delta}$. Finally let $\delta \geq 256$, so $e_\delta = \lceil \delta/100 \rceil \geq 3$ and $\delta > 100(e_\delta - 1)$. The integer inequality $4^{100} > 4096 \cdot 3^{100}$ says $(4/3)^{100} > 16^3$, so $(4/3)^\delta > 16^{3(e_\delta - 1)} \geq 4 \cdot 16^{e_\delta}$, and with (1), $16^{e_\delta} \text{ wt}(w) < \frac{1}{4}(4/3)^\delta \cdot 2(3/4)^\delta = \frac{1}{2}$.

For (4), which is [8, (3.5)] for general u , (2) gives $0 < \text{off}(w) < 2^\delta \leq n/2$, so $\text{wt}(w) \text{ src}(w, n) = n - \text{off}(w) \in (n/2, n)$; insert (1). $\square$

5.2. The selector and the family L_j . Consider the words $w = v_0 \cdots v_{j-1}$ with $v_i \in \text{FC}(\delta_i, R_i)$. Since each $\text{FC}(\delta_i, R_i)$ is prefix-disjoint, such a word determines its *blocks* v_i : if $v_0 \cdots v_{j-1}w = v'_0 \cdots v'_{j-1}w'$ with $v_i, v'_i \in \text{FC}(\delta_i, R_i)$, then one of v_0, v'_0 is a prefix of the other, so $v_0 = v'_0$, and cancelling and inducting gives $v_i = v'_i$ and $w = w'$. For $i \leq j$ write $w_{<i} = v_0 \cdots v_{i-1}$.

By (3.2), the offset of a long word is dominated by its first few blocks. The *five-block selector* bounds those offsets, which improves the histogram estimate (5.6). For a word v and an integer $\mathfrak{r}$, let

$$\text{rd}_\mathfrak{r}(v) = 2^{-\mathfrak{r}} \sum_{i=1}^{|v|} \left\lceil 2^\mathfrak{r} \cdot 3^{i-1} \cdot 2^{-\deg_i(v)} \right\rceil,$$

the offset with each summand rounded up to a multiple of $2^{-\mathfrak{r}}$; rounding makes the selector a condition on integer bins, which Appendix A.4 enumerates exactly. The selector retains a word with first five blocks $v_0, \dots, v_4$ if

$$\begin{aligned} \frac{3}{8} &\leq \text{rd}_9(v_0) + \text{wt}(v_0) \text{rd}_6(v_1) \leq \frac{19}{8}, \\ \text{rd}_6(v_1) &\leq 16, \quad \text{rd}_3(v_2) \leq 16, \quad \text{rd}_0(v_3) \leq 64, \quad \text{rd}_{-3}(v_4) \leq 256. \end{aligned}$$

The *selected words* form the set $\bar{L}_j$ of words $v_0 \cdots v_{j-1}$ as above that pass the selector when $j \geq 5$; $\bar{L}_0$ consists of the empty word. The selector involves only the first five blocks, so for $j \geq 5$ every word of $\bar{L}_j$ extended by a word of $\text{FC}(\delta_j, R_j)$ lies in $\bar{L}_{j+1}$, and every word of $\bar{L}_j$ has $\frac{85}{256} \leq \text{off}(w) \leq \frac{637}{256}$ (Lemma 16.2). Its retained fraction of the mass of all five-block words $v_0 \cdots v_4$ is

$$\frac{17295610054959855512665734818947672447921469210720928769975}{32309186679205268427697461585468266195406151955473216860286} > \frac{1}{2}, \quad (5.4)$$

computed in Appendix A.4.

Lemma 5.3 (Selected mass). *For every $j \geq 5$, $\text{mass}(\bar{L}_j) > 2^{-23}$.*

Proof. The degree is additive under concatenation, the families are finite, and the selector constrains only the first five blocks, so

$$\text{mass}(\bar{L}_j) = \text{mass}(\bar{L}_5) \prod_{5 \leq i < j} \text{mass}(\text{FC}(\delta_i, R_i)),$$

and $\text{mass}(\bar{L}_5)$ is $\prod_{i < 5} \text{mass}(\text{FC}(\delta_i, R_i))$ times the retained fraction (5.4), which exceeds $\frac{1}{2}$. All factors lie in $[0, 1]$ (Lemma 15.2). For $j \leq 444$ the product $\prod_{i < j} \text{mass}(\text{FC}(\delta_i, R_i))$ is at least the product over $i \leq 443$, which exceeds $\frac{72}{5} 2^{-25} > 2^{-22}$ (Lemma 15.12); for $j > 444$, Lemmas 15.12 and 15.13 bound it below by $\frac{72}{5} 2^{-25} \cdot \frac{5}{9} = 2^{-22}$. Hence $\text{mass}(\bar{L}_j) > \frac{1}{2} \cdot 2^{-22}$. $\square$

For a positive odd integer t , let $L_j = L_j(t)$ be the set of words of $\bar{L}_j$ valid at t , and for $w \in L_j$ write $s_w = \text{src}(w, t)$ for its source; the empty word has $s_w = t$. A prefix of a valid word is valid, so $w_{<i} \in L_i$ for $w \in L_j$ and $i \leq j$; the block v_i is valid at $s_{w_{<i}}$, with source $s_{w_{<i+1}}$ from there, so by (3.1)

$$s_{w_{<i+1}} = \frac{s_{w_{<i}} - \text{off}(v_i)}{\text{wt}(v_i)} \quad (i < j). \quad (5.5)$$

Lemma 5.4 (Growth of sources). *Let $t \geq 16^9$ be odd, $j \in \mathbb{N}$ and $w \in L_j$. For every $i < j$,*

$$s_{w_{<i}} \geq 16^{\delta_i}, \quad 16^{e_{\delta_i}} s_{w_{<i}} < s_{w_{<i+1}}, \quad \frac{1}{4}(4/3)^{\delta_i} s_{w_{<i}} \leq s_{w_{<i+1}} \leq 2^{R_i+1}(4/3)^{\delta_i} s_{w_{<i}}.$$

In particular $s_w \geq 16^{\delta_j}$.

Proof. We prove $s_{w_{<i}} \geq 16^{\delta_i}$ by induction on i ; the case $i = 0$ is $t \geq 16^9$. If $s_{w_{<i}} \geq 16^{\delta_i}$, then Lemma 5.2(2) gives $0 < \text{off}(v_i) < 2^{\delta_i} \leq 2^{-3\delta_i} s_{w_{<i}}$, and by (5.5) and Lemma 5.2(3),

$$s_{w_{<i+1}} > \frac{(1 - 2^{-3\delta_i}) s_{w_{<i}}}{\text{wt}(v_i)} > 16^{e_{\delta_i}} s_{w_{<i}} \geq 16^{\delta_i + e_{\delta_i}} = 16^{\delta_{i+1}}.$$

The third inequality is (5.3) with $u = r_{\delta_i}$ and $R = R_i$, and the induction up to $i = j$ gives $s_w \geq 16^{\delta_j}$. $\square$

The *weighted density* is

$$Z_j(t) = \sum_{w \in L_j} \text{wt}(w) \rho_{k_j}(s_w), \quad k_j = \lfloor \delta_j/4 \rfloor.$$

By (3.5), it weighs each word of L_j by the total weight of its valid continuations of length k_j . The *histogram estimate* (Proposition 16.5) says that the weighted sources do not concentrate on few residues: for $j \geq 10000$, $0 \leq q \leq 2\delta_j$, any $f : G_q \rightarrow \mathbb{R}$, and any subset of L_j ,

$$\sum_w \text{wt}(w) |f(s_w)| \leq \frac{4075}{56} \frac{277}{128} j^{5/2} g^{j/2} \langle |f| \rangle_q. \quad (5.6)$$

5.3. The terminal extension. To bring the sources into $[X, 32X]$, each $w \in L_j$ receives a last word whose line is shifted according to s_w . Let $X > 0$ be real. For an integer scale $\delta \geq 9$ and a real $n > 0$, set, as in [8, Section 5],

$$X_\delta^{\text{lo}}(n) = \frac{(4/3)^\delta n}{4 \cdot 2^{r_\delta}}, \quad X_\delta^{\text{hi}}(n) = 2^{2r_\delta} X_\delta^{\text{lo}}(n).$$

By (5.3), every $v \in \text{FC}_u(\delta, R)$ and every odd $n \geq 16^\delta$ satisfy $2^u X_\delta^{\text{lo}}(n) < \text{src}(v, n) < 2^{R+3+u} X_\delta^{\text{lo}}(n)$. The real X is *reachable* at generation j if $X_{\delta_j}^{\text{lo}}(s_w) < X \leq X_{\delta_j}^{\text{hi}}(s_w)$ for every $w \in L_j$. For $w \in L_j$ let the *terminal shift* u_w be the least integer $u \geq 0$ with $2^u X_{\delta_j}^{\text{lo}}(s_w) \geq X$. If X is reachable at j , then

$$X \leq 2^{u_w} X_{\delta_j}^{\text{lo}}(s_w) < 2X, \quad 0 \leq u_w \leq 2r_{\delta_j}.$$

For $R \in \mathbb{N}$, let $\text{FC}_X^{\text{term}}(w, \delta_j, R)$ be the words $v \in \text{FC}_{u_w}(\delta_j, R)$ that are valid at s_w and satisfy the *mask*

$$2|v| \leq H_{\delta_j, u_w}(|v|) + \lfloor \delta_j/8 \rfloor, \quad (5.7)$$

Mazur's terminal restriction [8, (5.2)], which bounds the length of v by its degree and is used only in Lemma 8.2(3). The *terminal extension* $L_{j,X}$ consists of the words wv with $w \in L_j$ and

$v \in \text{FC}_X^{\text{term}}(w, \delta_j, R_j)$; its sources are at least X . Define the *terminal weighted density* and the *terminal weight*, which uses only the cap 1 (the census in Section 6 reduces the cap R_j to 1),

$$\begin{aligned} Z_{j,X}^{\text{term}}(t) &= \sum_{w \in L_j} \text{wt}(w) \sum_{v \in \text{FC}_X^{\text{term}}(w, \delta_j, R_j)} \text{wt}(v) \rho_{k_j}(\text{src}(v, s_w)), \\ \beta_{j,X}^{\text{term}}(t) &= \sum_{w \in L_j} \text{wt}(w) \sum_{v \in \text{FC}_X^{\text{term}}(w, \delta_j, 1)} \text{wt}(v), \end{aligned}$$

and let $\text{Src}_{j,X}(t)$ be the set of sources $\text{src}(v, s_w)$ of the words wv , $w \in L_j$, $v \in \text{FC}_X^{\text{term}}(w, \delta_j, 1)$. If $t \geq 16^9$ and X is reachable at j , these sources lie in $[X, 32X]$ [8, Lemma 6.1], since $s_w \geq 16^{\delta_j}$ (Lemma 5.4): the factor 2 comes from the rounding of u_w , and the factor 16 from overshoot at most 1 and the slack 3 above. If moreover t is not S -periodic, (3.3) gives $\beta_{j,X}^{\text{term}}(t) \leq \beta_X(t)$ and

$$X \beta_{j,X}^{\text{term}}(t) \leq t \# \text{Src}_{j,X}(t). \quad (5.8)$$

6. KEY ESTIMATES

This section proves the key estimates of Section 2: propagation (Proposition 6.1), the average at generation j_0 (Proposition 6.5), reachability (Proposition 6.7), the terminal error (Proposition 6.9) and the census inequality (6.8). Apart from the census, which uses the bounded subdensity, the arguments are Mazur's [8, Proposition 3.1, Sections 4–6 and 8]. Throughout, t is a positive odd integer.

6.1. Propagation. Passing from Z_j to Z_{j+1} appends a block and moves the level from k_j to k_{j+1} ; the resulting transfer errors (Lemma 4.8) are summed over L_j by the histogram estimate. By Lemmas 15.1 and 15.11, for $j \geq 6000000$ the level k_j is large enough for Proposition 4.1 and the block family has small deficit:

$$k_j \geq \max(n_0, 47g^j), \quad 1 - \text{mass}(\text{FC}(\delta_j, R_j)) \leq (1024 + 2^{-16})g^{-3j/2}. \quad (6.1)$$

Hence the three errors in Lemma 4.8 are small: every level $m \geq k_j$ has $Cm^{-3/2} \leq C47^{-3/2}g^{-3j/2}$, and

$$2C \cdot 47^{-3/2} + 2^{121} + 1024 + 2^{-16} < \frac{28}{1000}(C + 1), \quad (6.2)$$

because $2000^2 < 27^2 \cdot 47^3$ gives $2 \cdot 47^{-3/2} < 27/1000$, and $C > 2^{87532}$ gives $C/1000 > 2^{121} + 1024 + 2^{-16}$. The term 2^{121} covers the terminal deficit of Proposition 15.14. Put

$$F = \frac{9}{2}(C + 1), \quad j_0 = \left\lceil \frac{23 + \log_2(3(C + 1)) + 535/8}{\log_2 g} \right\rceil = 6103914. \quad (6.3)$$

By (4.4), $\log_2 F < 87534.2$. The generation j_0 is chosen so that the propagation errors from j_0 on are small compared with 2^{-23} (Lemma 6.3).

Proposition 6.1 (Propagation). *For every $j \geq 6000000$,*

$$|Z_{j+1}(t) - Z_j(t)| \leq Fj^{5/2}g^{-j}. \quad (6.4)$$

Proof. This is [8, Proposition 3.1] with explicit constants. Put $Q = d_{\delta_j}^{\text{hi}} + k_{j+1}$. Since $\delta_j \geq 256$ (Lemma 15.1), $\delta_{j+1} \leq (51/50)\delta_j$, so $Q \leq (8/5 + 51/200)\delta_j < 2\delta_j$. Define on G_Q

$$\mathcal{D}_j = \sum_{v \in \text{FC}(\delta_j, R_j)} \mathcal{T}_{v,Q} \rho_{k_{j+1}} - \rho_{k_j} \circ \text{pr}_{Q,k_j}.$$

Since $j \geq 5$, the words of L_{j+1} are exactly the words wv with $w \in L_j$ and $v \in \text{FC}(\delta_j, R_j)$ valid at s_w , so (3.7) gives

$$Z_{j+1}(t) - Z_j(t) = \sum_{w \in L_j} \text{wt}(w) \mathcal{D}_j(s_w).$$

Lemma 4.8 at the levels $k_{j+1}, k_j \geq k_j$, (6.1) and (6.2) give $\langle |\mathcal{D}_j| \rangle_Q < \frac{28}{1000}(C+1)g^{-3j/2}$. Apply (5.6), using $j/2 - 3j/2 = -j$ and $\frac{4075}{56} \frac{277}{128} \frac{28}{1000} = \frac{45151}{10240} < \frac{9}{2}$. $\square$

Lemma 6.2 (Polynomial-geometric tails). *For real $0 \leq \epsilon \leq 9/2$ and integers $i \geq 6000000$, the terms $j^\epsilon g^{-j}$ decrease in $j \geq i$, and*

$$\sum_{j \geq i} j^\epsilon g^{-j} < 162 i^\epsilon g^{-i}.$$

Proof. For $j \geq i$ the ratio of consecutive terms is

$$\left(\frac{j+1}{j}\right)^\epsilon g^{-1} \leq \left(\frac{6000001}{6000000}\right)^{9/2} \frac{100}{101} < \frac{161}{162}.$$

Squared, the last inequality is the integer comparison $6000001^9 \cdot 100^2 \cdot 162^2 < 161^2 \cdot 6000000^9 \cdot 101^2$. Hence the term at j is at most $(161/162)^{j-i} i^\epsilon g^{-i}$, strictly for $j > i$, and the geometric series sums to $162 i^\epsilon g^{-i}$. $\square$

Lemma 6.3 (Propagation tail). *$F \sum_{j \geq j_0} j^{5/2} g^{-j} < \frac{5}{4} \cdot 2^{-24}$. Consequently, if $Z_{j_0}(t) > 2^{-23}$, then $Z_j(t) > \frac{3}{4} \cdot 2^{-24} \geq 2^{-25}$ for every $j \geq j_0$.*

Proof. By the ceiling in (6.3), $j_0 \log_2 g \geq 23 + \log_2(3(C+1)) + \frac{535}{8}$, so $g^{-j_0} \leq 2^{-23-535/8}/(3(C+1))$, and $F = \frac{3}{2} \cdot 3(C+1)$. Lemma 6.2 with $\epsilon = 5/2$ gives

$$F \sum_{j \geq j_0} j^{5/2} g^{-j} < \frac{3}{2} \cdot 162 j_0^{5/2} 2^{-535/8} \cdot 2^{-23}.$$

Raising to the eighth power, $243 j_0^{5/2} 2^{-535/8} < 5/8$ is the integer comparison $243^8 j_0^{20} 8^8 < 5^8 2^{535}$, so the bound is less than $\frac{5}{8} \cdot 2^{-23} = \frac{5}{4} \cdot 2^{-24}$. The consequence follows by summing (6.4) from j_0 to $j-1$. $\square$

6.2. The average at generation j_0 . At generation j_0 the words are short enough that $Z_{j_0}(t)$ depends only on t modulo 3^{d_0} , where $d_0 = 2^{87639}$, so its average over residues can be computed exactly. The comparison of lengths uses the exact numerical enclosure of Appendix A,

$$87639.473776839504865 < 16 + j_0 \log_2 g < 87639.473776839504878. \quad (6.5)$$

Lemma 6.4 (Period of Z_{j_0}). *Every word of $\bar{L}_{j_0}$ has length at most $d_0 - k_{j_0}$. Consequently, as a function of t , $Z_{j_0}(t)$ depends only on $t \bmod 3^{d_0}$, and on G_{d_0}*

$$Z_{j_0} = \sum_{w \in \bar{L}_{j_0}} \mathcal{T}_{w, d_0} \rho_{k_{j_0}}.$$

Proof. Block j has length at most $d_{\delta_j}^{\text{hi}} \leq \frac{8}{5} \delta_j$. By (5.1), $\delta_j \leq g^{-(j_0-j)} \delta_{j_0}$ for $j < j_0$, so $\sum_{j < j_0} \delta_j < 100 \delta_{j_0}$, and with $k_{j_0} \leq \delta_{j_0}/4$,

$$k_{j_0} + \sum_{j < j_0} d_{\delta_j}^{\text{hi}} < \left(\frac{1}{4} + \frac{8}{5} \cdot 100\right) \delta_{j_0} = \frac{641}{4} \delta_{j_0} < \frac{641 \cdot 201}{4} g^{j_0} < 2^{15} g^{j_0},$$

by $\delta_{j_0} < 201 g^{j_0}$ and $641 \cdot 201 = 128841 < 2^{17}$. By (6.5), $15 + j_0 \log_2 g < 87638.48 < \log_2 d_0$. For $w \in \bar{L}_{j_0}$, (3.7) shows that the term of w in the displayed sum, at a positive odd t , is $\text{wt}(w) \rho_{k_{j_0}}(s_w)$ if w is valid at t and zero otherwise, so the sum is $Z_{j_0}(t)$. $\square$

Proposition 6.5 (Average at generation j_0). *The average of Z_{j_0} over the units of G_{d_0} equals $\frac{3}{2} \text{mass}(\bar{L}_{j_0})$, and in particular exceeds 2^{-23} .*

Proof. We follow [8, Proposition 4.2]. By Lemma 6.4, (3.6) and $\langle \rho_k \rangle_k = 1$, the average of Z_{j_0} over all of G_{d_0} is $\sum_{w \in \bar{L}_{j_0}} 2^{-\deg(w)} = \text{mass}(\bar{L}_{j_0})$. Every $w \in \bar{L}_{j_0}$ is nonempty, so the image of ι_w consists of units: its offset is 2^{-a_1} modulo 3 and its linear coefficient $\text{wt}(w)$ vanishes modulo 3. Thus Z_{j_0} vanishes off the units, which are $2/3$ of G_{d_0} , and Lemma 5.3 gives $\frac{3}{2} \text{mass}(\bar{L}_{j_0}) > 2^{-23}$. $\square$

6.3. Reachability. Let $J(t) = 2 \cdot 10^{10} + 200 \log_2 t$, a generation large enough for Lemmas 6.6 and 7.1. The *generation- j ranges* $(X_{\delta_j}^{\text{lo}}(s_w), X_{\delta_j}^{\text{hi}}(s_w)]$, $w \in L_j$, with *lower ends* $X_{\delta_j}^{\text{lo}}(s_w)$ and *upper ends* $X_{\delta_j}^{\text{hi}}(s_w)$, intersect in the set of X reachable at j ; consecutive intersections overlap and move to infinity. The overlap holds because the sources of L_j differ by a factor at most 2^{R_Σ} (Lemma 16.1), where $R_\Sigma = \sum_{i < j} (R_i + 3)$, suppressing j . Since $R_i + 3 = 19 + \lfloor i/32 \rfloor$, $\frac{1}{56} - \frac{1}{64} = \frac{1}{448}$ and $448(19j+1) < j^2$ for $j \geq 10000$,

$$R_\Sigma + 1 \leq 19j + \frac{j(j-1)}{64} + 1 < \frac{j^2}{56} \quad (j \geq 10000). \quad (6.6)$$

Lemma 6.6 (Consecutive generations overlap). *Let $t \geq 16^9$ be odd and $j \geq 2 \cdot 10^{10}$. For all $w' \in L_j$ and $w \in L_{j+1}$, $X_{\delta_{j+1}}^{\text{lo}}(s_w) < X_{\delta_j}^{\text{hi}}(s_{w'})$.*

Proof. By Lemma 15.1, $\delta_j \geq 8000j^2$ and $\delta_j \geq 256$, so $e_{\delta_j} \leq \delta_j/100 + 1$ and $\delta_j \leq \delta_{j+1} \leq \frac{101}{100}\delta_j + 1$. Lemmas 5.4 and 16.1 and (6.6) give $s_w \leq 2^{R_j+1}(4/3)^{\delta_j} s_{w < j} \leq 2^{R_j+1+R_\Sigma}(4/3)^{\delta_j} s_{w'}$. Since $X_{\delta_j}^{\text{hi}}(s_{w'}) = 2^{r_{\delta_j}}(4/3)^{\delta_j} s_{w'}/4$, it suffices that $(4/3)^{\delta_{j+1}} 2^{R_j+1+R_\Sigma} < 2^{r_{\delta_j}+r_{\delta_{j+1}}}$. The inequality $2^{19} \leq 3^{12}$ gives $(4/3)^\delta \leq 2^{5\delta/12}$, so $(4/3)^{\delta_{j+1}} \leq 2^{505\delta_j/1200+5/12}$. By Lemma 5.1, $r_{\delta_j} + r_{\delta_{j+1}} \geq \frac{458}{1000}\delta_j - 8$. As $\frac{458}{1000} - \frac{505}{1200} = \frac{223}{6000} > \frac{37}{1000}$, it suffices that $R_j + R_\Sigma + 10 < \frac{37}{1000}\delta_j$, and indeed $R_j + R_\Sigma + 10 < 16 + \frac{j}{32} + \frac{j^2}{56} + 10 < 296j^2 \leq \frac{37}{1000}\delta_j$. $\square$

Proposition 6.7 (Reachability). *Let $t \geq 16^9$ be odd. Every real $X \geq 2^{2^{29}t^3}$ is reachable at some generation $j \geq J(t)$.*

Proof. We follow [8, Lemma 6.2] in the effective form of [8, Section 8.5]. Put $\text{gr}(\delta) = 2\delta - \lfloor 19\delta/12 \rfloor$, $\mathfrak{l}_0 = 0$ and $\mathfrak{l}_{j+1} = \mathfrak{l}_j + R_j + 1 + \text{gr}(\delta_j)$, Mazur's $t(u)$ and $E(j)$ [8, Section 8.2] with this paper's caps and scales. The inequality $2^{19} \leq 3^{12}$ gives $(4/3)^\delta \leq 2^{\text{gr}(\delta)}$, so Lemma 5.4 and induction give $s_w \leq 2^{\mathfrak{l}_j} t$ for $w \in L_j$. Let $J' = \lceil J(t) \rceil$ and $\delta = \delta_{J'}$. By Lemma 5.1 and $\text{gr}(\delta) \leq \frac{5\delta}{12} + 1$, $X_\delta^{\text{hi}}(n) = 2^{r_\delta}(4/3)^\delta n/4 \leq 2^{r_\delta+\text{gr}(\delta)-2} n$ with $r_\delta + \text{gr}(\delta) - 2 \leq \frac{97\delta}{150} - 1 < \lfloor 2\delta/3 \rfloor$. Hence every upper end at generation J' is at most $2^B t$, where $B = \mathfrak{l}_{J'} + \lfloor 2\delta_{J'}/3 \rfloor$. Using $R_j + 1 = 17 + \lfloor j/32 \rfloor$, $\text{gr}(\delta) \leq 2\delta$ and (5.1),

$$B < 17J' + \frac{J'^2}{64} + 2 \sum_{j < J'} \delta_j + \frac{2}{3}\delta_{J'} < 17J' + \frac{J'^2}{64} + 40334g^{J'}.$$

Since $\log_2 g < 1/69$ and $J' < J(t) + 1$, $g^{J'} < 2^{(2 \cdot 10^{10} + 1)/69} t^{200/69} < 2^{2.9 \cdot 10^8} t^3$; also $17J' + J'^2/64 < J'^2 < 2^{72} t^2$. Hence $B + \log_2 t < 2^{2.9 \cdot 10^8 + 17} t^3 < 2^{2^{29}} t^3$, and every upper end at generation J' is below X .

For $i \geq J'$ and $w \in L_i$, Lemma 5.4 gives $s_w \geq 16^{\delta_i}$, so $X_{\delta_i}^{\text{hi}}(s_w) \geq 2^{\delta_i} \rightarrow \infty$. Let $j \geq J'$ be least such that every upper end at generation j is at least X . If $j = J'$, then $L_{J'}$ is empty and X is reachable vacuously. Otherwise some upper end at generation $j-1 \geq J' \geq 2 \cdot 10^{10}$ is below X , and by Lemma 6.6 every lower end at generation j is below it, hence below X . $\square$

6.4. The terminal error. The terminal extension appends to w words of $\text{FC}_{u_w}(\delta_j, R_j)$, a family that depends on w only through its terminal shift; grouping L_j by shift reduces the terminal error to transfer errors, at the cost of a factor equal to the number of shifts.

Lemma 6.8 (Few terminal shifts). *Let $t \geq 16^9$ be odd, $j \geq 10000$ and $X > 0$. Then the terminal shifts u_w , $w \in L_j$, take fewer than $j^2/56$ values.*

Proof. We follow [8, proof of Proposition 5.2]. Assume $L_j \neq \emptyset$ and let $w_{\min} \in L_j$ have the least source. Since $X_{\delta_j}^{\text{lo}}(n)$ is positive and linear in n , u_w is nonincreasing in s_w , so $u_w \leq u_{w_{\min}}$. By Lemma 16.1, $s_w \leq 2^{R_\Sigma} s_{w_{\min}}$, so $2^{u_w + R_\Sigma} X_{\delta_j}^{\text{lo}}(s_{w_{\min}}) \geq X$ and $u_{w_{\min}} \leq u_w + R_\Sigma$. Thus all shifts lie in a set of at most $R_\Sigma + 1 < j^2/56$ integers, by (6.6). $\square$

Proposition 6.9 (Terminal error). *Let $t \geq 16^9$ be odd. For $j \geq 6000000$ and every X reachable at j ,*

$$|Z_{j,X}^{\text{term}}(t) - Z_j(t)| \leq 2Fj^{9/2}g^{-j},$$

and the right side is less than 2^{-1500} for $j \geq 10^7$.

Proof. We follow [8, Proposition 5.2]. Put $Q = d_{\delta_j}^{\text{hi}} + k_j < 2\delta_j$, and group the words of L_j by their terminal shift; by Lemma 6.8 there are fewer than $j^2/56$ groups. For each shift $0 \leq u \leq 2r_{\delta_j}$, Proposition 15.14 gives

$$\begin{aligned} 1 - \text{mass}(\text{FC}_u(\delta_j, R_j)) &\leq (2^{121} + 2^{-16})g^{-3j/2}, \\ \text{mass}(\{v \in \text{FC}_u(\delta_j, R_j) : v \text{ fails (5.7)}\}) &\leq 2^{95}g^{-3j/2}. \end{aligned}$$

For each shift define on G_Q

$$\mathcal{D}_u^{\text{term}} = \sum_{v \in \text{FC}_u(\delta_j, R_j)} \mathcal{T}_{v,Q} \rho_{k_j} - \rho_{k_j} \circ \text{pr}_{Q,k_j}, \quad \tilde{\mathcal{D}}_u^{\text{term}} = \sum_{\substack{v \in \text{FC}_u(\delta_j, R_j) \\ v \text{ fails (5.7)}}} \mathcal{T}_{v,Q} \rho_{k_j}.$$

By (3.7), $Z_{j,X}^{\text{term}}(t) - Z_j(t) = \sum_u \sum_{w:u_w=u} \text{wt}(w)(\mathcal{D}_u^{\text{term}} - \tilde{\mathcal{D}}_u^{\text{term}})(s_w)$. Lemma 4.8 and (6.2) give $\langle |\mathcal{D}_u^{\text{term}}| \rangle_Q < \frac{28}{1000}(C+1)g^{-3j/2}$; also $\tilde{\mathcal{D}}_u^{\text{term}} \geq 0$ and, by (3.6), its mean is the mass of the words failing the mask, so it obeys the same bound. Apply (5.6) to each group and sum over fewer than $j^2/56$ shifts: the power of j is $5/2 + 2 = 9/2$, and each of the two parts is at most $\frac{45151}{10240}(C+1)j^{9/2}g^{-j}/56 < Fj^{9/2}g^{-j}$.

By Lemma 6.2, $j^{9/2}g^{-j}$ decreases for $j \geq 6000000$. At $j = 10^7$, $2F < 2^{87536}$, $j^{9/2} = 10^{63/2} < 2^{105}$ and $g^{-10^7} < 2^{-10^7/70} < 2^{-142857}$ give $2Fj^{9/2}g^{-j} < 2^{-55216}$. $\square$

6.5. The census. The terminal weighted density weighs each source by ρ_{k_j} , which might be large on a few residues; the census bounds it by the terminal weight, splitting ρ_{k_j} into the bounded subdensity of a lower level and a remainder Φ of small average, on which distinct sources cannot concentrate. As in [8, Proposition 5.3], it is applied through the maps $\text{lift}_i(s) = 4^i s + (4^i - 1)/3$, $i \geq 0$.

Proposition 6.10 (Census with a bounded subdensity). *Let $0 \leq m \leq q$ be integers and let $X \geq 3^q$ be real. Let $\mathcal{S}$ be a set of integers in $[X, 32X)$ with nonnegative weights $v_s \leq t/X$. Suppose $\mathfrak{w}, \Phi \geq 0$ on G_q , $0 \leq \underline{\mathfrak{w}} \leq U$ on G_m , and $\mathfrak{w} \leq \underline{\mathfrak{w}} \circ \text{pr}_{q,m} + \Phi$. Then*

$$\sum_{s \in \mathcal{S}} v_s \sum_{i \geq 0} 4^{-i} \mathfrak{w}(\text{lift}_i(s)) \leq \frac{4}{3} U \sum_{s \in \mathcal{S}} v_s + 44t \langle \Phi \rangle_q. \quad (6.7)$$

Proof. The contribution from $\underline{\mathfrak{w}}$ is at most $U \sum_s v_s \sum_{i \geq 0} 4^{-i} = \frac{4}{3} U \sum_s v_s$. For each fixed i , lift_i is a permutation of G_q because 4^i is a unit. Each residue class contains at most $31X/3^q + 1 \leq 33X/3^q$ integers of $[X, 32X)$, since $X \geq 3^q$. As the elements of $\mathcal{S}$ are distinct and $\Phi \geq 0$,

$$\sum_{s \in \mathcal{S}} v_s \Phi(\text{lift}_i(s)) \leq \frac{t}{X} \frac{33X}{3^q} \sum_{y \in G_q} \Phi(y) = 33t \langle \Phi \rangle_q.$$

Summing against 4^{-i} gives $44t \langle \Phi \rangle_q$. This counting is that of [8, Proposition 5.3]. $\square$

Let $t \geq 16^9$ be odd and not S -periodic, and let X be reachable at j . A word $v \in \text{FC}_X^{\text{term}}(w, \delta_j, R_j)$ with overshoot o gives, after subtracting $2i$ from its last letter, $i = \lfloor o/2 \rfloor$, a word $v' \in \text{FC}_X^{\text{term}}(w, \delta_j, 1)$ with the same length, the same earlier letters and the mask, and with $\text{src}(v, s_w) = \text{lift}_i(\text{src}(v', s_w))$ and $\text{wt}(v) = 4^{-i} \text{wt}(v')$, since adding 2 to the last letter turns a source s' into $4s' + 1$ [8, proof of Proposition 5.3]; and (v', i) determines v . So $Z_{j,X}^{\text{term}}(t)$ is at most the left side of (6.7) for $\mathfrak{w} = \rho_{k_j}$, $\mathcal{S} = \text{Src}_{j,X}(t)$ and the weights $v_s = \text{wt}(ws_w) \leq t/X$ of the sources $s = \text{src}(v', s_w)$, whose total is $\beta_{j,X}^{\text{term}}(t)$.

If L_j is empty, $Z_{j,X}^{\text{term}}(t) = 0$ and (6.8) below is trivial; otherwise reachability, $s_w \geq 16^{\delta_j}$ and $r_{\delta_j} \leq 3\delta_j/10$ give $X > 16^{\delta_j}/(4 \cdot 3^{\delta_j}) > 3^{k_j}$. For an integer m with $n_0 \leq m \leq k_j$, apply Proposition 6.10 with $q = k_j$, $\underline{m} = \underline{\rho}_m$, $U = 8m^\alpha$ and $\Phi = |\rho_{k_j} - \rho_m \circ \text{pr}_{k_j,m}| + (\rho_m - \underline{\rho}_m) \circ \text{pr}_{k_j,m}$, which is nonnegative with $\rho_{k_j} \leq \underline{\rho}_m \circ \text{pr}_{k_j,m} + \Phi$. By (4.5) and (4.6), $\langle \Phi \rangle_{k_j} \leq (C+1)m^{-3/2}$, so

$$Z_{j,X}^{\text{term}}(t) \leq \frac{32}{3} m^\alpha \beta_{j,X}^{\text{term}}(t) + 44(C+1)m^{-3/2}t. \quad (6.8)$$

7. PROOF OF THE MAIN THEOREM

This section combines the key estimates into a count for a target t satisfying the hypotheses of Section 2 (Proposition 7.2), chooses a seed that satisfies them (Proposition 7.4), and deduces Theorem 1.3, using the step bounds of Section 8.

7.1. The count under the hypotheses. The census inequality (6.8) holds at every level m with $n_0 \leq m \leq k_j$; we take m large enough that its error term $44(C+1)m^{-3/2}t$ is negligible.

Lemma 7.1 (The census level is available). *Let $t \geq 1$ and $m = \lceil 2^{80000}t^{2/3} \rceil$. Then $m \geq n_0$, $m^{3/2} \geq 2^{120000}t$, $m < 2^{80001}t^{2/3}$, and $m \leq k_j$ for every integer $j \geq J(t)$.*

Proof. The first three claims are immediate, as $2^{80000}t^{2/3} \geq 1$. Let $j \geq J(t)$. Then $j \geq 200 \log_2 m$, as $200 \log_2 m < 200(80001 + \log_2 t) \leq J(t)$, so $\lfloor j/200 \rfloor \geq \lfloor \log_2 m \rfloor \geq \lceil \log_2 m \rceil - 1$. Bernoulli's inequality gives $g^{200} \geq 3 > 2$, so (5.1) yields $\delta_j \geq 9g^j \geq 9 \cdot 2^{\lfloor j/200 \rfloor} \geq 9 \cdot 2^{\lceil \log_2 m \rceil - 1} \geq \frac{9}{2}m$. Hence $m \leq \delta_j/4$, and $m \leq k_j$ because m is an integer. $\square$

Proposition 7.2 (Count under the hypotheses). *Let $t \geq 16^9$ be odd and not S -periodic, with $Z_{j_0}(t) > 2^{-23}$. Then every real $X \geq 2^{2^{29}t^3}$ is reachable at some generation $j \geq J(t)$, and at every such j*

$$\beta_{j,X}^{\text{term}}(t) > 2^{-160071}t^{-4097/3072}, \quad \# \text{Src}_{j,X}(t) > 2^{-160071}t^{-7169/3072}X. \quad (7.1)$$

Proof. Proposition 6.7 gives j . As $j \geq J(t) > 10^7 > j_0$, Lemma 6.3 gives $Z_j(t) > \frac{3}{4} \cdot 2^{-24}$, and Proposition 6.9 gives $Z_{j,X}^{\text{term}}(t) > \frac{3}{4} \cdot 2^{-24} - 2^{-1500} > 2^{-26}$. Take m as in Lemma 7.1, so (6.8) applies and its last term is at most $44(C+1)2^{-120000} < 2^{-32462}$. Hence

$$\beta_{j,X}^{\text{term}}(t) > \frac{3}{32}(2^{-26} - 2^{-32462})m^{-\alpha} > 2^{-29.42} \cdot 2^{-80001\alpha}t^{-2\alpha/3}.$$

Here $80001\alpha = 327764097/2048 < 160041.07$ and $2\alpha/3 = 4097/3072$, which gives the first bound; the second follows by (5.8). $\square$

7.2. Choosing a seed. A general target need not satisfy the hypotheses of Proposition 7.2, so we apply it to a *seed* t_1 that does, chosen among the integers that reach the target in one odd step. Let t be a positive integer with $3 \nmid t$. As $2^{38} = 4^{19} \equiv 1 \pmod{3}$, exactly one $a^\circ \in \{38, 39\}$ has $2^{a^\circ}t \equiv 1 \pmod{3}$, and the *candidate seeds* are the integers $(2^a t - 1)/3$ with $a \equiv a^\circ \pmod{2}$ and $a \geq a^\circ$. They increase with a , and $3 \cdot (2^a t - 1)/3 + 1 = 2^a t$, so a candidate reaches t after one odd step and a halvings.

Lemma 7.3 (Candidate residues). *Let t' be an integer with $t' \equiv 1 \pmod{3}$, and let $q, i_0 \in \mathbb{N}$. Then the integers $(4^{it'} - 1)/3$, $i_0 \leq i < i_0 + 3^q$, meet every residue class modulo 3^q exactly once.*

Proof. For $t' = 1$ this is [8, Lemma 4.1]. First, $\text{ord}_3(4^e - 1) = 1 + \text{ord}_3(e)$ for every integer $e \geq 1$, by induction on $\text{ord}_3(e)$. If $3 \nmid e$, then $(4^e - 1)/3 = \sum_{i < e} 4^i \equiv e \not\equiv 0 \pmod{3}$. If $e = 3e'$, put $x = 4^{e'} = 1 + 3z$; then $4^e - 1 = (x - 1)(x^2 + x + 1)$ with $x^2 + x + 1 = 3 + 9z + 9z^2$, so $\text{ord}_3(4^e - 1) = \text{ord}_3(4^{e'} - 1) + 1$. Now let $i_0 \leq i < i' < i_0 + 3^q$ and $e = i' - i$, so $\text{ord}_3(e) < q$. The difference of the two integers is $4^{it'}(4^e - 1)/3$, and since $3 \nmid 4^{it'}$ its 3-adic order is $\text{ord}_3(e) < q$. So the 3^q integers are pairwise incongruent modulo 3^q . $\square$

Proposition 7.4 (The seed). *Let t be a positive integer with $3 \nmid t$. There is a candidate seed $t_1 = (2^a t - 1)/3$ with $a < 38 + 4 \cdot 3^{d_0}$ that is odd, not S -periodic, satisfies*

$$2^{36}t < t_1 < 2^{36+4 \cdot 3^{d_0}}t \quad \text{and} \quad Z_{j_0}(t_1) > 2^{-23},$$

and reaches t in $1 + a$ ordinary steps; $S(t_1)$ is the odd part of t .

Proof. By Proposition 6.5 there is a unit $y_0 \in G_{d_0}$ with $Z_{j_0}(y_0) > 2^{-23}$. Apply Lemma 7.3 with $t' = 2^{a^\circ}t$ and $q = d_0$, once with $i_0 = 0$ and once with $i_0 = 3^{d_0}$: this gives two candidate seeds $(4^{it'} - 1)/3$, $i < 2 \cdot 3^{d_0}$, congruent to y_0 modulo 3^{d_0} , with $a = a^\circ + 2i \leq 37 + 4 \cdot 3^{d_0}$. Each is odd, as $2^a t - 1$ is odd, satisfies $Z_{j_0} = Z_{j_0}(y_0)$ by Lemma 6.4, and lies between $(2^{38}t - 1)/3 > 2^{36}t$ and $2^a t/3 < 2^{36+4 \cdot 3^{d_0}}t$. Since $3t_1 + 1 = 2^a t$, $S(t_1)$ is the odd part of t , so the two candidates have the same image under S , and by Lemma 3.2(1) at most one of them is S -periodic. $\square$

7.3. General targets. Every integer that reaches the seed reaches t , so Proposition 7.2 for the seed counts integers reaching t . Let t be a positive integer with $3 \nmid t$, and fix t_1 as in Proposition 7.4. Put the *internal density* and the *internal cutoff*

$$c_{\text{int}} = 2^{-160076} t_1^{-7169/3072}, \quad X_{\text{int}} = 32 \cdot 2^{2^{29}} t_1^3. \quad (7.2)$$

Proposition 7.5. *For every real $X \geq X_{\text{int}}$, at least $c_{\text{int}}X$ odd $n < X$ reach t within $\lfloor 10.46 \log(n/t) \rfloor$ ordinary steps and reach the odd part of t within $\lfloor 3.4881 \log(n/t) \rfloor$ accelerated steps.*

Proof. The seed t_1 is odd, not S -periodic, at least $2^{36} \geq 16^9$, and has $Z_{j_0}(t_1) > 2^{-23}$, so Proposition 7.2 applies to it at $X/32 \geq 2^{2^{29}} t_1^3$: at some $j \geq J(t_1)$ with $X/32$ reachable, $\#\text{Src}_{j,X/32}(t_1) > 2^{-160071} t_1^{-7169/3072} X/32 = c_{\text{int}}X$. Every source in $\text{Src}_{j,X/32}(t_1)$ is odd, lies in $[X/32, X)$, and satisfies both step bounds by Proposition 8.4. $\square$

The theorem's constants are rounded versions of c_{int} and X_{int} :

$$1/c_{\text{int}} < t^4 D_A, \quad X_{\text{int}} < 2^{t^3 D_A}. \quad (7.3)$$

Indeed $\log_2 D_A = 2^{87640} = 4^{d_0}$, and $(4/3)^{d_0} \geq (4/3)^{2^{20}} > 2^{10^5}$ gives $2^{10^5} \cdot 3^{d_0} < 4^{d_0}$. With $\log_2(t_1/t) < 36 + 4 \cdot 3^{d_0}$ (Proposition 7.4) and $t \geq 1$,

$$\log_2(1/c_{\text{int}}) < 160076 + \frac{7169}{3072}(36 + 4 \cdot 3^{d_0}) + 4 \log_2 t < 16 \cdot 3^{d_0} + 4 \log_2 t < 4^{d_0} + \log_2 t^4,$$

and $\log_2 X_{\text{int}} = 5 + 2^{29} t_1^3 < t^3(5 + 2^{29+108+12 \cdot 3^{d_0}}) < t^3 \cdot 2^{4^{d_0}} = t^3 D_A$, since $2^{29} + 109 + 12 \cdot 3^{d_0} < 4^{d_0}$.

Proof of Theorem 1.3. For $X \geq X_0(t) = 2^{t^3 D_A}$, (7.3) gives $X > X_{\text{int}}$ and $c_{\text{int}} > c(t)$, so Proposition 7.5 gives at least $c_{\text{int}}X \geq c(t)X$ odd $n < X$ with both step bounds. $\square$

8. THE NUMBER OF STEPS

This section proves the step bounds used in Proposition 7.5 (Proposition 8.4), following [8, Section 7]: the steps beyond the heuristic $\log(s/t)/\log(4/3)$ of Section 2 are paid by the deviation of the degree from twice the length, which the lines, caps and the mask bound, and by the steps from the seed to the target.

Throughout, t and t_1 are as in Section 7.3. For a generation j define, suppressing j ,

$$\delta_\Sigma = \sum_{i \leq j} \delta_i, \quad \ell_\Sigma^c = \sum_{i < j} \ell_{\delta_i}^c, \quad c_{\text{step}} = \log\left(3 + \frac{1}{4096}\right), \quad c_{\text{src}} = \frac{28559}{100000}.$$

The first three are Mazur's B_* , W_* (with ℓ_δ^c in place of his width) and $h = \log(12289/4096)$; c_{step} bounds each step $\log(3 + 1/y)$ of an inverse trajectory with values $y \geq 4096$, and c_{src} is the rate in Lemma 8.3. We use $(2 - \log_2 3) \log 2 = \log(4/3)$ and

$$\frac{287682}{10^6} < \log \frac{4}{3}, \quad \log 2 < \frac{347}{500} < \frac{7}{10}, \quad c_{\text{step}} - \log 3 < \frac{1}{12288}; \quad (8.1)$$

the first two are Lemma A.1(3), and the last is $\log(1 + 1/12288) < 1/12288$.

Lemma 8.1 (Scale sums). *Let $j \geq 2 \cdot 10^{10}$. Then (1) $\delta_j \geq 2^{40}j$, and hence $2^{22}(j+1) \leq \delta_\Sigma$; (2) $\delta_j \leq 9 + \delta_\Sigma/100 + 2j$; and (3) $9000 \ell_\Sigma^c \leq \delta_\Sigma$.*

Proof. These are this paper's forms of Mazur's bounds $n+1 \leq B_*/32^5$, $100b_n \leq 100b_0 + B_*$ and [8, (7.1)]. By Lemma 15.1, $\delta_i \geq 256$, hence $e_{\delta_i} = \lceil \delta_i/100 \rceil$, for $i \geq 42$, $2^{40} < \delta_{2254} = 1100658881854 < 2^{41}$, and $\delta_j \geq 2^{40}j$; then $2^{22}(j+1) \leq 2^{23}j \leq \delta_j \leq \delta_\Sigma$, proving (1). For (2), $e_{\delta_i} < \delta_i/100 + 1$ for $i \geq 42$, so

$$\delta_j = 256 + \sum_{42 \leq i < j} e_{\delta_i} \leq 256 + \frac{\delta_\Sigma}{100} + (j-42) \leq 9 + \frac{\delta_\Sigma}{100} + 2j.$$

For (3), $\ell_\delta^c \leq \ell_\delta \leq \delta$ for every $\delta \geq 9$. Let $\delta \geq 2^{40}$. Then $40 \leq \lfloor \log_2 \delta \rfloor$, $\lceil \log_2 \delta \rceil \leq \lfloor \log_2 \delta \rfloor + 1$ and $2^{\lfloor \log_2 \delta \rfloor} \leq \delta$, and $(y+1)/2^y$ is nonincreasing in $y \geq 1$, so

$$\frac{\ell_\delta^c}{\delta} < \sqrt{\frac{32 \lfloor \log_2 \delta \rfloor}{\delta}} + \frac{1}{\delta} \leq \sqrt{\frac{32(\lfloor \log_2 \delta \rfloor + 1)}{2^{\lfloor \log_2 \delta \rfloor}}} + \frac{1}{2^{40}} \leq \sqrt{\frac{1312}{2^{40}}} + \frac{1}{2^{40}} \leq \frac{1}{20000},$$

the last being the rational comparison $1312/2^{40} \leq (1/20000 - 2^{-40})^2$. Split ℓ_Σ^c at $i = 2254$. The terms with $2254 \leq i < j$ contribute at most $\delta_\Sigma/20000$. For $i < 2254$, (5.1) gives $\delta_i \leq g^{i-2254} \delta_{2254}$, so these terms total at most $100 \delta_{2254} < 100 \cdot 2^{41}$. By (1), $\delta_\Sigma \geq \delta_j \geq 2^{41} \cdot 10^{10}$, hence $9000 \ell_\Sigma^c \leq 9 \cdot 10^5 \cdot 2^{41} + \frac{9}{20} \delta_\Sigma \leq \delta_\Sigma$. $\square$

Let $j \geq J(t_1)$, let X be reachable at j for the target t_1 , and let $s \in \text{Src}_{j,X}(t_1)$ be the source of wv , with $w = v_0 \cdots v_{j-1} \in L_j(t_1)$ and $v \in \text{FC}_X^{\text{term}}(w, \delta_j, 1)$; we call s a *counted source*. The word wv is valid at t_1 . Its inverse trajectory from t_1 has one value after each number $k = 0, 1, \dots, |wv|$ of letters, the first being t_1 and the last s ; for $i \leq j$ the value after $|w_{<i}|$ letters is $s_{w_{<i}} \geq 16^{\delta_i}$ (Lemma 5.4).

Lemma 8.2 (Inverse trajectory of a counted source). *In this notation, (1) every value of the inverse trajectory of wv is at least 4096; (2) $\deg(wv) \log 2 \leq \log s - \log t_1 + |wv| c_{\text{step}}$; and (3) $\deg(wv) \geq 2|wv| - (2 - \log_2 3) \ell_\Sigma^c - j - \delta_j/8$.*

Proof. For (1), we follow [8, Section 7]. The first value is $t_1 > 2^{36}$. Let y be the value after k letters, where either $|w_{<i}| < k \leq |w_{<i+1}|$ with $i < j$, or $|w| < k \leq |wv|$ and $i = j$. Then $S^{o(k-|w_{<i}|)}(y) = s_{w_{<i}}$, and S at most doubles every positive odd integer, so $s_{w_{<i}} \leq 2^{k-|w_{<i}|} y$, where $k - |w_{<i}| \leq d_{\delta_i}^{\text{hi}} \leq 8\delta_i/5$; hence $y \geq 16^{\delta_i} 2^{-8\delta_i/5} = 2^{12\delta_i/5} \geq 2^{108/5} > 4096$.

For (2), write $wv = (a_1, \dots, a_{|wv|})$. If y and y' are the values after k and $k-1$ letters, then $3y+1 = 2^{a_k}y'$, so $a_k \log 2 = \log y - \log y' + \log(3+1/y)$. Summing over k , the first two terms telescope to $\log s - \log t_1$, and by (1) each last term is at most c_{step} ; this is [8, (7.4)].

For (3), which is [8, (7.3)] with ℓ_δ^c in place of his width, let $i < j$. By (5.2) with $u = r_{\delta_i}$, $\deg(v_i) \geq H_{\delta_i, r_{\delta_i}}(|v_i|) > 2|v_i| - (2 - \log_2 3)(|v_i| - \delta_i) - 1$, and $|v_i| - \delta_i \leq \ell_{\delta_i}^c$, by the length restriction for $\delta_i \geq 256$ and by $|v_i| \leq d_{\delta_i}^{\text{hi}}$ otherwise. As $2 - \log_2 3 > 0$, $\deg(v_i) \geq 2|v_i| - (2 - \log_2 3)\ell_{\delta_i}^c - 1$. For v , the first crossing and the mask (5.7) give $\deg(v) \geq H_{\delta_j, u_w}(|v|) \geq 2|v| - \delta_j/8$. Summing proves (3). $\square$

Lemma 8.3 (Size of a counted source). *Every $s \in \text{Src}_{j,X}(t_1)$ satisfies $\log(s/t) \geq c_{\text{src}}\delta_\Sigma$.*

Proof. This is [8, (7.2)] with this paper's scales. Apply (5.3) to v_i at $s_{w_{<i}} \geq 16^{\delta_i}$ with $u = r_{\delta_i}$ for $i < j$, and to v at s_w with $u = u_w \geq 0$: each word of scale δ multiplies the value by more than $2^{u-r_{\delta}-2}(4/3)^\delta$. Hence $\log(s/t) \geq \log(t_1/t) + \delta_\Sigma \log \frac{4}{3} - (r_{\delta_j} + 2j + 2) \log 2$. By Lemmas 5.1 and 8.1(2), $r_{\delta_j} \leq \frac{3}{10}\delta_j \leq \frac{27}{10} + \frac{3}{1000}\delta_\Sigma + \frac{3}{5}j$, and $2j + 2 + \frac{3}{5}j \leq \frac{26}{10}(j+1)$. Since $\log(t_1/t) > 36 \log 2 > \frac{27}{10} \log 2$ and $j+1 \leq \delta_\Sigma/2^{22}$ by Lemma 8.1(1),

$$\log(s/t) \geq \delta_\Sigma \left(\log \frac{4}{3} - \log 2 \left(\frac{3}{1000} + \frac{26}{10 \cdot 2^{22}} \right) \right),$$

and by (8.1) the bracket exceeds $\frac{287682}{10^6} - \frac{347}{500} \left(\frac{3}{1000} + \frac{26}{10 \cdot 2^{22}} \right) = \frac{598945709}{2097152000} = c_{\text{src}} + \frac{501733}{52428800000}$. $\square$

Proposition 8.4 (Step bounds). *Let t_{odd} be the odd part of t . Each $s \in \text{Src}_{j,X}(t_1)$ reaches t within $\left\lfloor \frac{523}{50} \log(s/t) \right\rfloor$ ordinary steps and t_{odd} within $\left\lfloor \frac{34881}{10000} \log(s/t) \right\rfloor$ accelerated steps.*

Proof. As $j \geq J(t_1) \geq 2 \cdot 10^{10}$, Lemmas 8.1 and 8.3 apply. For the accelerated steps we follow the proof of [8, Lemma 7.2]. Put $c_{\text{gain}} = 2 \log 2 - c_{\text{step}}$; by (8.1), $\log(4/3) - 1/12288 < c_{\text{gain}} < \log(4/3)$. Lemma 8.2(2) and (3), multiplied out with $(2 - \log_2 3) \log 2 = \log(4/3)$, give

$$c_{\text{gain}}(|wv| + 1) \leq \log s - \log t_1 + c_{\text{gain}} + \log \frac{4}{3} \ell_\Sigma^c + \left(j + \frac{\delta_j}{8} \right) \log 2.$$

By Lemma 8.1(3), $\log \frac{4}{3} \ell_\Sigma^c \leq \log \frac{4}{3} \delta_\Sigma/9000$. By Lemma 8.1(2), $j + \delta_j/8 \leq \frac{9}{8} + \frac{\delta_\Sigma}{800} + \frac{5}{4}j$, and by $\log 2 < 7/10$ and Lemma 8.1(1), $\frac{5}{4}j \log 2 < \delta_\Sigma/10^6$ and $\frac{\delta_\Sigma}{800} \log 2 < \frac{7\delta_\Sigma}{8000}$. Finally $c_{\text{gain}} + \frac{9}{8} \log 2 < 36 \log 2 < \log(t_1/t)$. As $1/10^6 + 7/8000 = 219/250000$, we obtain, as in [8, (7.6)],

$$c_{\text{gain}}(|wv| + 1) \leq \log(s/t) + \left(\frac{\log(4/3)}{9000} + \frac{219}{250000} \right) \delta_\Sigma. \quad (8.2)$$

Put $c_{\text{acc}} = 34881/10000$, the constant of [8, (7.7)]. Since $0 \leq \delta_\Sigma \leq \log(s/t)/c_{\text{src}}$, it suffices for $|wv| + 1 \leq c_{\text{acc}} \log(s/t)$ that $c_{\text{src}} + \frac{\log(4/3)}{9000} + \frac{219}{250000} \leq c_{\text{src}} c_{\text{acc}} c_{\text{gain}}$. Using $c_{\text{gain}} > \log(4/3) - 1/12288$ and $\log(4/3) > 287682/10^6$, the right side minus the left side exceeds the rational residual

$$\left(c_{\text{src}} c_{\text{acc}} - \frac{1}{9000} \right) \frac{287682}{10^6} - \frac{c_{\text{src}} c_{\text{acc}}}{12288} - c_{\text{src}} - \frac{219}{250000} = \frac{25365007801}{192000000000000000} > 0. \quad (8.3)$$

Hence $|wv| + 1 \leq c_{\text{acc}} \log(s/t)$. Since s reaches t_1 in $|wv|$ accelerated steps and $S(t_1) = t_{\text{odd}}$, s reaches t_{odd} within $|wv| + 1$ accelerated steps.

For the ordinary steps we follow the proof of [8, Lemma 7.3]. By Section 3, s reaches t_1 in $|wv| + \deg(wv)$ ordinary steps, and t_1 reaches t in $1 + a$ more, where $3t_1 + 1 = 2^a t$; so s reaches t within $N_{\text{ord}} = |wv| + 1 + \deg(wv) + a$ ordinary steps. As $t_1 > 4096$, $a \log 2 = \log(3t_1 + 1) - \log t \leq \log t_1 + c_{\text{step}} - \log t$, and adding Lemma 8.2(2) gives $(\deg(wv) + a) \log 2 \leq \log(s/t) + (|wv| + 1)c_{\text{step}}$. With $|wv| + 1 \leq c_{\text{acc}} \log(s/t)$,

$$N_{\text{ord}} \log 2 \leq (|wv| + 1)(\log 2 + c_{\text{step}}) + \log(s/t) \leq (c_{\text{acc}}(\log 2 + c_{\text{step}}) + 1) \log(s/t).$$

It remains to prove Mazur's inequality [8, (7.9)],

$$c_{\text{acc}}(\log 2 + c_{\text{step}}) + 1 \leq \frac{523}{50} \log 2. \quad (8.4)$$

By (8.1) and Lemma A.1(1), $c_{\text{step}} < (317/200) \log 2 + 1/12288$, so it suffices that

$$\left(\frac{523}{50} - \frac{517}{200} c_{\text{acc}} \right) \log 2 \geq \frac{c_{\text{acc}}}{12288} + 1.$$

The bracket is $2886523/2000000 > 0$, and by $\log 2 > 693147180/10^9$ (Lemma A.1(3)) the difference of the two sides exceeds

$$\frac{693147180}{10^9} \cdot \frac{2886523}{2000000} - \frac{c_{\text{acc}}}{12288} - 1 = \frac{43510569153}{40000000000000} > 0.$$

Both step counts are integers, so both bounds may be floored. $\square$

9. THE FOURIER BOUND: BLACK POINTS AND TRIANGLES

This section and the next five prove the Fourier bound (4.2), following Tao's argument [10, Section 7] with explicit constants. Section 11 bounds $|\hat{\mu}_n(\xi)|$ by the expected penalty of a random walk, which gains a factor below 1 at the white points of Definition 9.2, and reduces its decay to a descent step proved in three cases in Sections 12–14; Section 10 supplies the laws of the walk. Here we describe the geometry of the points at which the character sum gains no cancellation: these *black* points form disjoint triangles, and every point bordering a triangle is *white*. In Sections 9–14 the letter j denotes the first coordinate of a point and J is as in Definition 9.1, and from Section 10 on b (also b^k) denotes a hold word and $\mathbf{b}$ a list of hold words; the generations j , scales δ_j , families L_j and the bound $J(t)$ of Sections 5–7 do not occur there.

Definition 9.1 (Standing conventions). From now on, $n \geq 1$, $J = \lfloor n/2 \rfloor$, $\xi \in G_n$ is a frequency prime to 3, and $\varepsilon = 8/217$ is the constant of Appendix A.1. The objects defined below depend on n and ξ ; this dependence is suppressed from the notation.

Apart from Lemma 9.14, the arguments of this section use only $0 < \varepsilon < 1/27$.

9.1. Points, angles, white and black points. In the character sum, the cancellation contributed by the j th pair (a_{2j-1}, a_{2j}) of letters of the random word W is governed by the angle, defined next, of the point $(j, \deg_{2j}(W))$ (see Lemma 11.1).

Definition 9.2 (Points and angles). A *point* is an element $p = (j, l)$ of $\mathcal{P} = \mathbb{Z}_{\geq 1} \times \mathbb{Z}$; we write $j(p) = j$ and $l(p) = l$, the *raw length* and *height* of p , and add points coordinatewise. Points $p \neq v$ are *cardinal neighbours* if $|j(p) - j(v)| + |l(p) - l(v)| = 1$. For $y \in G_n$, let $\tilde{y} \in \mathbb{Z}$ be the representative of y with $|\tilde{y}| < 3^n/2$ (it is unique, as 3^n is odd) and put $\text{sfr}(y) = \tilde{y}/3^n$, the *signed fractional part*. The *angle* of $p = (j, l)$ is

$$\vartheta(p) = \text{sfr}(\text{res}(p)), \quad \text{res}(p) = [3^{2(j-1)} 2^{1-l}]_n \xi.$$

A point p is *white* if $j(p) \leq J$ and $|\vartheta(p)| > \varepsilon$, and *black* if $j(p) \leq J$ and $|\vartheta(p)| \leq \varepsilon$. Points with $j(p) > J$ are neither.

Since the numerator of $\text{sfr}(y)$ is the representative of least modulus,

$$y' \in \mathbb{Z}, y' \equiv y, |y'| < 3^n/2 \implies \text{sfr}(y) = \frac{y'}{3^n}; \quad |\text{sfr}(ky)| \leq k |\text{sfr}(y)| \quad (k \in \mathbb{N}), \quad (9.1)$$

the second because $k \cdot 3^n \text{sfr}(y)$ represents ky .

Moving a point one step right multiplies its residue by 9, and one step down by 2; the southeast distance measures the resulting growth of the angle (Lemma 9.4).

Definition 9.3 (Southeast distance and the constant λ). For points u, v with $j(u) \leq j(v)$ and $l(v) \leq l(u)$ put the *southeast distance*

$$\text{se}(u, v) = (j(v) - j(u)) \log 9 + (l(u) - l(v)) \log 2,$$

so that $e^{\text{se}(u, v)} = 9^{j(v)-j(u)} 2^{l(u)-l(v)}$ is a positive integer. Let $\lambda = \log 2 / \log 9$, the slope of the hypotenuse of a triangle (Definition 9.9).

Lemma 9.4 (Southeast transport). *Let $u, v \in \mathcal{P}$ with $j(u) \leq j(v)$ and $l(v) \leq l(u)$, and put $k = e^{\text{se}(u, v)}$. Then $\text{res}(v) = k \text{res}(u)$ and $|\vartheta(v)| \leq k |\vartheta(u)|$. If moreover $k |\vartheta(u)| < 1/2$, then $\vartheta(v) = k \vartheta(u)$.*

Proof. As in [10, (7.12)], $3^{2(j(v)-1)} 2^{1-l(v)} = k \cdot 3^{2(j(u)-1)} 2^{1-l(u)}$ in $\mathbb{Z}[1/2]$; apply $[\cdot]_n$ and multiply by ξ . The bound follows from the second part of (9.1), and the equality, as in [10, (7.18)], from its first part with $y' = k \cdot 3^n \vartheta(u)$. $\square$

Lemma 9.5 (Angles are bounded below). *If $j(p) \leq J$, then $3^{n+1-2j(p)} |\vartheta(p)| \geq 1/3$; in particular $\vartheta(p) \neq 0$.*

Proof. This is the proof of [10, (7.16)] with ε replaced by $|\vartheta(p)|$: for $p = (j, l)$ and $k = n+1-2j \geq 1$, $3^k \text{res}(p) = 3^{n-1} [2^{1-l}]_n \xi$ is represented by $\pm 3^{n-1}$, so $1/3 = |\text{sfr}(3^k \text{res}(p))| \leq 3^k |\vartheta(p)|$ by (9.1). $\square$

9.2. Angles of neighbouring points. The proof of [10, Lemma 7.4] assumes $0 < \varepsilon < 1/100$; our $\varepsilon = 8/217$ is larger, which is why we reprove its ingredients with explicit thresholds. The next lemma is a quantitative form of the claims (i)–(iii) of that proof, which concern points with $|\vartheta| \leq 1/100$: it recovers the angle of a point exactly from the angles of two cardinal neighbours.

Lemma 9.6 (Quantitative claims (i)–(iii)). *Let $p = (j, l) \in \mathcal{P}$ and let $\bar{\vartheta}_1, \bar{\vartheta}_2$ be real numbers with $|\vartheta(j, l-1)| \leq \bar{\vartheta}_2$. Each of the following implies $2\vartheta(p) = \vartheta(j, l-1)$:*

- (i) $|\vartheta(j, l)| \leq \bar{\vartheta}_1$ and $2\bar{\vartheta}_1 + \bar{\vartheta}_2 < 1$;
- (ii) $|\vartheta(j+1, l)| \leq \bar{\vartheta}_1$ and $2\bar{\vartheta}_1 + 9\bar{\vartheta}_2 < 1$;
- (iii) $j \geq 2$, $|\vartheta(j-1, l)| \leq \bar{\vartheta}_1$ and $18\bar{\vartheta}_1 + \bar{\vartheta}_2 < 1$.

Proof. Write $y_2 = 3^n \vartheta(j, l-1)$; by Lemma 9.4 it represents $2 \text{res}(p)$. In each case $\bar{\vartheta}_1 \geq 0$, so $\bar{\vartheta}_2 < 1$.

(i) The integer $2 \cdot 3^n \vartheta(p) - y_2$ is divisible by 3^n and has modulus at most $(2\bar{\vartheta}_1 + \bar{\vartheta}_2) 3^n < 3^n$, so it vanishes.

(ii) Let $y_1 = 3^n \vartheta(j+1, l)$, which represents $9 \text{res}(p)$. Then $2y_1 - 9y_2 \equiv 18 \text{res}(p) - 18 \text{res}(p) = 0$ and $|2y_1 - 9y_2| < 3^n$, so $2y_1 = 9y_2$. The integer $y_1 - 4y_2$ represents $9 \text{res}(p) - 8 \text{res}(p) = \text{res}(p)$ and equals $y_2/2$, of modulus at most $\bar{\vartheta}_2 3^n/2 < 3^n/2$. By (9.1), $\vartheta(p) = y_2/(2 \cdot 3^n)$.

(iii) Let $y_3 = 3^n \vartheta(j-1, l)$, which represents $\text{res}(j-1, l)$; then $\text{res}(p) = 9 \text{res}(j-1, l)$ and y_2 represents $18 \text{res}(j-1, l)$. As before $|18y_3 - y_2| < 3^n$ forces $y_2 = 18y_3$, and $9y_3 = y_2/2$ represents $\text{res}(p)$ with modulus less than $3^n/2$, so $\vartheta(p) = y_2/(2 \cdot 3^n)$. $\square$

Corollary 9.7 (Small angles along a row or column). *Let $(j, l) \in \mathcal{P}$ and $m \in \mathbb{N}$.*

- (a) *If $|\vartheta(j+m, l)| \leq \varepsilon$ and $|\vartheta(j+k, l-1)| \leq \varepsilon$ for $0 \leq k < m$, then $|\vartheta(j+k, l)| \leq \varepsilon$ for $0 \leq k \leq m$.*
- (b) *If $|\vartheta(j, l)| \leq \varepsilon$ and $|\vartheta(j+k, l-1)| \leq \varepsilon$ for $1 \leq k \leq m$, then $|\vartheta(j+k, l)| \leq \varepsilon$ for $0 \leq k \leq m$.*
- (c) *If $|\vartheta(j, l)| \leq \varepsilon$ and $|\vartheta(j+1, l+k)| \leq \varepsilon$ for $1 \leq k \leq m$, then $|\vartheta(j, l+k)| \leq \varepsilon$ for $0 \leq k \leq m$.*

Proof. Take $\bar{\vartheta}_1 = \bar{\vartheta}_2 = \varepsilon$, so that $2\bar{\vartheta}_1 + 9\bar{\vartheta}_2 = 11\varepsilon < 1$ and $18\bar{\vartheta}_1 + \bar{\vartheta}_2 = 19\varepsilon < 1$. In (a), induct downwards on k using Lemma 9.6(ii) at $(j+k, l)$: $|\vartheta(j+k, l)| = |\vartheta(j+k, l-1)|/2 \leq \varepsilon$. In (b), induct upwards on k using Lemma 9.6(iii) at $(j+k, l)$, where $j+k \geq 2$. In (c), induct on k using Lemma 9.6(ii) at $(j, l+k)$, whose right neighbour is $(j+1, l+k)$ and lower neighbour $(j, l+k-1)$. $\square$

9.3. Column tops and canonical triangles. To make the decomposition of the black points canonical, each black point is assigned the triangle whose corner is reached by moving up through black points of its column as far as possible and then left through black points of that row, as in the proof of [10, Lemma 7.4]. Moving up stops: if $p = (j, l)$ is black, then $(j, l + k)$ is white for some $1 \leq k \leq 2n$. Otherwise $(j, l + k)$ is black for $0 \leq k \leq 2n$, since $j \leq J$. Moving down one step doubles the angle exactly, by Lemma 9.4 and $2\varepsilon < 1/2$; hence $|\vartheta(p)| = 4^n |\vartheta(j, l + 2n)|$. By Lemma 9.5 and $3^{n+1-2j} \leq 3^{n-1} < 4^n$, this is at least $1/3 > \varepsilon$, contradicting blackness of p . As every point v with $j(v) \leq J$ is either black or white, the following objects are well defined.

Definition 9.8 (Column top, gap, row start). Let $p = (j, l)$ be black. Its *column top* $l_*(p)$ is the unique $l'' \geq l$ such that (j, l'') is black for $l \leq l'' \leq l''$ and $(j, l'' + 1)$ is white. Its *gap* is $\text{gap}(p) = l_*(p) - l(p) \in \mathbb{N}$. Its *row start* $j_*(p)$ is the unique $j' \in [1, j]$ such that $(i, l_*(p))$ is black for $j' \leq i \leq j$ and either $j' = 1$ or $(j' - 1, l_*(p))$ is white.

Definition 9.9 (Triangles and the canonical family). A *triangle* is a triple $\Delta = (j_\Delta, l_\Delta, S_\Delta) \in \mathbb{Z}_{\geq 1} \times \mathbb{Z} \times \mathbb{R}$, with *corner* $p_\Delta = (j_\Delta, l_\Delta)$ and *size* S_Δ . A point v *lies in* Δ , written $v \in \Delta$, if

$$j(v) \geq j_\Delta, \quad l(v) \leq l_\Delta, \quad \text{se}(p_\Delta, v) \leq S_\Delta.$$

The *top edge* of Δ is the row $l = l_\Delta$, and its *left edge* is the column $j = j_\Delta$. For a black point p , the *canonical triangle* is

$$\Delta(p) = \left(j_*(p), l_*(p), \log \frac{\varepsilon}{|\vartheta(j_*(p), l_*(p))|} \right),$$

and the *canonical family* is $\mathfrak{T} = \{\Delta(p) : p \text{ black}\}$.

Thus $v \in \Delta$ means that v lies weakly southeast of the corner and weakly left of the hypotenuse $j - j_\Delta = \lambda(l - l_\Delta) + S_\Delta / \log 9$. Membership is *closed to the northwest*: if $v \in \Delta$ and $j_\Delta \leq j(v') \leq j(v)$, $l(v) \leq l(v') \leq l_\Delta$, then $v' \in \Delta$, since $\text{se}(p_\Delta, \cdot)$ decreases. Also every $v \in \Delta$ satisfies $j(v) \leq j_\Delta + S_\Delta / \log 9$.

Lemma 9.10 (Canonical triangles). *Let p be black and $\Delta = \Delta(p)$. Then*

- (a) p_Δ is black, $|\vartheta(p_\Delta)| = \varepsilon e^{-S_\Delta}$ and $S_\Delta \geq 0$;
- (b) (the right-end bound)

$$j_\Delta + S_\Delta / \log 9 \leq n/2 + 1 - \log(1/\varepsilon) / \log 9 < n/2 - 1/2 \leq J; \quad (9.2)$$

- (c) every $v \in \Delta$ is black.

Proof. (a) The corner is black by Definition 9.8, so $j_\Delta \leq J$, $\vartheta(p_\Delta) \neq 0$ by Lemma 9.5, and the size is defined by the stated formula; $S_\Delta \geq 0$ as $|\vartheta(p_\Delta)| \leq \varepsilon$.

(b) Lemma 9.5 gives $\log(1/|\vartheta(p_\Delta)|) \leq (n + 2 - 2j_\Delta) \log 3 = (n/2 + 1 - j_\Delta) \log 9$. Subtract $\log(1/\varepsilon)$ to obtain the first inequality. The second follows from $\log(1/\varepsilon) > \log 27 = \frac{3}{2} \log 9$, and the third from $J = \lfloor n/2 \rfloor$.

(c) By Lemma 9.4 and (a), $|\vartheta(v)| \leq e^{\text{se}(p_\Delta, v)} \varepsilon e^{-S_\Delta} \leq \varepsilon$. By (b), $j(v) \leq j_\Delta + S_\Delta / \log 9 < n/2 - 1/2$, so $j(v) \leq J$. $\square$

9.4. The neighbourhood of a triangle and the partition of the black points. The *cardinal neighbourhood* of a triangle Δ consists of the points not in Δ that are cardinal neighbours of a point of Δ . Tao's claim (*) in the proof of [10, Lemma 7.4] makes every point within distance $\frac{1}{10} \log \frac{1}{\varepsilon}$ of a triangle white, for small ε . The next lemma is its cardinal-neighbour case, proved at $\varepsilon = 8/217$ by Tao's three cases.

Lemma 9.11 (White neighbourhood). *Let p be black. Every point of the cardinal neighbourhood of $\Delta(p)$ is white.*

Proof. We follow the proof of claim (*) in [10, Lemma 7.4], with distance 1 in place of $\frac{1}{10} \log \frac{1}{\varepsilon}$; since v below is itself black, Corollary 9.7 replaces Tao's claims (i)–(iii) and weakly black points. Write $\Delta = \Delta(p)$, and let $v \notin \Delta$ be a cardinal neighbour of $x \in \Delta$. As $j(x) \leq j_\Delta + S_\Delta / \log 9 < J$ by Lemma 9.10(b), $j(v) \leq J$, so it suffices to show that v is not black. Recall $j_\Delta \leq j(p)$, that (i, l_Δ) is black for $j_\Delta \leq i \leq j(p)$, and that $p' = (j(p), l_\Delta + 1)$ is white. Tao's three cases are exhaustive.

Case 1: $j(v) \geq j_\Delta$, $l(v) \leq l_\Delta$. Then $S_\Delta < \text{se}(p_\Delta, v) \leq S_\Delta + \log 9$, as a cardinal step changes $\text{se}(p_\Delta, \cdot)$ by at most $\log 9$. Since $9\varepsilon < 1/2$, Lemma 9.4 and Lemma 9.10(a) give $|\vartheta(v)| = \varepsilon e^{\text{se}(p_\Delta, v) - S_\Delta} > \varepsilon$.

Case 2: $j(v) \geq j_\Delta$, $l(v) > l_\Delta$. Then $v = x + (0, 1)$ and $l(x) = l_\Delta$; suppose v is black. If $j(v) < j(p)$, Corollary 9.7(b), started at v along the black points $(j(v) + k, l_\Delta)$, $1 \leq k \leq j(p) - j(v)$, gives $|\vartheta(p')| \leq \varepsilon$. If $j(v) > j(p)$, the points $(j(p) + k, l_\Delta)$, $0 \leq k < j(v) - j(p)$, lie in Δ by northwest closure from x and are black by Lemma 9.10(c), and Corollary 9.7(a), ending at v , gives the same. If $j(v) = j(p)$, then $p' = v$. So p' is black, as $j(p') \leq J$; but it is white.

Case 3: $j(v) < j_\Delta$. Then $v = x - (1, 0)$, $j_\Delta \geq 2$ and $l(v) \leq l_\Delta$, so $v' = (j_\Delta - 1, l_\Delta)$ is white; suppose v is black. The points $(j_\Delta, l(v) + k)$, $1 \leq k \leq l_\Delta - l(v)$, lie between x and p_Δ , hence in Δ , and are black; Corollary 9.7(c), started at v , gives $|\vartheta(v')| \leq \varepsilon$, so v' is black, a contradiction. $\square$

Lemma 9.12 (Membership and coherence). *Let p be black. Then $p \in \Delta(p)$. If $v \in \Delta(p)$ is black, then $l_*(v) = l_*(p)$ and $\Delta(v) = \Delta(p)$.*

Proof. As in the last paragraph of the proof of [10, Lemma 7.4], with Lemma 9.11 in place of claim (*). Write $\Delta = \Delta(p)$. By Lemma 9.11, a black cardinal neighbour of a point of Δ lies in Δ . The corner lies in Δ because $S_\Delta \geq 0$. Walking right through the black points (i, l_Δ) , $j_\Delta \leq i \leq j(p)$, and then down through the black points $(j(p), l')$, $l(p) \leq l' \leq l_\Delta$, every point lies in Δ ; the last one is p .

Let $v \in \Delta$ be black. By northwest closure and Lemma 9.10(c), the points $(j(v), l')$ with $l(v) \leq l' \leq l_\Delta$ are black. The point $(j(v), l_\Delta + 1)$ is not in Δ and is a cardinal neighbour of $(j(v), l_\Delta) \in \Delta$, so it is white by Lemma 9.11. Hence $l_*(v) = l_\Delta$ by Definition 9.8. Similarly the points (i, l_Δ) , $j_\Delta \leq i \leq j(v)$, lie in Δ and are black, and either $j_\Delta = 1$ or $(j_\Delta - 1, l_\Delta)$ is white, so $j_*(v) = j_\Delta$. The corners agree, and the size depends only on the corner. $\square$

Proposition 9.13 (Partition of the black points). *A point is black if and only if it lies in some $\Delta \in \mathfrak{T}$. Two members of $\mathfrak{T}$ with a common point are equal.*

Proof. A black p lies in $\Delta(p)$, and points of canonical triangles are black, by Lemmas 9.12 and 9.10(c). If $x \in \Delta(p) \cap \Delta(p')$, then x is black and Lemma 9.12 gives $\Delta(p) = \Delta(x) = \Delta(p')$. $\square$

9.5. White exit points above a triangle. In Sections 12 and 14 a random walk leaves a triangle $\Delta \in \mathfrak{T}$ from a point $p \in \Delta$ at height $l_\Delta - l(p)$ below the top edge and first lands above that edge at an exit point $(j(p) + r, l_\Delta + o)$ with $1 \leq o \leq 4$. Such an exit point is white when r is not too large.

Lemma 9.14 (White exit points). *For $\omega \in \mathbb{N}$ put $f_\omega = \lfloor (5\omega + 16)/16 \rfloor$. Let $\Delta \in \mathfrak{T}$, $p \in \Delta$ and $\omega = l_\Delta - l(p)$, which is $\text{gap}(p)$ by Lemmas 9.10(c) and 9.12. Then $(j(p) + r, l_\Delta + o)$ is white for all integers $1 \leq r \leq f_\omega$ and $1 \leq o \leq 4$.*

Proof. Since $f_\omega \leq 1 + 5\omega/16$ and $9^5 \leq 2^{16}$, $9^{f_\omega} \leq 9 \cdot 2^\omega$. Hence $3\varepsilon 9^r \leq 3\varepsilon 9^{f_\omega} \leq 27\varepsilon 2^\omega = \frac{216}{217} 2^\omega < 2^\omega$. Put $p' = (j(p) + r, l_\Delta + o)$. Membership of p gives $j(p) \leq j_\Delta + S_\Delta / \log 9 - \lambda\omega$, so Lemma 9.10(b) gives $j(p) \leq n/2 + 1 - \log(1/\varepsilon) / \log 9 - \lambda\omega$. Taking logarithms in $9^r < 2^\omega / (3\varepsilon)$ gives $r < \lambda\omega + \log(1/\varepsilon) / \log 9 - 1/2$. Adding, $2j(p') < n + 1$, so $j(p') \leq J$.

Suppose p' is not white; then it is black. Put $p'' = (j(p'), l_\Delta + 1)$. Lemma 9.4 gives $|\vartheta(p'')| \leq 2^{o-1}\varepsilon \leq 8\varepsilon$. By northwest closure, $(j(p), l_\Delta) \in \Delta$. We use that points of Δ are black and points of its cardinal neighbourhood are white (Lemmas 9.10(c) and 9.11).

Case 1. Suppose $(j(p'), l_\Delta) \in \Delta$. Then $|\vartheta(j(p'), l_\Delta)| \leq \varepsilon$, and Lemma 9.6(i) at p'' with $\bar{\vartheta}_1 = 8\varepsilon$, $\bar{\vartheta}_2 = \varepsilon$ ($17\varepsilon < 1$) gives $|\vartheta(p'')| \leq \varepsilon/2$. So p'' is black; but $p'' \notin \Delta$ is a cardinal neighbour of $(j(p'), l_\Delta)$, hence white.

Case 2. Suppose $(j(p'), l_\Delta) \notin \Delta$. Let i be the largest integer in $[j(p), j(p')]$ with $(i, l_\Delta) \in \Delta$; then $i < j(p')$, and put $m = j(p') - i \geq 1$. For $0 < k < m$ the point $x = (i + k, l_\Delta)$ satisfies, by Lemma 9.4, Lemma 9.10(a) and $S_\Delta \geq (j(p) - j_\Delta) \log 9 + \omega \log 2$,

$$|\vartheta(x)| \leq 9^{j(x)-j_\Delta} \varepsilon e^{-S_\Delta} \leq \varepsilon 9^{j(x)-j(p)} 2^{-\omega} \leq \varepsilon 9^{r-1} 2^{-\omega} < \frac{1}{27}.$$

Also $|\vartheta(i, l_\Delta)| \leq \varepsilon < 1/27$. Put $\bar{\vartheta}_1 = \max(8\varepsilon, 1/54)$ and $\bar{\vartheta}_2 = 1/27$, so $2\bar{\vartheta}_1 + 9\bar{\vartheta}_2 \leq 25/27 < 1$. Starting from $|\vartheta(i+m, l_\Delta+1)| = |\vartheta(p'')| \leq \bar{\vartheta}_1$, Lemma 9.6(ii) at $(i+k, l_\Delta+1)$ for $k = m-1, \dots, 0$ gives successively $2\vartheta(i+k, l_\Delta+1) = \vartheta(i+k, l_\Delta)$, hence $|\vartheta(i+k, l_\Delta+1)| \leq 1/54 \leq \bar{\vartheta}_1$. At $k=0$ this gives $|\vartheta(i, l_\Delta+1)| \leq \varepsilon/2$, so $(i, l_\Delta+1)$ is black, as $i < j(p') \leq J$. But it is not in Δ and is a cardinal neighbour of $(i, l_\Delta) \in \Delta$, hence white. $\square$

The range $1 \leq r \leq f_\omega$, $1 \leq o \leq 4$ of exit points is the *exit window*.

10. THE RENEWAL PROCESS

This section splits the raw letters (sums of consecutive pairs of letters) of a random word after each raw letter 4 or 5, where [10, §7.3] splits after each raw letter 3; the pieces are independent, so their hold points form a random walk with independent steps, the *renewal process* or hold walk of Lemma 10.4, along which the penalty of Section 11 factorizes. In place of the local limit bounds of [10, Lemmas 2.2, 7.6, 7.7], we prove exact laws, exponential moments and explicit tail bounds for it, used in Sections 11–14 and Appendix A.

Nothing here depends on n or ξ . Throughout, $a_1, a_2, \dots$ are the independent random letters of the random word W , and $\deg_i = \deg_i(W)$ are its partial degrees (Section 3), and $\binom{N}{k} = 0$ unless $0 \leq k \leq N$.

10.1. Raw letters and holds.

Definition 10.1 (Raw letters). The *raw letters* of a word are the pair sums $c_i = a_{2i-1} + a_{2i}$ ($i \geq 1$), so that $c_1 + \dots + c_k = \deg_{2k}$ ($\deg_0 = 0$). For $c \in \mathbb{Z}$ put $\mathbf{q}(c) = (c-1)2^{-c}$ if $c \geq 2$ and $\mathbf{q}(c) = 0$ otherwise. Since exactly $c-1$ pairs of positive integers have sum c , each of probability 2^{-c} , the raw letters are independent with law $\mathbb{P}(c_i = c) = \mathbf{q}(c)$. Below, a *word of raw letters* is a deterministic word $\mathbf{c} = (\mathbf{c}_1, \dots, \mathbf{c}_k) \in \mathbb{Z}^k$. For $k \in \mathbb{N}$ and $l \in \mathbb{Z}$ put

$$\mathbf{R}(k, l) = \mathbb{P}(\deg_{2k} = l) = \sum_{\substack{\mathbf{c} \in \mathbb{Z}^k, \mathbf{c}_i \geq 2 \\ \mathbf{c}_1 + \dots + \mathbf{c}_k = l}} \prod_{i=1}^k \mathbf{q}(\mathbf{c}_i).$$

The raw letters 4 and 5 are called *closing*.

All exact formulas below come from one classical observation [3, Section VI.8]: the letters are the waiting times between successes in a sequence of independent fair Bernoulli trials, so the set $\mathcal{A} = \{\deg_1, \deg_2, \dots\}$ of success times contains each positive integer independently with probability $1/2$, and $|\mathcal{A} \cap [1, l]|$ is binomial with parameters l and $1/2$.

Lemma 10.2 (Sums of raw letters). (1) $\sum_c \mathbf{q}(c) = 1$, and $\sum_{c>m} \mathbf{q}(c) = (m+1)2^{-m}$ for $m \in \mathbb{N}$.

(2) $\mathbf{R}(0, l) = \mathbf{1}[l = 0]$; for $k \geq 1$, $\mathbf{R}(k, l) = \binom{l-1}{2k-1} 2^{-l}$ if $l \geq 1$ and $\mathbf{R}(k, l) = 0$ if $l \leq 0$.

(3) $\sum_{k \geq 1} \mathbf{R}(k, l) = \frac{1}{4}$ for every integer $l \geq 2$.

(4) For $l, k \in \mathbb{N}$, $\mathbb{P}(\deg_{2k} \leq l < \deg_{2k+2}) = \binom{l+1}{2k+1} 2^{-l}$.

Proof. The raw letters are almost surely finite, so $\sum_c \mathbf{q}(c) = 1$, and $c_1 > m$ means $|\mathcal{A} \cap [1, m]| \leq 1$, of probability $(1+m)2^{-m}$. Part (2) is the negative binomial law [3, Section VI.8]: for $k, l \geq 1$, $\deg_{2k} = l$ means $l \in \mathcal{A}$ and $|\mathcal{A} \cap [1, l-1]| = 2k-1$. Summing over $k \geq 1$, $\sum_k \mathbf{R}(k, l)$ is the probability that $l \in \mathcal{A}$ and $|\mathcal{A} \cap [1, l-1]|$ is odd, which is $\frac{1}{2} \cdot \frac{1}{2}$ when $l-1 \geq 1$. Finally $\deg_{2k} \leq l < \deg_{2k+2}$ means $|\mathcal{A} \cap [1, l]| \in \{2k, 2k+1\}$, of probability $(\binom{l}{2k} + \binom{l}{2k+1})2^{-l} = \binom{l+1}{2k+1} 2^{-l}$. $\square$

We split the sequence of raw letters after each closing letter; the possible pieces are the hold words below.

Definition 10.3 (Hold words). For $(j, l) \in \mathcal{P}$, let $\mathfrak{B}_{j,l}$ be the set of *hold words* $b = (\mathbf{c}_1, \dots, \mathbf{c}_j) \in \mathbb{Z}^j$ with all $\mathbf{c}_i \geq 2$, with $\mathbf{c}_i \notin \{4, 5\}$ for $i < j$, with $\mathbf{c}_j \in \{4, 5\}$, and with $\mathbf{c}_1 + \dots + \mathbf{c}_j = l$; such a b has *length* $|b| = j$ and *hold point* $\text{pt}(b) = (j, l)$, its raw length and height. Let $\mathfrak{B} = \bigcup_{(j,l) \in \mathcal{P}} \mathfrak{B}_{j,l}$, a countable set. For a list $\mathbf{b} = (b^1, \dots, b^K) \in \mathfrak{B}^K$ put $\text{pt}_k(\mathbf{b}) = \text{pt}(b^1) + \dots + \text{pt}(b^k)$ for $0 \leq k \leq K$, so that $\text{pt}_0(\mathbf{b}) = (0, 0) \notin \mathcal{P}$ and $j(\text{pt}_k(\mathbf{b})) \geq k$; the *raw word* of $\mathbf{b}$ is the concatenation $b^1 b^2 \dots b^K$. For a word of raw letters $\mathbf{c} = (\mathbf{c}_1, \dots, \mathbf{c}_k)$ put $\mathbf{q}(\mathbf{c}) = \prod_{i=1}^k \mathbf{q}(\mathbf{c}_i)$, and for a list let $\mathbf{q}$ be $\mathbf{q}$ of its raw word. Put

$$\eta(y) = \sum_{b \in \mathfrak{B}: \text{pt}(b)=y} \mathbf{q}(b) \quad (y \in \mathcal{P}), \quad \eta^{\otimes N}(\mathbf{y}) = \prod_{i=1}^N \eta(y_i) \quad (\mathbf{y} \in \mathcal{P}^N), \quad (10.1)$$

with $\eta^{\otimes 0} = 1$, and $\nu_{45}(r) = \frac{5}{16} \left(\frac{11}{16}\right)^{r-1}$ for $r \geq 1$, $\nu_{45}(r) = 0$ for $r \leq 0$.

The pieces of the random raw letters are the *random holds*. A hold word is also called a *hold*.

Lemma 10.4 (Holds are independent). *Almost surely infinitely many raw letters are closing. The successive random holds are independent, each equal to a given hold word b with probability $\mathbf{q}(b)$. Consequently:*

- (1) *the hold points $\mathbf{h}_1, \mathbf{h}_2, \dots$ are independent with law η , the hold law, so $\sum_{y \in \mathcal{P}} \eta(y) = 1$ and $\eta^{\otimes N}$ is the law of $(\mathbf{h}_1, \dots, \mathbf{h}_N)$;*
- (2) *the raw length $j(\mathbf{h}_1)$ of the first hold point has the law*

$$\mathbb{P}(j(\mathbf{h}_1) = r) = \sum_l \eta(r, l) = \nu_{45}(r) \quad (r \geq 1); \quad (10.2)$$

- (3) *each $\mathfrak{B}_{j,l}$ is finite, and $\eta(j, l) \neq 0$ implies $l \geq 2j + 2 \geq 4$;*
- (4) *$\mathbf{W}_N = \mathbf{h}_1 + \dots + \mathbf{h}_N$, the hold walk, equals $(i, \deg_{2i})$, where i is the index of the N th closing raw letter; it is $\text{pt}_N(\mathbf{b})$ for the list $\mathbf{b}$ of the first N random holds;*
- (5) *for every $K \in \mathbb{N}$, $\sum_{\mathbf{b} \in \mathfrak{B}^K} \mathbf{q}(\mathbf{b}) = 1$, and if $0 \leq m \leq K$ and $\mathbf{c} \in \mathbb{Z}^m$, then*

$$\sum_{\mathbf{b} \in \mathfrak{B}^K} \mathbf{q}(\mathbf{b}) \mathbf{1}[\text{the raw word of } \mathbf{b} \text{ begins with } \mathbf{c}] = \prod_{i=1}^m \mathbf{q}(\mathbf{c}_i). \quad (10.3)$$

Proof. Each raw letter is closing with probability $\mathbf{q}(4) + \mathbf{q}(5) = \frac{3}{16} + \frac{1}{8} = \frac{5}{16}$, independently, so almost surely infinitely many are. Since a hold word has exactly one closing letter, at its end, the first N holds are given hold words $b^1, \dots, b^N$ if and only if the raw letters begin with their concatenation, an event of probability $\prod_{i=1}^N \mathbf{q}(b^i)$. Summing $\mathbf{q}(b)$ over the hold words b of length r factorizes as

$$\left(\sum_{c \notin \{4,5\}} \mathbf{q}(c) \right)^{r-1} (\mathbf{q}(4) + \mathbf{q}(5)) = \left(\frac{11}{16} \right)^{r-1} \frac{5}{16} = \nu_{45}(r),$$

by Lemma 10.2(1). Summing over $r \geq 1$ gives $\sum_b \mathbf{q}(b) = 1$, the sum running over all hold words, so the holds are independent with law $\mathbf{q}$ on hold words. This proves (1) and (2). In (3), the raw letters of a word in $\mathfrak{B}_{j,l}$ lie in $[2, l]$, and $l \geq 2(j-1) + 4$. Part (4) holds because the N th hold ends at the N th closing letter. In (5), the first K random holds equal $\mathbf{b}$ with probability $\mathbf{q}(\mathbf{b})$, which gives the total mass 1; their concatenation, the raw word of $\mathbf{b}$, has length at least $K \geq m$, so the event in (10.3) is that the first m raw letters equal $\mathbf{c}$, of probability $\prod_{i=1}^m \mathbf{q}(\mathbf{c}_i)$. $\square$

10.2. First passage over a height. The exit from a triangle in Section 9 is modelled by the first time the hold walk, started at gap ω below the top edge, rises above that edge.

Definition 10.5 (First passage). Let $\omega \in \mathbb{N}$. The *first-passage index* N_ω is the least $N \geq 1$ with $l(W_N) > \omega$. Since every $l(h_i) \geq 4$, it exists and $N_\omega \leq \lfloor \omega/4 \rfloor + 1$. The *first-passage law* is $F_\omega(x) = \mathbb{P}(W_{N_\omega} = x)$ for $x \in \mathbb{Z} \times \mathbb{Z}$, where, as for points, $j(x)$ and $l(x)$ denote the coordinates of x . Equivalently, call $\mathbf{y} \in \mathcal{P}^K$ a *first-passage list over height ω* if $l(y_1) + \dots + l(y_k) \leq \omega$ for $k < K$ and $l(y_1) + \dots + l(y_K) > \omega$; then $F_\omega(x)$ is the sum of $\eta^{\otimes K}(\mathbf{y})$ over $K \geq 1$ and the first-passage lists $\mathbf{y}$ with $y_1 + \dots + y_K = x$.

The *first-passage overshoot* over height ω is $l(W_{N_\omega}) - \omega \geq 1$. The *horizontal first-passage coordinate* over height ω is $j(W_{N_\omega})$, and its law $r \mapsto \sum_{x: j(x)=r} F_\omega(x)$ is the *horizontal first-passage law*.

Lemma 10.6 (First-passage law). *Let $\omega \in \mathbb{N}$.*

- (1) F_ω is a probability on $\mathbb{Z} \times \mathbb{Z}$, and $F_\omega(x) \neq 0$ implies $x \in \mathcal{P}$ (that is, $j(x) \geq 1$) and $l(x) > \omega$.
- (2) $W_{N_\omega} = (i, \deg_{2i})$ for the least index i with c_i closing and $\deg_{2i} > \omega$.

Proof. Part (1) holds since N_ω is finite, $j(W_{N_\omega}) \geq N_\omega \geq 1$, and $l(W_{N_\omega}) > \omega$. Part (2) follows from Lemma 10.4(4): the points W_N , $N \geq 1$, are the points $(i, \deg_{2i})$ at the closing indices i , in increasing order. $\square$

10.3. Exponential moments. Chernoff bounds for sums of hold points need the exponential moments of one hold point.

Definition 10.7. For real t and $y > 0$ put

$$M_{45}(t) = \sum_{r \geq 1} \nu_{45}(r) e^{tr} = \mathbb{E} e^{tj(h_1)}, \quad V(y) = \sum_{(j,l) \in \mathcal{P}} \eta(j,l) y^l = \mathbb{E} y^{l(h_1)},$$

with values in $[0, \infty]$.

Lemma 10.8 (Moments of one hold point). (1) If $11e^t < 16$, then $M_{45}(t) = 5e^t/(16 - 11e^t) < \infty$.

(2) If $0 \leq t \leq 1/1024$, then $M_{45}(t) \leq e^{13t/4}$.

(3) $M_{45}(1/16) \leq 5/4$.

(4) If $0 < y < 2$ and $V_{nc}(y) = \frac{y^2}{(2-y)^2} - \frac{3}{16}y^4 - \frac{1}{8}y^5 < 1$, then $V(y) = (\frac{3}{16}y^4 + \frac{1}{8}y^5)/(1 - V_{nc}(y))$.

(5) $V(27/25) < 8$ and $V(4095/4096)^{10} < (4095/4096)^{127}$.

Proof. Part (1) is a geometric series. For (2), $e^{-t} \geq 1 - t$ gives $16e^{-t} - 11 \geq 5 - 16t > 0$, so $M_{45}(t) = 5/(16e^{-t} - 11)$, and it suffices that $5 - 16t \geq 5e^{-13t/4}$. Since $e^{-13t/4} \leq 1 - \frac{13}{4}t + \frac{1}{2}(\frac{13}{4}t)^2$ for $t \geq 0$, the right side is at most $5 - \frac{65}{4}t + \frac{845}{32}t^2$, which is at most $5 - 16t$ because $t \leq 1/1024 < 8/845$. For (3), $e^{1/16} \leq 16/15$ by $e^{-1/16} \geq 15/16$, and by (1) $M_{45}(t) = 5e^t/(16 - 11e^t)$ increases with e^t on $(0, 16/11)$, with value $5/4$ at $e^t = 16/15$.

For (4), $y^l = \prod_{i=1}^j y^{c_i}$ for $b = (c_1, \dots, c_j) \in \mathfrak{B}_{j,l}$, so the factorization in the proof of Lemma 10.4 gives $V(y) = \sum_{j \geq 1} V_{nc}(y)^{j-1} (\frac{3}{16}y^4 + \frac{1}{8}y^5)$. Here $\sum_c q(c)y^c = \sum_{c \geq 2} (c-1)(y/2)^c = y^2/(2-y)^2$, so $V_{nc}(y)$ is the sum of $q(c)y^c$ over $c \notin \{4, 5\}$, a nonnegative number. Sum the geometric series.

For (5), exact evaluation gives $V_{nc}(27/25) = 77640184719/82656250000 < 1$ and $V(27/25) = 36266065281/5016065281 < 8$. Similarly $V_{nc}(4095/4096) < 1$ and

$$V(4095/4096) = \frac{48328698012322974813669375}{48479813739774803460507647};$$

the claim is the exact integer comparison

$$48328698012322974813669375^{10} 4096^{127} < 48479813739774803460507647^{10} 4095^{127}$$

(the ratio of the two sides is 0.99979...). $\square$

Lemma 10.9 (Tails of sums of hold points). *Let $N \in \mathbb{N}$, and let $\mathfrak{m}$, $\mathfrak{k}$ and $\mathfrak{t}$ be real. Then $\mathbb{E}e^{\mathfrak{t}(j(\mathbf{h}_1)+\dots+j(\mathbf{h}_N))} = \mathbf{M}_{45}(\mathfrak{t})^N$ in $[0, \infty]$, and*

$$\mathbb{P}\left(\mathfrak{m} \leq 10 \sum_{i \leq N} j(\mathbf{h}_i)\right) \leq \exp\left(\frac{N}{2} - \frac{\mathfrak{m}}{160}\right),$$

$$\mathbb{P}\left(\mathfrak{k} \leq \sum_{i \leq N} l(\mathbf{h}_i)\right) \leq 8^N \left(\frac{25}{27}\right)^{\mathfrak{k}},$$

$$\mathbb{P}\left(\sum_{i \leq N} l(\mathbf{h}_i) \leq \mathfrak{k}\right) \leq \left(\frac{4095}{4096}\right)^{\frac{127}{10}N - \mathfrak{k}}.$$

For $\omega \in \mathbb{N}$ and real $\mathfrak{t} \geq 0$ with $11e^{\mathfrak{t}} < 16$, $\sum_{x \in \mathcal{P}} \mathbf{F}_{\omega}(x) e^{\mathfrak{t}j(x)} \leq \mathbf{M}_{45}(\mathfrak{t})^{\omega+1} < \infty$.

Proof. The identity is independence. Each bound is the exponential Markov inequality: the three probabilities are at most

$$e^{-\mathfrak{m}/160} \mathbf{M}_{45}(1/16)^N, \quad (25/27)^{\mathfrak{k}} \mathbf{V}(27/25)^N, \quad y^{-\mathfrak{k}} \mathbf{V}(y)^N \quad (y = 4095/4096).$$

Apply Lemma 10.8, with $5/4 < 1 + \frac{1}{2} + \frac{1}{8} \leq e^{1/2}$ in the first case and $0 < \mathbf{V}(y) < y^{127/10}$ in the last. In the final claim the sum is $\mathbb{E}e^{\mathfrak{t}j(\mathbf{W}_{N_{\omega}})}$; since $N_{\omega} \leq \omega + 1$ and $\mathfrak{t} \geq 0$, $\mathfrak{t}j(\mathbf{W}_{N_{\omega}}) \leq \mathfrak{t} \sum_{i \leq \omega+1} j(\mathbf{h}_i)$, so the identity and Lemma 10.8(1) apply. $\square$

10.4. First-passage overshoot. Section 13 needs the first passage over a height not to overshoot it by much (Lemma 13.10).

Lemma 10.10 (First-passage overshoot tail). *For integers $\omega \geq 64$ and $\mathfrak{k} \geq 1$,*

$$\sum_{x: l(x) \geq \omega + \mathfrak{k}} \mathbf{F}_{\omega}(x) < 9 \left(\frac{25}{27}\right)^{\mathfrak{k}}.$$

Proof. Put $\mathbf{W}_0 = (0, 0)$ and, for $m \in \mathbb{Z}$, let $\mathbf{u}(m) = \sum_{N \geq 0} \mathbb{P}(l(\mathbf{W}_N) = m)$ be the expected number of $N \geq 0$ with $l(\mathbf{W}_N) = m$. By Lemma 10.4(4), the heights $l(\mathbf{W}_N)$, $N \geq 1$, are the $\deg_{2i}$ at the closing indices i . Since c_{k+1} is independent of $\deg_{2k}$,

$$\mathbf{u}(m) = \mathbf{1}[m = 0] + \sum_{k \geq 0} (\mathbf{q}(4)\mathbf{R}(k, m - 4) + \mathbf{q}(5)\mathbf{R}(k, m - 5)).$$

By Lemma 10.2(2),(3), $\sum_{k \geq 0} \mathbf{R}(k, l)$ is 1 at $l = 0$, is $1/4$ for $l \geq 2$, and vanishes otherwise. With $\mathbf{q}(4) = 3/16$ and $\mathbf{q}(5) = 1/8$, for $m \in \mathbb{N}$,

$$\mathbf{u}(m) = \mathbf{1}[m = 0] + \frac{3}{16}\mathbf{1}[m = 4] + \frac{1}{8}\mathbf{1}[m = 5] + \frac{3}{64}\mathbf{1}[m \geq 6] + \frac{1}{32}\mathbf{1}[m \geq 7],$$

so $\mathbf{u}(m) \leq \frac{5}{64} + \mathbf{1}[m = 0] + \frac{3}{16}\mathbf{1}[m = 4] + \frac{1}{8}\mathbf{1}[m = 5]$.

The left side of the claim is $\mathbb{P}(l(\mathbf{W}_{N_{\omega}}) \geq \omega + \mathfrak{k})$. On this event, $N = N_{\omega} - 1$ is the unique $N \geq 0$ with $l(\mathbf{W}_N) \leq \omega$ and $l(\mathbf{h}_{N+1}) \geq \omega + \mathfrak{k} - l(\mathbf{W}_N)$; conversely, any such N equals $N_{\omega} - 1$. Since $\mathbf{h}_{N+1}$ is independent of $\mathbf{W}_N$, with $y = 27/25$,

$$\mathbb{P}(l(\mathbf{W}_{N_{\omega}}) \geq \omega + \mathfrak{k}) = \sum_{m=0}^{\omega} \mathbf{u}(m) \mathbb{P}(l(\mathbf{h}_1) \geq \omega + \mathfrak{k} - m) \leq \mathbf{V}(y) y^{-\mathfrak{k}} \sum_{m=0}^{\omega} \mathbf{u}(m) y^{m-\omega},$$

and

$$\sum_{m=0}^{\omega} \mathbf{u}(m) y^{m-\omega} \leq \frac{5}{64} \sum_{i \geq 0} y^{-i} + y^{-\omega} \left(1 + \frac{3}{16}y^4 + \frac{1}{8}y^5\right) \leq \frac{135}{128} + \frac{9}{128} = \frac{9}{8}.$$

The last step uses the exact rational comparison $(25/27)^{64} \cdot 224805889/156250000 < 9/128$ (the left side is $0.0104\dots$), where $224805889/156250000 = 1 + \frac{3}{16}y^4 + \frac{1}{8}y^5$. Lemma 10.8(5) gives $\mathbf{V}(y) < 8$. $\square$

10.5. The horizontal first-passage law. Section 13 needs a packing bound and a tail bound for the horizontal first-passage law (Lemmas 10.14 and 10.15). Both follow from its exact form, a mixture of two binomial rows (Lemma 10.13); we first bound binomial rows.

Lemma 10.11 (Binomial bounds). *For $N \in \mathbb{N}$ and $k \in \mathbb{Z}$, $\left(\binom{N}{k} 2^{-N}\right)^2 \leq 2/(3N+2)$. For $N \geq 1$ and real $u \geq 0$,*

$$\sum_{|k-N/2| \geq u} \binom{N}{k} 2^{-N} \leq 2e^{-2u^2/N}.$$

Proof. We have $(3m+1)\binom{2m}{m}4^{-m} \leq 1$ by induction on m : it holds at $m=0$, and $\binom{2m+2}{m+1}4^{-m-1} = \binom{2m}{m}4^{-m}(2m+1)/(2m+2)$ with $(3m+4)(2m+1)^2 = 12m^3 + 28m^2 + 19m + 4 \leq (3m+1)(2m+2)^2$. The largest coefficient of row N is the middle one. For $N=2m$ this gives the claim; for $N=2m+1$, $\binom{2m+1}{m}2^{-N} = \binom{2m+2}{m+1}4^{-m-1}$, whose square is at most $1/(3m+4) = 2/(3N+5)$.

The second bound is Hoeffding's inequality [4, Theorem 2] for a sum of N independent fair Bernoulli variables, applied to each of the two tails. $\square$

Definition 10.12 (Binomial masses). For $l \in \mathbb{N}$ and $i \in \mathbb{Z}$ put

$$p_l(i) = \binom{l+1}{2i+1} 2^{-l}, \quad p_l^{\text{cl}}(i) = \frac{1}{8} \binom{l-1}{2i-1} 2^{-l} \quad (l, i \geq 1),$$

and $p_l^{\text{cl}}(i) = 0$ otherwise.

By Lemma 10.2, $p_l(i)$ is the probability that $\deg_{2i} \leq l < \deg_{2i+2}$, and for $l \geq 2$, $p_l^{\text{cl}}(i) = q(5)R(i, l)$ is the probability that $\deg_{2i} = l$ and $c_{i+1} = 5$. Odd-index sums of binomial rows give $\sum_i p_l(i) = 1$ and, for $l \geq 2$, $\sum_i p_l^{\text{cl}}(i) = 1/32$.

Lemma 10.13 (Exact horizontal first-passage law). *Let $\omega \geq 6$ and $r \geq 1$ be integers, and $l = \omega - 4$. Then*

$$\sum_{x: j(x)=r} F_\omega(x) = p_l^{\text{cl}}(r-1) + \sum_{k=0}^{r-2} (p_l(k) - p_l^{\text{cl}}(k)) \nu_{45}(r-1-k).$$

Proof. By Lemma 10.6(2), the left side is $\mathbb{P}(i_1 = r)$, where $i_1 = j(W_{N_\omega})$. Let i_0 be the largest $i \geq 0$ with $\deg_{2i} \leq l$; then $i_0 < i_1$. Closing indices $i \leq i_0$ have $\deg_{2i} \leq l \leq \omega$. If $c_{i_0+1} = 5$ and $\deg_{2i_0} = l$, then $\deg_{2i_0+2} = \omega + 1$ and $i_1 = i_0 + 1$. Otherwise a closing letter at $i_0 + 1$ gives $\deg_{2i_0+2} \leq \omega$, while every closing index $i \geq i_0 + 2$ has $\deg_{2i} \geq \deg_{2i_0+2} + 4 > \omega$; so i_1 is the first closing index after $i_0 + 1$. Therefore $\{i_1 = r\}$ is the disjoint union of $\{\deg_{2r-2} = l, c_r = 5\}$, of probability $p_l^{\text{cl}}(r-1)$ (zero for $r=1$, since $l \neq 0$), and, for $0 \leq k \leq r-2$, of the events

$$(\{\deg_{2k} \leq l < \deg_{2k+2}\} \setminus \{\deg_{2k} = l, c_{k+1} = 5\}) \cap \{c_{k+2}, \dots, c_{r-1} \text{ not closing}, c_r \text{ closing}\}.$$

The two parts are independent, with probabilities $p_l(k) - p_l^{\text{cl}}(k)$ and $\nu_{45}(r-1-k)$. $\square$

Lemma 10.14 (Packing of the horizontal first-passage coordinate). *Let $\omega \geq 1024$ be an integer, $d > 0$, and $\mathcal{K} \subseteq \mathbb{Z}$ with $|\zeta - \zeta'| \geq d$ for distinct $\zeta, \zeta' \in \mathcal{K}$. Then*

$$\sum_{\zeta \in \mathcal{K}} \sum_{x: j(x)=\zeta} F_\omega(x) < \frac{17}{10\sqrt{\omega}} + \frac{33}{32d}.$$

Proof. Put $l = \omega - 4$. By Lemma 10.13, with $p_l^{\text{cl}} \geq 0$,

$$\sum_{x: j(x)=r} F_\omega(x) \leq p_l^{\text{cl}}(r-1) + \sum_{k \geq 1} \nu_{45}(k) p_l(r-1-k)$$

for $r \geq 1$, and the left side vanishes for $r \leq 0$. Let $\mathbf{g} : \mathbb{Z} \rightarrow [0, \infty)$ have finite support, with every set $\{\mathbf{g} \geq y\}$, $y > 0$, consisting of consecutive integers. Such a set I contains at most $1 + |I|/d$ points of $\mathcal{K}$, so

$$\sum_{\zeta \in \mathcal{K}} \mathbf{g}(\zeta) = \int_0^{\max \mathbf{g}} |\mathcal{K} \cap \{\mathbf{g} \geq y\}| dy \leq \int_0^{\max \mathbf{g}} \left(1 + \frac{|\{\mathbf{g} \geq y\}|}{d}\right) dy = \max \mathbf{g} + \frac{1}{d} \sum_x \mathbf{g}(x).$$

Binomial rows are unimodal, so every superlevel set of $i \mapsto \binom{N}{i}$ is an interval, and so are its preimages under $i \mapsto 2i \pm 1$ and their translates. Thus this bound applies to $i \mapsto \mathbf{p}_l^{\text{cl}}(i-1)$ and to each $i \mapsto \mathbf{p}_l(i-1-k)$, of total masses $1/32$ and 1 . Since $\sum_k \nu_{45}(k) = 1$, the left side of the claim is at most $\max \mathbf{p}_l^{\text{cl}} + \max \mathbf{p}_l + \frac{33}{32d}$.

Finally, Lemma 10.11 with rows $l+1$ and $l-1$ gives, for all $i \in \mathbb{Z}$,

$$\mathbf{p}_l(i) \leq 2\sqrt{\frac{2}{3\omega-7}}, \quad \mathbf{p}_l^{\text{cl}}(i) \leq \frac{1}{16}\sqrt{\frac{2}{3\omega-13}}.$$

Both $\omega/(3\omega-7)$ and $\omega/(3\omega-13)$ decrease in $\omega > 13/3$. At $\omega = 1024$ the exact comparisons $8192/3065 < (41/25)^2$ and $8/3059 < (3/50)^2$ give $\max \mathbf{p}_l < \frac{41}{25\sqrt{\omega}}$ and $\max \mathbf{p}_l^{\text{cl}} < \frac{3}{50\sqrt{\omega}}$, and $\frac{41}{25} + \frac{3}{50} = \frac{17}{10}$. $\square$

Lemma 10.15 (Horizontal tail of the first passage). *For $\omega \in \mathbb{N}$ and real $u \geq 0$,*

$$\sum_{x: |j(x)-\omega/4| \geq u} F_\omega(x) \leq 2^{54} \left[\exp\left(-\frac{u^2}{2^{15}(1+\omega)}\right) + \exp\left(-\frac{u}{256}\right) \right].$$

Proof. The left side is at most 1. If $u \leq 4096$, then $e^{-u/256} \geq e^{-16} > 3^{-16} > 2^{-26}$, so the right side exceeds 1. Let $u > 4096$.

Case 1. Suppose $\omega \leq 5$. A nonzero term has $j(x) \geq 1$, so $j(x) - \omega/4 > -u$; hence $j(x) - \omega/4 \geq u$ and $j(x) \geq u$. Markov's inequality, Lemma 10.9 with exponent $1/16$ and Lemma 10.8(3) bound the left side by $(5/4)^{\omega+1} e^{-u/16} \leq (5/4)^6 e^{-u/16} \leq 4e^{-u/256}$.

Case 2. Suppose $\omega \geq 6$. Put $l = \omega - 4 \geq 2$; then $r - \omega/4 = (r-1) - l/4$ for $r = j(x)$. Let $\mathcal{J} = \{i \geq 0 : |i - l/4| \geq u/2\}$. By Pascal's rule $\mathbf{p}_l(i) = (\binom{l}{2i} + \binom{l}{2i+1})2^{-l} \geq \mathbf{p}_l^{\text{cl}}(i)$. For $i \in \mathcal{J}$ and $i' \in \{2i, 2i+1\}$ we have $|i' - l/2| \geq 2|i - l/4| - 1 \geq u/2$, so Lemma 10.11 gives $\sum_{i \in \mathcal{J}} \mathbf{p}_l(i) \leq 2e^{-u^2/(2l)}$. Sum Lemma 10.13 over the r with $|(r-1) - l/4| \geq u$. The $\mathbf{p}_l^{\text{cl}}$ -part is at most $\sum_{i \in \mathcal{J}} \mathbf{p}_l(i)$. In the remaining part, with $k = r - 1 - i \geq 1$, $|i + k - l/4| \geq u$ forces $i \in \mathcal{J}$ or $k \geq u/2$, so it is at most

$$\sum_{i \in \mathcal{J}} \mathbf{p}_l(i) + \sum_{k \geq u/2} \nu_{45}(k) \leq 2e^{-u^2/(2l)} + \left(\frac{11}{16}\right)^{u/2-1} \leq 2e^{-u^2/(2l)} + 2e^{-5u/32},$$

using $\log(16/11) \geq 5/16$. The total is at most $4e^{-u^2/(2l)} + 2e^{-u/256}$, and $2l < 2^{15}(1+\omega)$. $\square$

10.6. The exit window. The exit mass is the probability that the first passage lands in the exit window (Lemma 10.17), whose points are white (Lemma 9.14); it bounds below the reward of one passage (Lemma 12.5) and the gain at black points near the top of a triangle (Lemma 14.4).

Definition 10.16 (Exit mass). For $\omega \in \mathbb{N}$, with f_ω as in Lemma 9.14, put

$$\mathbf{q}_{45}(\omega) = \sum_{c=4}^5 \mathbf{q}(c) \sum_{o=1}^4 \sum_{r=1}^{f_\omega} R(r-1, \omega + o - c).$$

Lemma 10.17 (Exit masses as first-passage probabilities). *Let $\omega \in \mathbb{N}$. Then*

$$\sum_{x: j(x) \leq f_\omega, \omega < l(x) \leq \omega+4} F_\omega(x) = \mathbf{q}_{45}(\omega).$$

Proof. If c_r is closing and $\omega < \deg_{2r} \leq \omega + 4$, then $\deg_{2r-2} = \deg_{2r} - c_r \leq \omega$, so every earlier closing index i has $\deg_{2i} \leq \omega$ and $r = j(W_{N_\omega})$. Hence, by Lemma 10.6(2), the left side is the probability that some $r \leq f_\omega$ has $c_r \in \{4, 5\}$ and $\deg_{2r} = \omega + o$ with $1 \leq o \leq 4$. These events are disjoint in (r, c_r, o) , and c_r is independent of $\deg_{2r-2}$, which gives $\mathbf{q}_{45}(\omega)$. $\square$

Lemma 10.18 (Lower bounds for $\mathbf{q}_{45}$). *For every integer $\omega \geq 256$, $\mathbf{q}_{45}(\omega) \geq \frac{5}{16} - \frac{125}{8\omega} \geq \frac{515}{2048}$. For every $\omega \in \mathbb{N}$, $\mathbf{q}_{45}(\omega) \geq \frac{3}{16}$.*

Proof. Let $\omega \geq 256$, fix $c \in \{4, 5\}$ and $o \in \{1, \dots, 4\}$, and put $l = \omega + o - c \in [\omega - 4, \omega]$. By Lemma 10.2(2),(3),

$$\sum_{r=1}^{f_\omega} R(r-1, l) \geq \frac{1}{4} - \sum_{k \geq f_\omega} R(k, l), \quad \sum_{k \geq f_\omega} R(k, l) \leq \frac{1}{2} \mathbb{P}(|\mathcal{A} \cap [1, l-1]| \geq 2f_\omega - 1),$$

where $|\mathcal{A} \cap [1, l-1]|$ is binomial with parameters $l-1$ and $1/2$. Since $16f_\omega \geq 5\omega + 1$ and $l \leq \omega$, $2f_\omega - 1 - \frac{l-1}{2} \geq \frac{5\omega+1}{8} - 1 - \frac{\omega-1}{2} = \frac{\omega-3}{8} \geq \frac{\omega}{10}$, so Chebyshev's inequality bounds the probability by $\frac{l-1}{4}(10/\omega)^2 \leq 25/\omega$. Summing over the eight pairs (c, o) , with $\mathbf{q}(4) + \mathbf{q}(5) = 5/16$, gives $\mathbf{q}_{45}(\omega) \geq \frac{5}{4}(\frac{1}{4} - \frac{25}{2\omega}) = \frac{5}{16} - \frac{125}{8\omega}$. For $\omega \geq 256$ this is at least $515/2048 > 3/16$.

For $\omega \leq 255$, $\mathbf{q}_{45}(\omega)$ is an explicit rational computed from the definition with R given by Lemma 10.2(2). At $\omega = 0$ only $r = 1$, $c = o = 4$ contributes, and $\mathbf{q}_{45}(0) = \mathbf{q}(4) = 3/16$. For $1 \leq \omega \leq 255$, exact evaluation of these 255 rationals shows $\mathbf{q}_{45}(\omega) > 3/16$; the minimum is $\mathbf{q}_{45}(6) = 227/1024$. $\square$

11. THE CHARACTER SUM AND THE WEIGHTED MAXIMA

This section bounds $|\hat{\mu}_n(\xi)|$ by the expectation $\sum_y \eta(y)Q(y)$ of the expected penalty Q (Proposition 11.8, used in Theorem 14.6), following [10, §§7.1, 7.3] but with cancellation at the raw letters 4 and 5 as well as 3 (Lemma 11.2). It then introduces the weighted maxima Q_m , whose descent step $Q_m \leq Q_{m-1}$ is proved in Sections 12–14.

The constants $z = \frac{21}{500}$ and $\kappa = \frac{4}{25}$ are those of Appendix A.1. Every series below is indexed by a countable set and has nonnegative terms or converges absolutely, so it may be rearranged and regrouped freely.

11.1. Pairing. Grouping the letters in pairs, as in [10, §7.1], bounds $|\hat{\mu}_n(\xi)|$ by an average of products of the following factors, one for each raw letter $\mathbf{c}_i$ at its point $(i, \mathbf{c}_1 + \dots + \mathbf{c}_i)$. For $p \in \mathcal{P}$ and an integer $c \geq 2$ put

$$\varphi_c(p) = \frac{1}{c-1} \left| \sum_{k=1}^{c-1} e^{-2\pi i(2^{k-1}-1)\vartheta(p)} \right|, \quad (11.1)$$

and put $\varphi_c(p) = 1$ for $c \leq 1$. By the triangle inequality, $0 \leq \varphi_c(p) \leq 1$.

Lemma 11.1 (Pairing).

$$|\hat{\mu}_n(\xi)| \leq \sum_{\mathbf{c} \in \mathbb{Z}^J} \prod_{i=1}^J \mathbf{q}(\mathbf{c}_i) \varphi_{\mathbf{c}_i}(i, \mathbf{c}_1 + \dots + \mathbf{c}_i). \quad (11.2)$$

Proof. Let $\chi(y) = e^{-2\pi i \xi y / 3^n}$ for $y \in G_n$, evaluated at elements of $\mathbb{Z}[1/2]$ after reduction (Section 3). Then $\hat{\mu}_n(\xi) = \mathbb{E}\chi(\text{off}(W_{\leq n}))$ is the sum $S_\chi(n)$ of [10, (7.2)], since the expression in [10, (1.29)] is $\text{off}(W_{\leq n})$ letter for letter. Tao's pair sums $a_{2i-1} + a_{2i}$ are the raw letters, with law $\mathbf{q}$ (Definition 10.1), and given $a_{2i-1} + a_{2i} = c$ the letter a_{2i} is uniform on $\{1, \dots, c-1\}$. So, with $p_i = (i, \mathbf{c}_1 + \dots + \mathbf{c}_i)$ and $x_i = 3^{2(i-1)}2^{-l(p_i)}$, [10, (7.4)–(7.5)] read

$$|\hat{\mu}_n(\xi)| \leq \sum_{\mathbf{c} \in \mathbb{Z}^J} \prod_{i=1}^J \mathbf{q}(\mathbf{c}_i) |f(x_i, \mathbf{c}_i)|, \quad f(x, c) = \frac{1}{c-1} \sum_{k=1}^{c-1} \chi((2^k + 3)x),$$

where the terms with some $\mathbf{c}_i \leq 1$ vanish. Since $2^k + 3 = 5 + 2(2^{k-1} - 1)$ and $\chi(2x_i) = e^{-2\pi i \vartheta(p_i)}$ (Definition 9.2), we have $\chi((2^k + 3)x_i) = \chi(5x_i) e^{-2\pi i (2^{k-1} - 1)\vartheta(p_i)}$, so $|f(x_i, \mathbf{c}_i)| = \varphi_{\mathbf{c}_i}(p_i)$. $\square$

11.2. Cancellation at white points. At a white point the factors $\varphi_3, \varphi_4, \varphi_5$ are bounded away from 1.

Lemma 11.2 (White cancellation). *If $p \in \mathcal{P}$ is white, then*

$$\varphi_3(p) < e^{-\kappa z} = e^{-21/3125}, \quad \varphi_4(p) < e^{-z}, \quad \varphi_5(p) < e^{-z} = e^{-21/500}.$$

Proof. Put $x = \pi|\vartheta(p)| \in (\pi\varepsilon, \frac{\pi}{2}]$, $\mathbf{a} = e^{-2ix}$ and $x_o = 3.1415\varepsilon = \frac{6283}{54250}$. Replacing $\vartheta(p)$ by $-\vartheta(p)$ conjugates the sum in (11.1), so $\varphi_c(p) = \frac{1}{c-1} |\sum_{k=1}^{c-1} \mathbf{a}^{2^{k-1}-1}|$. Since $\pi > 3.1415$,

$$0 < x_o < \pi\varepsilon < x \leq \frac{\pi}{2}.$$

We use the Taylor bounds $\sin y \geq y - y^3/6$ for $y \geq 0$, $\cos y \leq 1 - y^2/2 + y^4/24$, and $e^{-y} \geq 1 - y + y^2/2 - y^3/6$ for $y \geq 0$. The exponents $2^{k-1} - 1$ in (11.1) are 0, 1, 3, 7 for $k = 1, 2, 3, 4$.

The raw letter 3. As in [10, Lemma 7.2], $\varphi_3(p) = \frac{1}{2}|1 + \mathbf{a}| = \cos x$. As $\cos$ decreases on $[0, \pi]$, $\cos x < \cos x_o \leq 1 - x_o^2/2 + x_o^4/24$. Since $\kappa z = \frac{21}{3125}$, the exact rational comparison

$$1 - \frac{x_o^2}{2} + \frac{x_o^4}{24} < 1 - \kappa z + \frac{(\kappa z)^2}{2} - \frac{(\kappa z)^3}{6} \leq e^{-\kappa z}$$

(decimal values 0.9933009 and 0.9933025) proves the first bound.

The raw letter 4. Here $\varphi_4(p) = \frac{1}{3}|1 + \mathbf{a} + \mathbf{a}^3|$, and since $|\mathbf{a}| = 1$,

$$|1 + \mathbf{a} + \mathbf{a}^3|^2 = 3 + 2\cos 2x + 2\cos 4x + 2\cos 6x = 9 - 4\mathfrak{h}(\mathbf{u}), \quad \mathfrak{h}(\mathbf{u}) = 14\mathbf{u} - 28\mathbf{u}^2 + 16\mathbf{u}^3,$$

by the multiple-angle formulas in $\mathbf{u} = \sin^2 x$. As $\sin$ increases on $[0, \frac{\pi}{2}]$, $\sin x > \sin x_o \geq x_o - x_o^3/6 > 0$, so $\mathbf{u} > \mathbf{u}_o = (x_o - x_o^3/6)^2$. If $\mathbf{u} \leq \frac{1}{4}$, then $\mathfrak{h}$ is nondecreasing on $[0, \frac{1}{4}]$ (as $\mathfrak{h}'(y) \geq 14 - 56y$), so $\mathfrak{h}(\mathbf{u}) \geq \mathfrak{h}(\mathbf{u}_o)$. If $\mathbf{u} > \frac{1}{4}$, then $\mathfrak{h}(\mathbf{u}) - \frac{7}{4}\mathbf{u} = 16\mathbf{u}(\mathbf{u} - \frac{7}{8})^2 \geq 0$ gives $\mathfrak{h}(\mathbf{u}) > \frac{7}{16}$. Exact rational arithmetic gives $\frac{18199}{100000} < \mathfrak{h}(\mathbf{u}_o) < \frac{7}{16}$ and

$$9 - 4 \cdot \frac{18199}{100000} < 9 \left(1 - 2z + 2z^2 - \frac{4}{3}z^3 \right) \leq 9e^{-2z}$$

(decimal values 8.27204 and 8.27486; the last step is the Taylor bound at $y = 2z$). In both cases $\mathfrak{h}(\mathbf{u}) > \frac{18199}{100000}$, hence $9\varphi_4(p)^2 = 9 - 4\mathfrak{h}(\mathbf{u}) < 9e^{-2z}$.

The raw letter 5. By the triangle inequality and $|1 + e^{-2iy}| = 2|\cos y|$,

$$\varphi_5(p) = \frac{1}{4}|1 + \mathbf{a} + \mathbf{a}^3 + \mathbf{a}^7| \leq \frac{1}{4}(|1 + \mathbf{a}| + |1 + \mathbf{a}^4|) = \frac{1}{2}(|\cos x| + |\cos 4x|).$$

If $x \leq \frac{\pi}{8}$, then $4x_o < 4x \leq \frac{\pi}{2}$, so $0 \leq \cos 4x < \cos 4x_o$, and $\varphi_5(p) < 1 - (4x_o)^2/4 + (4x_o)^4/48$, which is less than 0.94731. If $\frac{\pi}{8} < x < \frac{\pi}{6}$, then $|\cos 4x| < \frac{1}{2}$ and $\varphi_5(p) < \frac{3}{4}$. If $x \geq \frac{\pi}{6}$, then $0 \leq \cos x \leq \frac{\sqrt{3}}{2} < \frac{87}{100}$ and $\varphi_5(p) < \frac{187}{200}$. All three bounds are below $1 - z = \frac{479}{500} \leq e^{-z}$. $\square$

The penalty below records these bounds along a word of raw letters.

Definition 11.3 (White factors and penalty). For $p \in \mathcal{P}$ the *white factors* are $w(p) = e^{-z}$ and $w_3(p) = e^{-\kappa z}$ if p is white, and $w(p) = w_3(p) = 1$ otherwise. For a word of raw letters $\mathbf{c} = (\mathbf{c}_1, \dots, \mathbf{c}_j) \in \mathbb{Z}^j$, the *penalty* is

$$\text{Pen}(\mathbf{c}) = \prod_{i: \mathbf{c}_i \in \{4,5\}} w(i, \mathbf{c}_1 + \dots + \mathbf{c}_i) \prod_{i: \mathbf{c}_i = 3} w_3(i, \mathbf{c}_1 + \dots + \mathbf{c}_i) \in (0, 1].$$

11.3. Splitting the penalty into holds. Splitting the raw letters into holds turns the penalty into a product along the hold walk: the closing letters carry w at the hold points, and the raw letters 3 inside a hold carry the internal factor (Lemma 11.5).

Definition 11.4 (Internal factors). For $x \in \mathbb{N} \times \mathbb{Z}$ and a hold $b = (\mathbf{c}_1, \dots, \mathbf{c}_{|b|})$ (Definition 10.3), the *internal factor*

$$l(x; b) = \prod_{i: \mathbf{c}_i=3} w_3(x + (i, \mathbf{c}_1 + \dots + \mathbf{c}_i)) \in (0, 1]$$

records the raw letters 3 inside the hold.

Lemma 11.5 (Penalty along holds). *Let $K \geq J$, let $\mathbf{b} \in \mathfrak{B}^K$, and let $\mathbf{c}$ be the first J raw letters of its raw word. Then*

$$\text{Pen}(\mathbf{c}) = \prod_{k=1}^K w(\text{pt}_k(\mathbf{b})) \prod_{k=1}^K l(\text{pt}_{k-1}(\mathbf{b}); b^k).$$

Proof. The raw letters $\mathbf{c}_1^k, \dots, \mathbf{c}_{|b^k|}^k$ of b^k occupy positions $j(\text{pt}_{k-1}(\mathbf{b})) + 1, \dots, j(\text{pt}_k(\mathbf{b}))$ of the raw word, and the sum of the raw letters of the raw word up to $\mathbf{c}_i^k$ is $l(\text{pt}_{k-1}(\mathbf{b})) + \mathbf{c}_1^k + \dots + \mathbf{c}_i^k$. Hence the point attached to $\mathbf{c}_i^k$ in Definition 11.3 is $\text{pt}_{k-1}(\mathbf{b}) + (i, \mathbf{c}_1^k + \dots + \mathbf{c}_i^k)$, which for the closing letter $i = |b^k|$ is $\text{pt}_k(\mathbf{b})$. The raw letters in $\{4, 5\}$ are exactly the closing letters, and the raw letters equal to 3 are among the others. So the right side is the penalty of the entire raw word, whose length is at least $K \geq J$. The raw letters beyond position J have points of raw length greater than J , which are not white, so their factors are 1. $\square$

11.4. Reduction to the expected penalty. The character sum is bounded by the average penalty of the hold walk (Proposition 11.8), the analogue of [10, (7.34)]. We truncate Tao's infinite product after K holds, so that $\mathbf{U}_K$ below is a sum over finite lists of holds, with no infinite product of random factors.

Definition 11.6 (The expected penalty). For $K \in \mathbb{N}$ and $p \in \mathcal{P}$ let

$$\mathbf{U}_K(p) = \sum_{\mathbf{b} \in \mathfrak{B}^K} \mathbf{q}(\mathbf{b}) \prod_{k=0}^K w(p + \text{pt}_k(\mathbf{b})) \prod_{k=1}^K l(p + \text{pt}_{k-1}(\mathbf{b}); b^k),$$

and call $\mathbf{Q}(p) = \mathbf{U}_J(p)$ the *expected penalty* of a hold walk started at p , counting the closing letter at p itself.

Lemma 11.7 (Properties of the expected penalty). *Let $p \in \mathcal{P}$.*

- (1) $0 \leq \mathbf{U}_K(p) \leq 1$ for all K , and $\mathbf{U}_K(p) = \mathbf{Q}(p)$ whenever $j(p) + K \geq J$. If $j(p) > J$, then $\mathbf{Q}(p) = 1$.
- (2) $\mathbf{Q}(p) = w(p) \sum_{b \in \mathfrak{B}} \mathbf{q}(b) l(p; b) \mathbf{Q}(p + \text{pt}(b))$.
- (3) $\mathbf{Q}(p) \leq w(p) \sum_{y \in \mathcal{P}} \eta(y) \mathbf{Q}(p + y)$.

Proof. Part (2) is the analogue of [10, (7.35)], with internal factors and truncated sums. All factors in Definition 11.6 lie in $(0, 1]$ and the lists have total mass 1 (Lemma 10.4(5)), so $0 \leq \mathbf{U}_K \leq 1$. Writing a list in $\mathfrak{B}^{K+1}$ as $(b, \mathbf{b}')$ with $\text{pt}_i(b, \mathbf{b}') = \text{pt}(b) + \text{pt}_{i-1}(\mathbf{b}')$ for $i \geq 1$ gives the one-step identity

$$\mathbf{U}_{K+1}(p) = w(p) \sum_{b \in \mathfrak{B}} \mathbf{q}(b) l(p; b) \mathbf{U}_K(p + \text{pt}(b)). \quad (11.3)$$

Splitting off the last hold instead shows that $\mathbf{U}_{K+1}(p) = \mathbf{U}_K(p)$ when $j(p) + K \geq J$: the new factors $w(p + \text{pt}_{K+1}(\mathbf{b}))$ and $l(p + \text{pt}_K(\mathbf{b}); b^{K+1})$ are evaluated at points of raw length at least $j(p) + K + 1 > J$, which are not white, so they equal 1, and the last hold then sums to mass 1. By

induction, $U_K(p)$ does not depend on K once $j(p) + K \geq J$; as $K = J$ qualifies, the common value is $Q(p)$. If $j(p) > J$, the same reasoning makes every factor equal to 1, and $Q(p) = 1$. This proves (1). For (2), apply (11.3) with $K = J$, noting $U_{J+1}(p) = Q(p)$ by (1). For (3), drop $l \leq 1$ in (2) and group the holds by their point using (10.1). $\square$

Proposition 11.8 (Character sum and expected penalty).

$$|\hat{\mu}_n(\xi)| \leq \sum_{y \in \mathcal{P}} \eta(y) Q(y). \quad (11.4)$$

Proof. If $J = 0$, then $Q(y) = 1$ for every $y \in \mathcal{P}$ (Lemma 11.7(1)), so the right side is 1, while $|\hat{\mu}_n(\xi)| \leq \sum_x \mu_n(x) = 1$.

Let $J \geq 1$ and $\mathbf{c} \in \mathbb{Z}^J$. The factor of $\text{Pen}(\mathbf{c})$ at i is $1 \geq \varphi_{\mathbf{c}_i}(i, \mathbf{c}_1 + \cdots + \mathbf{c}_i)$ unless $\mathbf{c}_i \in \{3, 4, 5\}$ and $(i, \mathbf{c}_1 + \cdots + \mathbf{c}_i)$ is white, and then Lemma 11.2 applies. So (11.2) gives

$$|\hat{\mu}_n(\xi)| \leq \sum_{\mathbf{c} \in \mathbb{Z}^J} \prod_{i=1}^J q(\mathbf{c}_i) \text{Pen}(\mathbf{c}).$$

By (10.3) with $m = K = J$, the mass $\prod_i q(\mathbf{c}_i)$ is the total mass of the lists $\mathbf{b} \in \mathfrak{B}^J$ whose raw word begins with $\mathbf{c}$. Summing over $\mathbf{c}$ first and applying Lemma 11.5, the right side becomes

$$\sum_{\mathbf{b} \in \mathfrak{B}^J} q(\mathbf{b}) \prod_{k=1}^J w(\text{pt}_k(\mathbf{b})) \prod_{k=1}^J l(\text{pt}_{k-1}(\mathbf{b}); b^k).$$

Write $\mathbf{b} = (b^1, \mathbf{b}')$ and $y = \text{pt}(b^1)$, so that $\text{pt}_k(\mathbf{b}) = y + \text{pt}_{k-1}(\mathbf{b}')$ for $k \geq 1$. Summing over $\mathbf{b}' \in \mathfrak{B}^{J-1}$ turns the summand into $q(b^1) l((0, 0); b^1) U_{J-1}(y)$, by Definition 11.6. Since $j(y) \geq 1$, Lemma 11.7(1) gives $U_{J-1}(y) = Q(y)$. Finally drop $l \leq 1$ and group the holds b^1 by their point, using (10.1). $\square$

Two cases of the descent step halt the hold walk started at a black point p when its height first exceeds $l(p) + \omega$, the top edge of a triangle at gap ω above p , and continue from there (Lemmas 14.4 and 13.16). The expected penalty at p is at most its average at the halting point.

Lemma 11.9 (Halting at a first passage). *For every $\omega \in \mathbb{N}$ and $p \in \mathcal{P}$,*

$$Q(p) \leq \sum_{x \in \mathcal{P}} F_\omega(x) Q(p + x).$$

Proof. Compare [10, (7.45)]. Put $W_0 = (0, 0)$. The hold points are independent with law η (Lemma 10.4(1)), so Lemma 11.7(3) and $w \leq 1$ give $\mathbb{E}[Q(p + W_{k+1}) \mid h_1, \dots, h_k] = \sum_y \eta(y) Q(p + W_k + y) \geq Q(p + W_k)$: the process $k \mapsto Q(p + W_k) \in [0, 1]$ is a submartingale. The first-passage index N_ω is a stopping time in the sense of [2], with $N_\omega \leq \lfloor \omega/4 \rfloor + 1$ (Definition 10.5), so optional stopping [2, Theorem 4.4.1] gives $Q(p) \leq \mathbb{E} Q(p + W_{N_\omega}) = \sum_x F_\omega(x) Q(p + x)$, where the sum may be taken over $x \in \mathcal{P}$ by Lemma 10.6(1). $\square$

11.5. Weighted maxima near the horizon. To show that $Q(p)$ decays like a power of the distance to the horizon $j = J$, we follow [10, (7.38)–(7.41)] and track the largest weighted value $d_J^{A_*} Q$ within distance m of the horizon; a bound on it that does not grow with m is the decay we need.

Definition 11.10 (Weighted maxima). For $p \in \mathcal{P}$ let $d_J(p) = \max(J - j(p), 1)$, the *distance to the horizon* $j = J$. For an integer $m \geq 1$, a *boundary point* for m is a $p \in \mathcal{P}$ with $j(p) + m = J$. With the exponent $A_* = \frac{10241}{4096}$ of (4.1), for an integer $m \geq 0$ the *weighted maxima* are

$$Q_m = \sup\{d_J(p)^{A_*} Q(p) : p \in \mathcal{P}, j(p) \geq J - m\}.$$

Lemma 11.11 (Basic properties of the weighted maxima). *Let $m \geq 0$ be an integer.*

- (1) $0 \leq Q_m \leq \max(m, 1)^{A^*}$.
- (2) If $j(p) \geq J - m$, then $d_J(p)^{A^*} Q(p) \leq Q_m$.
- (3) If $0 \leq m \leq m'$, then $Q_m \leq Q_{m'}$.

Proof. As in [10, (7.39)]: if $j(p) \geq J - m$, then $1 \leq d_J(p) \leq \max(m, 1)$ and $0 \leq Q(p) \leq 1$ (Lemma 11.7(1)), and the set in Definition 11.10 contains the point $(J + 1, 0)$. This gives (1) and (2); for (3), the sets increase with m . $\square$

Lemma 11.12 (Boundary reduction). *Let $m \geq 1$ be an integer.*

- (1) If $r \geq 1$ and $j(p) + m = J + r$, then $Q(p) \leq \max(m - r, 1)^{-A^*} Q_{m-1}$.
- (2) If $Q(p) \leq m^{-A^*} Q_{m-1}$ for every boundary point p for m , then $Q_m \leq Q_{m-1}$.

Proof. As in [10, §7.4] (the bound opening Case 1, and the reduction of (7.40) to (7.41)), by Lemma 11.11(2): in (1), $j(p) \geq J - (m - 1)$ and $d_J(p) = \max(m - r, 1)$; in (2), a point with $j(p) \geq J - m$ either has $j(p) \geq J - (m - 1)$ or is a boundary point for m , where $d_J(p) = m$. $\square$

By Lemma 11.12(2), the descent step $Q_m \leq Q_{m-1}$ (Proposition 14.5) reduces to the bound $Q(p) \leq m^{-A^*} Q_{m-1}$ at every boundary point p for m , which is white or black as $j(p) < J$. We prove it in three cases: *Case 1*, p is white; *Case 2*, p is black with $\text{gap}(p) \leq m/(\log m)^2$; and *Case 3*, p is black with $\text{gap}(p) > m/(\log m)^2$. Cases 1 and 2 need one hold or one passage over the top edge of a triangle (Section 14); Case 3 needs the trace (Section 12) and a packing estimate (Section 13).

12. THE TRACE

In Case 3 a hold walk with low reward must enter many triangles (Section 13). Each entry is preceded by a passage over the top edge of the previous triangle, which earns a reward with probability bounded below, so many entries are unlikely. This section proves this in the explicit form of [10, Lemma 7.9]: from any starting point, the average of the discounted reward $e^{-\gamma \text{Rw}}$ at the $(k + 1)$ st trace time is at most $r_2^{\lfloor k/2 \rfloor} (1 - d_{\text{tr}})^{k-2\lfloor k/2 \rfloor}$ (Lemma 12.11). The proof rests on bounds for the powers of the entry kernel (Proposition 12.10). The constants are those of Definitions A.2 and A.12. All sums below have nonnegative terms and values in $[0, \infty]$ unless stated otherwise, and are rearranged freely.

12.1. Paths and rewards. A path follows the hold walk for a fixed number of holds, and its reward counts the white points where the penalty of Section 11 receives a factor below 1.

Definition 12.1 (Paths and rewards). For a list $\mathbf{b} = (b^1, \dots, b^L) \in \mathfrak{B}^L$, $L \in \mathbb{N}$, write $|\mathbf{b}| = L$ and $\mathbf{b}_{[i, i']}$ $= (b^i, \dots, b^{i'})$, and write $\mathbf{w}\mathbf{w}'$ for the concatenation of lists $\mathbf{w}$ and $\mathbf{w}'$. The *path* of $\mathbf{b}$ from $\bar{o} \in \mathcal{P}$ is $y_u(\bar{o}, \mathbf{b}) = \bar{o} + \text{pt}_{\min(u, L)}(\mathbf{b})$, $u \in \mathbb{N}$. The *reward* of a hold $b = (\mathbf{c}_1, \dots, \mathbf{c}_{|b|})$ at a point x is

$$\text{rw}(x, b) = \mathbf{1}[x + \text{pt}(b) \text{ white}] + \kappa \#\{i < |b| : \mathbf{c}_i = 3, x + (i, \mathbf{c}_1 + \dots + \mathbf{c}_i) \text{ white}\},$$

and the *accumulated reward* is $\text{Rw}(\bar{o}, \mathbf{b}; u) = \sum_{i=1}^{\min(u, L)} \text{rw}(y_{i-1}(\bar{o}, \mathbf{b}), b^i)$.

In the notation of Definitions 11.3 and 11.4, $w(x + \text{pt}(b))l(x; b) = e^{-z \text{rw}(x, b)}$.

Lemma 12.2 (Elementary properties). *Let $\bar{o} \in \mathcal{P}$, $\mathbf{w} \in \mathfrak{B}^{L_1}$, $\mathbf{w}' \in \mathfrak{B}^{L_2}$, and $\mathbf{b} \in \mathfrak{B}^L$.*

- (1) For $\Psi \geq 0$ on $\mathfrak{B}^{L_1} \times \mathfrak{B}^{L_2}$,

$$\sum_{\mathbf{b} \in \mathfrak{B}^{L_1+L_2}} q(\mathbf{b}) \Psi(\mathbf{b}_{[1, L_1]}, \mathbf{b}_{[L_1+1, L_1+L_2]}) = \sum_{\mathbf{w}, \mathbf{w}'} q(\mathbf{w}) q(\mathbf{w}') \Psi(\mathbf{w}, \mathbf{w}').$$

- (2) $y_u(\bar{o}, \mathbf{w}\mathbf{w}') = y_u(\bar{o}, \mathbf{w})$ for $u \leq L_1$, $y_{L_1+u}(\bar{o}, \mathbf{w}\mathbf{w}') = y_u(y_{L_1}(\bar{o}, \mathbf{w}), \mathbf{w}')$ and $\text{Rw}(\bar{o}, \mathbf{w}\mathbf{w}'; L_1 + u) = \text{Rw}(\bar{o}, \mathbf{w}; L_1) + \text{Rw}(y_{L_1}(\bar{o}, \mathbf{w}), \mathbf{w}'; u)$ for $u \in \mathbb{N}$.
- (3) For $u \leq u'$, with $y_u = y_u(\bar{o}, \mathbf{b})$ and $k = \min(u', L) - \min(u, L)$, we have $j(y_{u'}) - j(y_u) \geq k$ and $l(y_{u'}) - l(y_u) \geq 4k$.

- (4) For $L' \in \mathbb{N}$, the masses of the lists in a prefix-disjoint set of lists of length at most L' sum to at most 1.
- (5) For $u \leq L$, $\text{Rw}(\bar{o}, \mathbf{b}; u) \geq \#\{1 \leq i \leq u : y_i(\bar{o}, \mathbf{b}) \text{ white}\}$, and $u \mapsto \text{Rw}(\bar{o}, \mathbf{b}; u)$ is nondecreasing.

Proof. (1) Concatenation is a bijection $\mathfrak{B}^{L_1} \times \mathfrak{B}^{L_2} \rightarrow \mathfrak{B}^{L_1+L_2}$ and the mass is multiplicative. (2) is immediate from $\text{pt}_{L_1+u}(\mathbf{w}\mathbf{w}') = \text{pt}_{L_1}(\mathbf{w}) + \text{pt}_u(\mathbf{w}')$. (3) Every $\text{pt}(b) = (j, l)$ has $j \geq 1$ and $l \geq 4$ (Lemma 10.4(3)). (4) Extend every list $\mathbf{w}$ of the set by all lists of length $L' - |\mathbf{w}|$. By (1) the extensions of $\mathbf{w}$ carry mass $\mathbf{q}(\mathbf{w})$, and the extensions of distinct lists of the set are disjoint subsets of $\mathfrak{B}^{L'}$, which has total mass 1 (Lemma 10.4(5)). (5) The first term of $\text{rw}(y_{i-1}, b^i)$ is the indicator that y_i is white, and all rewards are nonnegative. $\square$

12.2. Gaps and the trace times. The trace times mark the successive entries of a path into triangles; each new entry lies above the top edge of the triangle entered before, so between two entries the path makes a full passage over that edge (Lemma 12.5). A black point p lies in $\Delta(p)$ and $l_{\Delta(p)} = l_*(p)$ (Lemma 9.12 and Definition 9.9), so its gap (Definition 9.8) is the distance from p up to the top edge of $\Delta(p)$. If (j, l) and $(j, l+1)$ are black, they have the same column top, so

$$\text{gap}((j, l)) = \text{gap}((j, l+1)) + 1. \quad (12.1)$$

Call the path $y_u = y_u(\bar{o}, \mathbf{b})$ of $\mathbf{b} \in \mathfrak{B}^L$ *complete* if $j(\bar{o}) + L > J$, so that it passes the horizon and its trace times below are finite. If y_u is black on a complete path, then $u < L$ and $j(\bar{o}) + u \leq J$; otherwise $j(y_u) \geq j(\bar{o}) + \min(u, L) > J$ by Lemma 12.2(3). A time u is a *hit* if y_u is black.

Definition 12.3 (Trace times). The *trace times* $u_1 < u_2 < \dots$ of a complete path are defined as follows: u_1 is the least hit, if there is one; and once u_i is defined, u_{i+1} is the least hit $u > u_i$ with $l(y_u) > l_*(y_{u_i})$, if there is one. The sequence ends at the first index where no such time exists. Since hits are less than L , it is finite. The points y_{u_i} for $i \geq 2$ are the *entries*; their sequence is the *trace*.

For a point $p \in \mathcal{P}$ and a list $\mathbf{w}$ write $p_{\mathbf{w}} = y_{|\mathbf{w}|}(p, \mathbf{w})$, and $p_b = p + \text{pt}(b)$ for a hold b . The first two trace times are read off from prefixes of the list, as follows. For $\bar{o} \in \mathcal{P}$ let $\mathcal{U}(\bar{o})$, the *bridge lists*, be the set of lists $\mathbf{w}$ such that $\bar{o}_{\mathbf{w}}$ is black and $y_u(\bar{o}, \mathbf{w})$ is not black for $u < |\mathbf{w}|$. For a black v let $\mathcal{E}(v)$, the *entry lists*, be the set of nonempty lists $\mathbf{w}$ such that $v_{\mathbf{w}}$ is black with $l(v_{\mathbf{w}}) > l_*(v)$, and no $y_u(v, \mathbf{w})$ with $1 \leq u < |\mathbf{w}|$ has these two properties. The sets $\mathcal{U}(\bar{o})$ and $\mathcal{E}(v)$ are prefix-disjoint. If $\mathbf{w} \in \mathcal{U}(\bar{o})$, then $j(\bar{o}) + |\mathbf{w}| \leq j(\bar{o}_{\mathbf{w}}) \leq J$, and similarly for $\mathcal{E}(v)$. The path of a list agrees with the path of each of its prefixes up to the prefix's length. So, for the complete path of $\mathbf{b} \in \mathfrak{B}^L$ from $\bar{o}$, u_1 exists and equals $\tilde{u}$ if and only if $\tilde{u} \leq L$ and $\mathbf{b}_{[1, \tilde{u}]} \in \mathcal{U}(\bar{o})$; and if $\bar{o} = v$ is black, then $u_1 = 0$, and u_2 exists and equals $\tilde{u}$ if and only if $1 \leq \tilde{u} \leq L$ and $\mathbf{b}_{[1, \tilde{u}]} \in \mathcal{E}(v)$.

Lemma 12.4 (Restart at a trace time). *Let $y_u = y_u(\bar{o}, \mathbf{b})$ be the complete path of $\mathbf{b} \in \mathfrak{B}^L$ from $\bar{o}$. Suppose $u_i = \tilde{u}$ exists, and let $y'_u = y_u(y_{\tilde{u}}, \mathbf{b}_{[\tilde{u}+1, L]})$ be the path of $\mathbf{b}_{[\tilde{u}+1, L]}$ from $y_{\tilde{u}}$. Then $j(y_{\tilde{u}}) + (L - \tilde{u}) > J$, and for every $k \geq 1$, u_{i-1+k} exists if and only if u'_k (the trace times of the path y'_u) does, and then $u_{i-1+k} = \tilde{u} + u'_k$.*

Proof. By Lemma 12.2(3),(2), $j(y_{\tilde{u}}) \geq j(\bar{o}) + \tilde{u}$ and $y'_u = y_{\tilde{u}+u}$. Since $y'_0 = y_{\tilde{u}}$ is black, $u'_1 = 0$, matching $u_i = \tilde{u}$, and both sequences are continued by the same rule applied to the same points. $\square$

12.3. One passage. Between two trace times the path climbs from a black point v above its column top. This subsection shows that such a passage earns a reward with probability bounded below, because it leaves through the white exit points of Lemma 9.14. For $\omega \in \mathbb{N}$, let Π_ω be the set of nonempty lists $\mathbf{f}$ with $l(\text{pt}_i(\mathbf{f})) \leq \omega$ for $i < |\mathbf{f}|$ and $l(\text{pt}_{|\mathbf{f}|}(\mathbf{f})) > \omega$, that is, the lists whose hold

points form a first-passage list over height ω (Definition 10.5). Grouping lists by their hold points and using (10.1) gives, for $\Psi \geq 0$,

$$\sum_{\mathbf{f} \in \Pi_\omega} \mathbf{q}(\mathbf{f}) \Psi(\text{pt}_{|\mathbf{f}|}(\mathbf{f})) = \sum_{x \in \mathcal{P}} F_\omega(x) \Psi(x). \quad (12.2)$$

In particular $\sum_{\mathbf{f} \in \Pi_\omega} \mathbf{q}(\mathbf{f}) = 1$. Let $\mathbf{c}$ be the raw word of $\mathbf{f}$. Its raw letters are at least 2, its raw letters in $\{4, 5\}$ are exactly the closing letters of the holds, $\mathbf{q}(\mathbf{f}) = \prod_i \mathbf{q}(\mathbf{c}_i)$, and the closing letter of the i th hold is $\mathbf{c}_r$ with $r = j(\text{pt}_i(\mathbf{f}))$.

For $\omega \in \mathbb{N}$, with f_ω as in Lemma 9.14 and the exit mass $\mathbf{q}_{45}$ of Definition 10.16, put

$$\mathbf{q}_3(\omega) = \mathbf{q}(3) \sum_{o=1}^3 \sum_{r=1}^{f_\omega} R(r-1, \omega+o-3), \quad (12.3)$$

the mass of the passages over height ω that first exceed it at a raw letter 3 at a white exit point (Lemma 12.5), and define the *passage surplus*

$$\delta_{\text{tr}}(\omega) = E_8(\gamma) \mathbf{q}_{45}(\omega) + E_8(\kappa\gamma) \mathbf{q}_3(\omega) - d_{\text{tr}},$$

the excess of the one-passage gain at gap ω over its value d_{tr} at gap 0.

Lemma 12.5 (Reward over one passage). *Let v be black and $\omega = \text{gap}(v)$. Then*

$$\sum_{\mathbf{f} \in \Pi_\omega} \mathbf{q}(\mathbf{f}) e^{-\gamma \text{Rw}(v, \mathbf{f}; |\mathbf{f}|)} \leq 1 - d_{\text{tr}} - \delta_{\text{tr}}(\omega).$$

Proof. Let Π_ω^{45} be the set of $\mathbf{f} \in \Pi_\omega$ with $\text{pt}_{|\mathbf{f}|}(\mathbf{f}) = (r, \omega+o)$, $1 \leq r \leq f_\omega$, $1 \leq o \leq 4$. By (12.2) and Lemma 10.17, its mass is $\mathbf{q}_{45}(\omega)$. Let Π_ω^3 be the set of $\mathbf{f} \in \Pi_\omega$ whose raw word $\mathbf{c}$ first exceeds ω at a position $r \leq f_\omega$ with $\mathbf{c}_r = 3$, that is, $\mathbf{c}_1 + \dots + \mathbf{c}_{r-1} \leq \omega < \mathbf{c}_1 + \dots + \mathbf{c}_r$. The two sets are disjoint. Indeed, for $\mathbf{f} \in \Pi_\omega^{45}$, the raw word has length r and ends with a closing letter $k \geq 4 \geq o$, so $\mathbf{c}_1 + \dots + \mathbf{c}_{r-1} = \omega + o - k \leq \omega$ and the raw word first exceeds ω at the closing letter.

The mass of Π_ω^3 is $\mathbf{q}_3(\omega)$. Consider the choices of $(r, o, \mathbf{c}^-, \mathbf{c}^+, k)$ with $1 \leq r \leq f_\omega$, $o \in \{1, 2, 3\}$, a word $\mathbf{c}^-$ of $r-1$ raw letters ≥ 2 with sum $\omega + o - 3$, a word $\mathbf{c}^+$ of raw letters ≥ 2 outside $\{4, 5\}$, and $k \in \{4, 5\}$. Splitting the raw word $\mathbf{c}^-$, then 3, then $\mathbf{c}^+$, then k after each closing letter gives a list $\mathbf{f}$. Every hold of $\mathbf{f}$ except the last ends inside $\mathbf{c}^-$, at height at most $\omega + o - 3 \leq \omega$, and the last hold ends at height at least $\omega + o > \omega$. So $\mathbf{f} \in \Pi_\omega$, and its raw word first exceeds ω at position r , with raw letter 3. Hence $\mathbf{f} \in \Pi_\omega^3$. Conversely, let $\mathbf{f} \in \Pi_\omega^3$, and let r be the position at which its raw word first exceeds ω . The raw letter 3 at r is not closing, so it lies in some hold of $\mathbf{f}$ before that hold's closing letter. Earlier holds end before r and at height at most ω , and this hold ends after r , at height greater than ω . By the definition of Π_ω , it is the last hold of $\mathbf{f}$. Thus $\mathbf{c}$ is $\mathbf{c}^-$, then 3, then $\mathbf{c}^+$, then k , as above, the two maps are inverse, and summing the product masses over the length N of $\mathbf{c}^+$ gives

$$\mathbf{q}(3) \sum_{o=1}^3 \sum_{r=1}^{f_\omega} R(r-1, \omega+o-3) \sum_{N \geq 0} \left(\frac{11}{16}\right)^N \frac{5}{16} = \mathbf{q}_3(\omega),$$

since $\sum_{c \notin \{4, 5\}} \mathbf{q}(c) = 1 - \frac{5}{16}$. The argument also shows that, for $\mathbf{f} \in \Pi_\omega^3$, the raw letter $\mathbf{c}_r = 3$ is the i th raw letter of the last hold of $\mathbf{f}$ for some i , and

$$y_{|\mathbf{f}|-1}(v, \mathbf{f}) + (i, \mathbf{c}_{r-i+1} + \dots + \mathbf{c}_r) = v + (r, \omega + o).$$

Rewards. Since $v \in \Delta(v)$ and $l_{\Delta(v)} = l(v) + \omega$ (Lemma 9.12 and Definition 9.9), Lemma 9.14 shows that $v + (r, \omega + o)$ is white for $1 \leq r \leq f_\omega$, $1 \leq o \leq 4$. If $\mathbf{f} \in \Pi_\omega^{45}$, this is its last point, so the last hold has reward at least 1. If $\mathbf{f} \in \Pi_\omega^3$, it is the point of the raw letter $\mathbf{c}_r = 3$ displayed above, so the last hold has reward at least κ . Since rewards are nonnegative,

$$e^{-\gamma \text{Rw}(v, \mathbf{f}; |\mathbf{f}|)} \leq 1 - (1 - e^{-\gamma}) \mathbf{1}[\mathbf{f} \in \Pi_\omega^{45}] - (1 - e^{-\kappa\gamma}) \mathbf{1}[\mathbf{f} \in \Pi_\omega^3].$$

Summing against $\mathbf{q}$ and using $E_8(y) < 1 - e^{-y}$ for $y > 0$ (Definition A.12) gives the bound $1 - E_8(\gamma)\mathbf{q}_{45}(\omega) - E_8(\kappa\gamma)\mathbf{q}_3(\omega) = 1 - d_{\text{tr}} - \delta_{\text{tr}}(\omega)$. $\square$

Lemma 12.6 (Entry lists split). *Let v be black and $\omega = \text{gap}(v)$. The map $(\mathbf{f}, \mathbf{w}) \mapsto \mathbf{fw}$ is a bijection from the pairs with $\mathbf{f} \in \Pi_\omega$ and $\mathbf{w} \in \mathcal{U}(v_{\mathbf{f}})$ onto $\mathcal{E}(v)$.*

Proof. For a list $\mathbf{b}$, $l(y_u(v, \mathbf{b})) > l_*(v)$ holds if and only if $l(\text{pt}_u(\mathbf{b})) > \omega$. Given $\mathbf{b} \in \mathcal{E}(v)$, let k be the least index with $l(\text{pt}_k(\mathbf{b})) > \omega$; it exists because the last point lies above $l_*(v)$. Then $\mathbf{f} = \mathbf{b}_{[1,k]} \in \Pi_\omega$, and $\mathbf{f}$ is the only prefix of $\mathbf{b}$ in Π_ω . By Lemma 12.2(3), every $y_u(v, \mathbf{b})$ with $u \geq k$ lies above $l_*(v)$. So the defining conditions of $\mathcal{E}(v)$ say exactly that $\mathbf{b}_{[k+1, |\mathbf{b}|]} \in \mathcal{U}(y_k(v, \mathbf{b}))$. The same two facts show that every $\mathbf{fw}$ as in the statement lies in $\mathcal{E}(v)$. $\square$

12.4. The entry kernel. By Lemma 12.11, the $\mathbf{q}$ -average of $e^{-\gamma \text{Rw}}$ at the $(k+1)$ st trace time from a black point v is $\mathbf{K}^k \mathbf{1}(v)$, for the entry kernel $\mathbf{K}$ below and the constant function $\mathbf{1}$, so it suffices to bound the powers of $\mathbf{K}$. The bridge loss measures how far a bridge list followed by one passage falls below $1 - d_{\text{tr}}$; it produces the improvement of two passages over one. For a point $p \in \mathcal{P}$ and a list $\mathbf{w}$ write $\mathbf{q}^p(\mathbf{w}) = \mathbf{q}(\mathbf{w})e^{-\gamma \text{Rw}(p, \mathbf{w}; |\mathbf{w}|)}$. The *entry kernel* acts on nonnegative functions on black points by

$$\mathbf{K}\Psi(v) = \sum_{\mathbf{w} \in \mathcal{E}(v)} \mathbf{q}^v(\mathbf{w}) \Psi(v_{\mathbf{w}}).$$

It is positive: $0 \leq \Psi \leq \Psi'$ implies $\mathbf{K}\Psi \leq \mathbf{K}\Psi'$. For $p \in \mathcal{P}$ define the *bridge mass*, *bridge gain* and *bridge loss*

$$\begin{aligned} \text{Br}(p) &= \sum_{\mathbf{w} \in \mathcal{U}(p)} \mathbf{q}^p(\mathbf{w}), & \text{Gain}(p) &= \sum_{\mathbf{w} \in \mathcal{U}(p)} \mathbf{q}^p(\mathbf{w}) \delta_{\text{tr}}(\text{gap}(p_{\mathbf{w}})), \\ \text{Loss}(p) &= (1 - d_{\text{tr}})(1 - \text{Br}(p)) + \text{Gain}(p). \end{aligned}$$

Lemma 12.7 (One kernel step and the bridge loss). (1) $\text{Br}(p) \leq 1$ for every p , and $\mathbf{K}\mathbf{1}(v) \leq 1 - d_{\text{tr}} - \delta_{\text{tr}}(\text{gap}(v)) \leq 1 - d_{\text{tr}}$ for every black v .
(2) $0 \leq \text{Loss}(p) \leq 1 - d_{\text{tr}}$ for every p ; $\text{Loss}(p) = \delta_{\text{tr}}(\text{gap}(p))$ if p is black, and $\text{Loss}(p) = 1 - d_{\text{tr}}$ if $j(p) > J$.
(3) For every black v , with $\omega = \text{gap}(v)$,

$$\mathbf{K}^2 \mathbf{1}(v) \leq (1 - d_{\text{tr}})^2 - (1 - d_{\text{tr}})\delta_{\text{tr}}(\omega) - \sum_{\mathbf{f} \in \Pi_\omega} \mathbf{q}^v(\mathbf{f}) \text{Loss}(v_{\mathbf{f}}).$$

Proof. (1) Rewards are nonnegative, so $\mathbf{q}^p(\mathbf{w}) \leq \mathbf{q}(\mathbf{w})$. Lists in $\mathcal{U}(p)$ have length less than J and form a prefix-disjoint family, so $\text{Br}(p) \leq 1$ by Lemma 12.2(4). By Lemma 12.6 with $\omega = \text{gap}(v)$ and Lemma 12.2(1),(2), $\mathbf{q}^v(\mathbf{fw}) = \mathbf{q}^v(\mathbf{f})\mathbf{q}^{v_{\mathbf{f}}}(\mathbf{w})$, and therefore

$$\mathbf{K}\mathbf{1}(v) = \sum_{\mathbf{f} \in \Pi_\omega} \mathbf{q}^v(\mathbf{f}) \text{Br}(v_{\mathbf{f}}) \leq \sum_{\mathbf{f} \in \Pi_\omega} \mathbf{q}^v(\mathbf{f}).$$

Lemma 12.5 and Lemma A.13(1) conclude.

(2) By (1), $0 \leq \delta_{\text{tr}}(\text{gap}(v)) \leq 1 - d_{\text{tr}}$ for every black v , so $0 \leq \text{Gain}(p) \leq (1 - d_{\text{tr}}) \text{Br}(p)$, which gives the two bounds. If p is black, then $\mathcal{U}(p)$ consists of the empty list only, so $\text{Br}(p) = 1$ and $\text{Gain}(p) = \delta_{\text{tr}}(\text{gap}(p))$. If $j(p) > J$, no point of a path from p is black, so $\mathcal{U}(p) = \emptyset$.

(3) Apply (1) at each $v_{\mathbf{w}}$ and Lemma 12.6:

$$\mathbf{K}^2 \mathbf{1}(v) \leq \sum_{\mathbf{f} \in \Pi_\omega} \mathbf{q}^v(\mathbf{f}) \sum_{\mathbf{w} \in \mathcal{U}(v_{\mathbf{f}})} \mathbf{q}^{v_{\mathbf{f}}}(\mathbf{w}) (1 - d_{\text{tr}} - \delta_{\text{tr}}(\text{gap}((v_{\mathbf{f}})_{\mathbf{w}}))).$$

The inner sum is $(1 - d_{\text{tr}}) \text{Br}(v_{\mathbf{f}}) - \text{Gain}(v_{\mathbf{f}}) = (1 - d_{\text{tr}}) - \text{Loss}(v_{\mathbf{f}})$, a difference of finite numbers. Since $1 - d_{\text{tr}} > 0$, Lemma 12.5 bounds $(1 - d_{\text{tr}}) \sum_{\mathbf{f}} \mathbf{q}^v(\mathbf{f})$ by $(1 - d_{\text{tr}})(1 - d_{\text{tr}} - \delta_{\text{tr}}(\omega))$. $\square$

12.5. Two passages. Squaring the one-passage bound would give only $(1 - d_{\text{tr}})^2$; Proposition 12.10 improves it to $r_2 = (1 - d_{\text{tr}})^2 - \Omega$ (Definition A.12). The correction Ω comes from the loss after the first of the two passages. If that passage ends at a black point, the loss is the surplus of the next passage. If it ends at a white point, the path earns a reward before its next trace time. Put

$$d_{\text{pair}} = \frac{3}{16}\delta_{\text{tr}}(6) + \frac{1}{8}\delta_{\text{tr}}(5), \quad d_{\text{wh}} = \frac{10000}{3299}d_{\text{pair}} :$$

d_{pair} bounds the surplus after the two closing letters 4, 5 of a hold (Lemma A.13(2)), and d_{wh} the resulting loss at a white point (Lemma 12.8).

In this subsection every hold is written $(\mathbf{d}, k) = (\mathbf{c}_1, \dots, \mathbf{c}_N, k)$, where $\mathbf{d} = (\mathbf{c}_1, \dots, \mathbf{c}_N)$, $N \in \mathbb{N}$, is the word of its raw letters before the closing letter $k \in \{4, 5\}$ (these are at least 2 and outside $\{4, 5\}$), and $\mathbf{q}((\mathbf{d}, k)) = \mathbf{q}(\mathbf{d})\mathbf{q}(k)$. For such a word $\mathbf{d}$ let $\#_3(\mathbf{d})$ be the number of its raw letters equal to 3; for $p \in \mathcal{P}$ let $\#_3^{\text{w}}(p, \mathbf{d})$ be the number of i with $\mathbf{c}_i = 3$ and $p + (i, \mathbf{c}_1 + \dots + \mathbf{c}_i)$ white, and put $\text{disc}(p, \mathbf{d}) = e^{-\kappa\gamma\#_3^{\text{w}}(p, \mathbf{d})}$. Since $1 - \kappa\gamma = 1163/1250$, summing over the words of length N factorizes:

$$\sum_{|\mathbf{d}|=N} \mathbf{q}(\mathbf{d})(1 - \kappa\gamma)^{\#_3(\mathbf{d})} = \left(1 - \frac{1}{4} - \frac{5}{16} + \frac{1}{4} \cdot \frac{1163}{1250}\right)^N = \left(\frac{6701}{10000}\right)^N. \quad (12.4)$$

Lemma 12.8 (A white point loses at least d_{wh}). *For every white p , $\text{Loss}(p) \geq d_{\text{wh}}$.*

Proof. Since p is not black, every $\mathbf{w} \in \mathcal{U}(p)$ is nonempty, and $\mathbf{w} \mapsto (b, \mathbf{w}')$ with $\mathbf{w} = (b)\mathbf{w}'$ is a bijection onto the pairs with $b \in \mathfrak{B}$ and $\mathbf{w}' \in \mathcal{U}(p_b)$. Since $\mathbf{q}^p((b)\mathbf{w}') = \mathbf{q}(b)e^{-\gamma\text{rw}(p, b)}\mathbf{q}^{p_b}(\mathbf{w}')$ and $\sum_b \mathbf{q}(b) = 1$,

$$\text{Loss}(p) = \sum_{b \in \mathfrak{B}} \mathbf{q}(b)\mathbf{L}(p, b), \quad \mathbf{L}(p, b) = (1 - d_{\text{tr}})(1 - e^{-\gamma\text{rw}(p, b)}) + e^{-\gamma\text{rw}(p, b)}\text{Loss}(p_b) \geq 0. \quad (12.5)$$

Fix $\mathbf{d}$ as above. Then $\text{rw}(p, (\mathbf{d}, k)) = \mathbf{1}[p(\mathbf{d}, k) \text{ white}] + \kappa\#_3^{\text{w}}(p, \mathbf{d})$. We claim

$$\frac{3}{16}\mathbf{L}(p, (\mathbf{d}, 4)) + \frac{1}{8}\mathbf{L}(p, (\mathbf{d}, 5)) \geq \text{disc}(p, \mathbf{d})d_{\text{pair}}. \quad (12.6)$$

The points $p(\mathbf{d}, 4)$ and $p(\mathbf{d}, 5) = p(\mathbf{d}, 4) + (0, 1)$ lie in the same column. If $j(p(\mathbf{d}, 4)) > J$, then $\mathbf{L}(p, (\mathbf{d}, k)) = 1 - d_{\text{tr}}$ by Lemma 12.7(2), and $\frac{5}{16}(1 - d_{\text{tr}}) > d_{\text{pair}}$. If both points are black, then $\mathbf{L}(p, (\mathbf{d}, k)) \geq \text{disc}(p, \mathbf{d})\delta_{\text{tr}}(\text{gap}(p(\mathbf{d}, k)))$. By (12.1) the two gaps are ω and $\omega - 1$ for some $\omega \geq 1$, and Lemma A.13(2) applies. Otherwise some $p(\mathbf{d}, k')$ is white. Then $e^{-\gamma\text{rw}(p, (\mathbf{d}, k'))} = \text{disc}(p, \mathbf{d})e^{-\gamma}$, so $\mathbf{L}(p, (\mathbf{d}, k')) \geq (1 - d_{\text{tr}})(1 - \text{disc}(p, \mathbf{d})e^{-\gamma}) \geq (1 - d_{\text{tr}})\text{disc}(p, \mathbf{d})(1 - e^{-\gamma}) > (1 - d_{\text{tr}})\text{disc}(p, \mathbf{d})E_8(\gamma)$. Since $\mathbf{q}(k') \geq 1/8$, Lemma A.13(3) gives (12.6).

Finally $\text{disc}(p, \mathbf{d}) \geq e^{-\kappa\gamma\#_3^{\text{w}}(\mathbf{d})} \geq (1 - \kappa\gamma)^{\#_3(\mathbf{d})}$. Multiplying (12.6) by $\mathbf{q}(\mathbf{d})$, summing over $\mathbf{d}$ and using (12.4) gives $\text{Loss}(p) \geq d_{\text{pair}} \sum_{N \geq 0} (6701/10000)^N = d_{\text{wh}}$. $\square$

Lemma 12.9 (A black point of gap zero). *If v is black with $\text{gap}(v) = 0$, then*

$$\sum_{\mathbf{f} \in \Pi_0} \mathbf{q}^v(\mathbf{f}) \text{Loss}(v_{\mathbf{f}}) \geq e^{-\gamma}d_{\text{wh}}\left(\frac{3}{16} + \frac{6701}{26392}\right).$$

Proof. Holds climb at least 4 (Lemma 10.4(3)), so Π_0 consists of the one-hold lists (b) , $b \in \mathfrak{B}$. Their terms are $\mathbf{q}(b)e^{-\gamma\text{rw}(v, b)}\text{Loss}(v_b) \geq 0$. For the hold $b = (4)$ the point $v + (1, 4)$ is white by Lemma 9.14 ($\omega = 0$, $r = 1$, $o = 4$). So its reward is 1, and by Lemma 12.8 its term is at least $\frac{3}{16}e^{-\gamma}d_{\text{wh}}$. For a nonempty $\mathbf{d}$, the two holds $(\mathbf{d}, 4)$, $(\mathbf{d}, 5)$ contribute at least $\mathbf{q}(\mathbf{d})\text{disc}(v, \mathbf{d}) \cdot \frac{5}{16}(1 - d_{\text{tr}})$ if $j(v(\mathbf{d}, 4)) > J$, at least $\mathbf{q}(\mathbf{d})\text{disc}(v, \mathbf{d})d_{\text{pair}}$ if $v(\mathbf{d}, 4), v(\mathbf{d}, 5)$ are black, and otherwise at least $\frac{1}{8}\mathbf{q}(\mathbf{d})\text{disc}(v, \mathbf{d})e^{-\gamma}d_{\text{wh}}$, by Lemma 12.7(2), (12.1) and Lemma A.13(2), and Lemma 12.8 respectively. Since $\frac{5}{16}(1 - d_{\text{tr}}) > d_{\text{pair}} > d_{\text{wh}}/8$, each case gives at least $\frac{1}{8}\mathbf{q}(\mathbf{d})(1 - \kappa\gamma)^{\#_3(\mathbf{d})}e^{-\gamma}d_{\text{wh}}$. Summing over $N \geq 1$ by (12.4) gives the factor $\frac{1}{8} \cdot \frac{6701}{3299} = \frac{6701}{26392}$. $\square$

Proposition 12.10 (The entry kernel). *For every black point v ,*

$$\mathbf{K}\mathbf{1}(v) \leq 1 - d_{\text{tr}}, \quad \mathbf{K}^2\mathbf{1}(v) \leq r_2,$$

and hence $\mathbf{K}^i\mathbf{1}(v) \leq r_2^{\lfloor i/2 \rfloor} (1 - d_{\text{tr}})^{i-2\lfloor i/2 \rfloor}$ for every $i \in \mathbb{N}$.

Proof. The first bound is Lemma 12.7(1). Let $\omega = \text{gap}(v)$. Losses are nonnegative, so if $\omega \geq 1$, Lemma 12.7(3) and Lemma A.13(1),(3) give $\mathbf{K}^2\mathbf{1}(v) \leq (1 - d_{\text{tr}})^2 - (1 - d_{\text{tr}})\delta_{\text{tr}}(6) < r_2$. If $\omega = 0$, then $\delta_{\text{tr}}(0) = 0$, and Lemma 12.9 and $e^{-\gamma} > 1 - \gamma > 0$ give

$$\mathbf{K}^2\mathbf{1}(v) \leq (1 - d_{\text{tr}})^2 - e^{-\gamma}d_{\text{wh}}\left(\frac{3}{16} + \frac{6701}{26392}\right) < (1 - d_{\text{tr}})^2 - \Omega = r_2$$

by Lemma A.13(3) and the definition of r_2 . For the powers, induct on i using positivity: if $\mathbf{K}^i\mathbf{1}$ is at most the constant of the statement pointwise, then $\mathbf{K}^{i+2}\mathbf{1} = \mathbf{K}^2(\mathbf{K}^i\mathbf{1})$ is at most r_2 times that constant, which is the constant for $i + 2$. $\square$

12.6. Accumulated rewards at the trace times. For $\mathbf{b} \in \mathfrak{B}^L$ whose path from $\bar{o}$ is complete, and $k \in \mathbb{N}$, put $\mathcal{Z}_{k+1}(\bar{o}, \mathbf{b}) = e^{-\gamma \text{Rw}(\bar{o}, \mathbf{b}; u_{k+1})}$ if u_{k+1} exists, and $\mathcal{Z}_{k+1}(\bar{o}, \mathbf{b}) = 0$ otherwise.

Lemma 12.11 (Explicit form of Tao's Lemma 7.9). *Let $\bar{o} \in \mathcal{P}$ and $L \in \mathbb{N}$ with $j(\bar{o}) + L > J$, so that the paths from $\bar{o}$ of all $\mathbf{b} \in \mathfrak{B}^L$ are complete. Then*

- (1) *if $\bar{o} = v$ is black, then $\sum_{\mathbf{b} \in \mathfrak{B}^L} \mathbf{q}(\mathbf{b}) \mathcal{Z}_{k+1}(v, \mathbf{b}) = \mathbf{K}^k\mathbf{1}(v)$ for every $k \in \mathbb{N}$;*
- (2) *for every $k \in \mathbb{N}$, $\sum_{\mathbf{b} \in \mathfrak{B}^L} \mathbf{q}(\mathbf{b}) \mathcal{Z}_{k+1}(\bar{o}, \mathbf{b}) \leq r_2^{\lfloor k/2 \rfloor} (1 - d_{\text{tr}})^{k-2\lfloor k/2 \rfloor}$.*

Proof. (1) We induct on k , for all black v and all L with $j(v) + L > J$ at once. For $k = 0$, $u_1 = 0$ and $\mathcal{Z}_1(v, \mathbf{b}) = 1$, and the total mass is 1. For the step, by the description of u_2 before Lemma 12.4, that lemma with $i = 2$, and Lemma 12.2(2),

$$\mathcal{Z}_{k+2}(v, \mathbf{b}) = \sum_{u=1}^L \mathbf{1}[\mathbf{b}_{[1,u]} \in \mathcal{E}(v)] e^{-\gamma \text{Rw}(v, \mathbf{b}_{[1,u]}; u)} \mathcal{Z}_{k+1}(y_u(v, \mathbf{b}), \mathbf{b}_{[u+1,L]}),$$

with at most one nonzero term. Split $\mathbf{b}$ at u by Lemma 12.2(1). For $\mathbf{w} \in \mathcal{E}(v)$, the point $v_{\mathbf{w}}$ is black and $j(v_{\mathbf{w}}) + L - |\mathbf{w}| > J$, so the induction hypothesis evaluates the sum over the remainder as $\mathbf{K}^k\mathbf{1}(v_{\mathbf{w}})$. Every $\mathbf{w} \in \mathcal{E}(v)$ has $|\mathbf{w}| \leq J - j(v) < L$. Hence the total is $\sum_{\mathbf{w} \in \mathcal{E}(v)} \mathbf{q}^v(\mathbf{w}) \mathbf{K}^k\mathbf{1}(v_{\mathbf{w}}) = \mathbf{K}^{k+1}\mathbf{1}(v)$.

(2) In the same way, the description of u_1 and Lemma 12.4 with $i = 1$ give

$$\sum_{\mathbf{b}} \mathbf{q}(\mathbf{b}) \mathcal{Z}_{k+1}(\bar{o}, \mathbf{b}) = \sum_{\mathbf{w} \in \mathcal{U}(\bar{o})} \mathbf{q}^{\bar{o}}(\mathbf{w}) \sum_{\mathbf{w}' \in \mathfrak{B}^{L-|\mathbf{w}|}} \mathbf{q}(\mathbf{w}') \mathcal{Z}_{k+1}(\bar{o}_{\mathbf{w}}, \mathbf{w}').$$

By (1) and Proposition 12.10, the inner sum is at most the claimed bound. Finally $\sum_{\mathbf{w} \in \mathcal{U}(\bar{o})} \mathbf{q}^{\bar{o}}(\mathbf{w}) = \text{Br}(\bar{o}) \leq 1$. $\square$

13. BOUNDARY POINTS FAR BELOW THE TOP EDGE

This section proves Case 3 of the descent step (Section 11): the boundary point p for m is black and lies in a triangle $\Delta \in \mathfrak{T}$ at gap $\text{gap}(p) = l_{\Delta} - l(p)$ greater than $m/(\log m)^2$ below its top edge (Lemma 13.17). Following Tao [10, Case 3 of Proposition 7.8 and Lemma 7.10], we restart the hold walk after its first passage above the top edge; outside rare events, a low reward then forces many entries into triangles (Lemma 13.4), which Lemma 12.11 makes unlikely, and the main rare event, a visit to a large triangle, is bounded by the packing estimate (Lemma 13.13). The constants are those of Definitions A.7, A.2, A.3, A.5 and A.12, and Q_{pk} of Lemma A.10(2).

13.1. The fresh law and the Case 3 events. In Case 3 the path is restarted at a boundary point $p \in \mathcal{P}$ after a first passage over height $\omega \in \mathbb{N}$, with the law $\mathbf{F}_\omega$.

Definition 13.1 (Fresh law). The *atoms* are the pairs $a = (x, \mathbf{b})$ with *displacement* $x \in \mathcal{P}$ and $\mathbf{b} = (b^1, \dots, b^J) \in \mathfrak{B}^J$. The *fresh law* is $\mathbb{P}_{p,\omega}(a) = \mathbf{F}_\omega(x) \mathbf{q}(\mathbf{b})$ (independent of p), and for a set $\mathcal{Q}$ of atoms, $\mathbb{P}_{p,\omega}(\mathcal{Q}) = \sum_{a \in \mathcal{Q}} \mathbb{P}_{p,\omega}(a)$. The *fresh path* and its accumulated reward are

$$y_u^p(a) = p + x + \text{pt}_{\min(u,J)}(\mathbf{b}), \quad \text{Rw}_u^p(a) = \text{Rw}(p + x, \mathbf{b}; u).$$

Thus $y_u^p(a) = y_u(p + x, \mathbf{b})$ is the path of $\mathbf{b}$ from $p + x$.

By Lemmas 10.6(1) and 10.4(5), $\mathbb{P}_{p,\omega}$ is a probability on the atoms. Since $j(p + x) \geq 1$, every fresh path is complete.

Lemma 13.2 (Displacement marginal). *For $0 \leq k \leq J$ and $\Psi \geq 0$ on $\mathcal{P} \times \mathcal{P}^k$,*

$$\sum_{a=(x,\mathbf{b})} \mathbb{P}_{p,\omega}(a) \Psi(x, (\text{pt}(b^1), \dots, \text{pt}(b^k))) = \sum_{x \in \mathcal{P}} \mathbf{F}_\omega(x) \sum_{\mathbf{h} \in \mathcal{P}^k} \eta^{\otimes k}(\mathbf{h}) \Psi(x, \mathbf{h}).$$

Proof. Split $\mathbf{b}$ after k holds and sum out the remainder, which has total mass 1 (Lemmas 12.2(1) and 10.4(5)). Then group the first k holds by their points, using (10.1). $\square$

Definition 13.3 (Case 3 events). For $p \in \mathcal{P}$ and $m \in \mathbb{N}$, define the sets of atoms

$$\begin{aligned} \text{Low}^p &= \{a : \text{Rw}_{P-1}^p(a) \leq T_{\text{tr}}\}, \\ \text{Bad}_m^p &= \{a : 400(j(y_P^p(a)) - j(p)) \geq 317m\}, \\ \text{E}^{*,p} &= \{a : y_u^p(a) \in \Delta \text{ and } S_\Delta \geq \beta(u) \text{ for some } 0 \leq u \leq P-1, \Delta \in \mathfrak{T}\}, \\ \text{Clr} &= \left\{a : \sum_{i=u+1}^{u+h_{\text{tr}}(u)} l(\text{pt}(b^i)) \leq \frac{13}{9}\beta(u) \text{ for some } u \text{ with } u + h_{\text{tr}}(u) \leq P-1\right\}. \end{aligned}$$

We call Low^p the *low-reward event*, $\text{E}^{*,p}$ the *oversize event* and Clr the *slow-clearing event*.

Off Bad_m^p the advance $j(y_P^p) - j(p)$ stays below $\frac{317}{400}m$, so the factor $\max(m - \cdot, 1)^{-A^*}$ of the boundary reduction stays comparable to m^{-A^*} ; off $\text{E}^{*,p} \cup \text{Clr}$ the path climbs out of every triangle it enters on schedule.

13.2. The Markov bound. On Low^p outside $\text{Bad}_m^p \cup \text{E}^{*,p} \cup \text{Clr}$ the fresh path has $r_{\text{tr}} + 1$ trace times before time P (Lemma 13.4), so Lemma 12.11 and Markov's inequality bound its mass, the *Markov bound* (Lemma 13.6).

Lemma 13.4 (The trace times come on schedule). *Let $p \in \mathcal{P}$ and $m \in \mathbb{N}$ with $j(p) + m = J$, suppose $P \leq J$, and let $a = (x, \mathbf{b}) \in \text{Low}^p \setminus (\text{Bad}_m^p \cup \text{E}^{*,p} \cup \text{Clr})$. Then the fresh path $y_u^p = y_u^p(a)$ has at least $r_{\text{tr}} + 1 = 581$ trace times, and $u_i \leq \mathbf{p}_{i-1}$ for $1 \leq i \leq r_{\text{tr}} + 1$.*

Proof. Since $a \notin \text{Bad}_m^p$, $j(y_P^p) < j(p) + m = J$, so $j(y_u^p) \leq J$ for $u \leq P$. Every point y_u^p with $u \leq P$ is therefore black or white. Also $\mathbf{p}_i$ increases with i (Lemma A.4(1)), and $P - 1 = \mathbf{p}_{r_{\text{tr}}} \leq J - 1$. By Lemma 12.2(5), for every $u' \leq P - 1$ the number of white points among $y_1^p, \dots, y_{u'}^p$ is at most $\text{Rw}_{P-1}^p(a) \leq T_{\text{tr}}$. In particular, every run of $T_{\text{tr}} + 1$ consecutive times in $[1, P - 1]$ contains a hit.

For $i = 1$, the run $[1, T_{\text{tr}} + 1]$, where $T_{\text{tr}} + 1 = 102 \leq \mathbf{p}_0$, contains a hit, so $u_1 \leq \mathbf{p}_0$. Suppose $u_i = \tilde{u} \leq \mathbf{p}_{i-1}$ with $i \leq r_{\text{tr}}$. Since $\mathbf{g}$ is nondecreasing (Lemma A.4(1)), $\tilde{u} + h_{\text{tr}}(\tilde{u}) + T_{\text{tr}} + 1 = \mathbf{g}(\tilde{u}) \leq \mathbf{g}(\mathbf{p}_{i-1}) = \mathbf{p}_i \leq P - 1$. Let $\Delta = \Delta(y_{\tilde{u}}^p)$. Then $y_{\tilde{u}}^p \in \Delta$ (Lemma 9.12) and $l_\Delta = l_*(y_{\tilde{u}}^p)$. As $a \notin \text{E}^{*,p}$, $S_\Delta < \beta(\tilde{u})$, and membership gives $\text{gap}(y_{\tilde{u}}^p) \log 2 \leq S_\Delta$. Since $\log 2 > 9/13$, we get $\text{gap}(y_{\tilde{u}}^p) < \frac{13}{9}\beta(\tilde{u})$. As $a \notin \text{Clr}$ and $\tilde{u} + h_{\text{tr}}(\tilde{u}) \leq P - 1 \leq J$,

$$l(y_{\tilde{u}+h_{\text{tr}}(\tilde{u})}^p) - l(y_{\tilde{u}}^p) = \sum_{k=\tilde{u}+1}^{\tilde{u}+h_{\text{tr}}(\tilde{u})} l(\text{pt}(b^k)) > \frac{13}{9}\beta(\tilde{u}).$$

So $l(y_u^p) > l_*(y_u^p)$ for all $u \geq \tilde{u} + h_{\text{tr}}(\tilde{u})$, by Lemma 12.2(3). The run $[\tilde{u} + h_{\text{tr}}(\tilde{u}) + 1, \tilde{u} + h_{\text{tr}}(\tilde{u}) + T_{\text{tr}} + 1]$ contains a hit u'' ; the point $y_{u''}^p$ lies above $l_*(y_u^p)$, so u_{i+1} exists and $u_{i+1} \leq u'' \leq p_i$. $\square$

Lemma 13.5 (Slow clearing is rare). *Let $p \in \mathcal{P}$ and $\omega \in \mathbb{N}$. If $P \leq J$, then $\mathbb{P}_{p,\omega}(\text{Clr}) < 2^{-43}$.*

Proof. Fix u with $u + h_{\text{tr}}(u) \leq P - 1$. By Lemma 13.2 with $k = u + h_{\text{tr}}(u) \leq J$, the points $\text{pt}(b^i)$, $u < i \leq u + h_{\text{tr}}(u)$, are independent with law η under $\mathbb{P}_{p,\omega}$. So, by the third bound of Lemma 10.9, the atoms with $\sum_{i=u+1}^{u+h_{\text{tr}}(u)} l(\text{pt}(b^i)) \leq \frac{13}{9}\beta(u)$ have mass at most $(4095/4096)^{\frac{127}{10}h_{\text{tr}}(u) - \frac{13}{9}\beta(u)}$. Since $h_{\text{tr}}(u) \geq \frac{4}{35}\beta(u)$ and $\beta(u) \geq K_{\text{tr}} \cdot (u + 1)$,

$$\frac{127}{10}h_{\text{tr}}(u) - \frac{13}{9}\beta(u) \geq \left(\frac{508}{350} - \frac{13}{9}\right)\beta(u) = \frac{11}{1575}\beta(u) \geq 32 \cdot 4096(u + 1),$$

using the integer comparison $11K_{\text{tr}} = 108493309639 \geq 32 \cdot 1575 \cdot 4096 = 206438400$. As $\log(1 - 1/4096) \leq -1/4096$, this mass is at most $e^{-32(u+1)}$. Summing over $u \geq 0$ gives at most $2e^{-32}$, and $2e^{-32} < 2^{-43}$ because $44 \log 2 < 44 \cdot 347/500 < 32$. $\square$

Lemma 13.6 (Case 3 low-reward bound). *Let $p \in \mathcal{P}$ and $\omega, m \in \mathbb{N}$ with $j(p) + m = J$. If $P \leq J$, then*

$$\mathbb{P}_{p,\omega}(\text{Low}^p \setminus \text{Bad}_m^p) \leq \mathbb{P}_{p,\omega}(\text{E}^{*,p}) + \frac{1}{P_{\text{tr}}} + 2^{-43}.$$

Proof. Let $\mathcal{Q} = \text{Low}^p \setminus (\text{Bad}_m^p \cup \text{E}^{*,p} \cup \text{Clr})$. Since $\mathbb{P}_{p,\omega}(\text{Clr}) < 2^{-43}$ (Lemma 13.5), it suffices to prove the Markov bound $\mathbb{P}_{p,\omega}(\mathcal{Q}) \leq e^{T_{\text{tr}}\gamma} r_2^{r_{\text{tr}}/2} = e^{101\gamma} r_2^{290}$, which is less than $1/P_{\text{tr}}$ by Lemma A.14. Let $a = (x, \mathbf{b}) \in \mathcal{Q}$ and put $\bar{o} = p + x$. By Lemma 13.4, $u_{r_{\text{tr}}+1}$ exists and $u_{r_{\text{tr}}+1} \leq P - 1$. The accumulated reward is nondecreasing, so $\text{Rw}_{u_{r_{\text{tr}}+1}}^p(a) \leq \text{Rw}_{P-1}^p(a) \leq T_{\text{tr}}$, and hence $\mathbf{1}[a \in \mathcal{Q}] \leq e^{T_{\text{tr}}\gamma} \mathcal{Z}_{r_{\text{tr}}+1}(\bar{o}, \mathbf{b})$. Sum against $\mathbb{P}_{p,\omega}$: Lemma 12.11(2) with $k = r_{\text{tr}}$, which is even, and $L = J$ bounds the sum over $\mathbf{b}$ by $r_2^{r_{\text{tr}}/2}$, and $\sum_x \mathbf{F}_\omega(x) = 1$. $\square$

13.3. Row geometry of triangles. For a triangle Δ and $l \in \mathbb{Z}$ put

$$\text{tip}_\Delta = l_\Delta - \frac{S_\Delta}{\log 2}, \quad \text{row}_\Delta(l) = j_\Delta + \lambda(l - \text{tip}_\Delta).$$

Thus tip_Δ is the (real) height at which the hypotenuse meets the left edge, and $\text{row}_\Delta(l)$ is the right end of row l . After a first passage over height ω the fresh path lies a little above the top edge of Δ , about $\omega/4$ to the right, while the rows of Δ widen by $\lambda\omega$ over that height; the margin $\lambda_0 = \lambda - \frac{1}{4}$ satisfies $\frac{1}{16} \leq \lambda_0 \leq \frac{3}{20}$ (Lemma A.10(1)). Making explicit the geometric steps of the proof of [10, Lemma 7.10], we show that large triangles met there have tips at or above l_Δ (Lemma 13.8) and corners in well-separated columns (Lemma 13.9), so that the packing bound of Lemma 10.14 applies.

Lemma 13.7 (Rows of a triangle). *Let Δ be a triangle and $(j, l) \in \mathcal{P}$ with $l \leq l_\Delta$. Then $(j, l) \in \Delta$ if and only if $j_\Delta \leq j \leq \text{row}_\Delta(l)$. In particular every point $(j, l) \in \Delta$ satisfies $l \geq \text{tip}_\Delta$ and $(j - j_\Delta) \log 9 \leq (l - \text{tip}_\Delta) \log 2$.*

Proof. Since $S_\Delta - (l_\Delta - l) \log 2 = (l - \text{tip}_\Delta) \log 2$, the third membership condition reads, as in [10, (7.65)], $(j - j_\Delta) \log 9 \leq (l - \text{tip}_\Delta) \log 2$, that is, $j \leq \text{row}_\Delta(l)$. With $j \geq j_\Delta$ it forces $l \geq \text{tip}_\Delta$. $\square$

Lemma 13.8 (Tips of nearby triangles). *Let $\Delta \neq \Delta'$ be members of $\mathfrak{T}$, let $p \in \Delta$ and $\omega = \text{gap}(p)$. Let $(j, l) \in \Delta'$ with $l_\Delta \leq l \leq l_\Delta + W$ and $|j - (j(p) + \omega/4)| \leq Y$, where $W, Y \geq 0$ and $Y + \lambda W + 1 \leq \omega\lambda_0$. Then $\text{tip}_{\Delta'} \geq l_\Delta$.*

Proof. Suppose $\text{tip}_{\Delta'} < l_\Delta$; as in the proof of [10, Lemma 7.10], we find a point in both triangles, contradicting Proposition 9.13. Put $\bar{j} = \text{row}_{\Delta'}(l_\Delta)$, so $\bar{j} > j_{\Delta'}$. By Lemma 13.7, $j \leq \text{row}_{\Delta'}(l) = \bar{j} + \lambda(l - l_\Delta) \leq \bar{j} + \lambda W$. Let $j' = \min(j, \lfloor \bar{j} \rfloor)$. Then $j_{\Delta'} \leq j' \leq \bar{j}$ and $0 \leq j - j' \leq \lambda W + 1$. Since $l_\Delta \leq l \leq l_{\Delta'}$, Lemma 13.7 gives $(j', l_\Delta) \in \Delta'$.

On the other hand $\omega \geq 0$ and $\lambda_0 \leq 1/4$, so $Y + \lambda W + 1 \leq \omega \lambda_0 \leq \omega/4$. Hence $j' \geq j - \lambda W - 1 \geq j(p) + \omega/4 - (Y + \lambda W + 1) \geq j(p) \geq j_\Delta$, and $j' \leq j \leq j(p) + \omega/4 + (Y + \lambda W + 1) \leq j(p) + (\frac{1}{4} + \lambda_0)\omega = j(p) + \lambda\omega$. Multiplying the last inequality by $\log 9$ and using the membership of p ,

$$(j' - j_\Delta) \log 9 \leq (j(p) - j_\Delta) \log 9 + \omega \log 2 \leq S_\Delta.$$

Thus $(j', l_\Delta) \in \Delta$ as well. $\square$

For $\Delta \in \mathfrak{T}$ and reals S and $W \geq 0$, the *anchor set* $\text{Anc}(\Delta, S, W) \subseteq \mathbb{Z}$ consists of the raw lengths $j_{\Delta'}$ of the corners of the $\Delta' \in \mathfrak{T}$ with $S_{\Delta'} \geq S$ and $l_\Delta \leq \text{tip}_{\Delta'} \leq l_\Delta + W$.

Lemma 13.9 (Anchors are separated). *Let $\Delta \in \mathfrak{T}$, let $X \geq 1$ be an integer and $S \geq 4096X$. Distinct elements $\zeta < \zeta'$ of $\text{Anc}(\Delta, S, 2X)$ satisfy $\zeta' - \zeta > \frac{36}{25}\lambda S$.*

Proof. Let Δ', Δ'' be triangles defining $\zeta = j_{\Delta'}$ and $\zeta' = j_{\Delta''}$; they are distinct. As in the proof of [10, Lemma 7.10], we compare their rows at one height, here $\hat{l} = l_\Delta + \lfloor S/\log 2 \rfloor$. Since $S/\log 2 > S \geq 2X + 1$, we have $\text{tip}_{\Delta'} \leq l_\Delta + 2X \leq \hat{l}$ and $l_{\Delta'} = \text{tip}_{\Delta'} + S_{\Delta'}/\log 2 \geq \hat{l}$, and likewise for Δ'' . By Lemma 13.7, $(\zeta', \hat{l}) \in \Delta''$. If $\zeta' \leq \text{row}_{\Delta'}(\hat{l})$, then $\zeta \leq \zeta'$ and Lemma 13.7 give $(\zeta', \hat{l}) \in \Delta'$ as well, contradicting Proposition 9.13. Hence

$$\zeta' - \zeta > \lambda(\hat{l} - \text{tip}_{\Delta'}) \geq \lambda\left(\frac{S}{\log 2} - 1 - 2X\right).$$

Finally $\log 2 < 347/500$ gives $\frac{S}{\log 2} - 1 - 2X - \frac{36}{25}S > \frac{8S}{8675} - 2X - 1 \geq \frac{32768}{8675}X - 2X - 1 > 0$, so $\zeta' - \zeta > \frac{36}{25}\lambda S$. $\square$

13.4. The packing estimate. Outside a small deviation event the first passage and the next v holds stay close to their typical displacements, so a visit to a large triangle pins $j(x)$ near an anchor (Lemma 13.11) and the packing bound applies; $\text{Exc}(v)$ collects the tail probabilities. Let $\omega, v \in \mathbb{N}$ with $v \leq J$. For real $X > 0$, the *deviation event* $\text{Dev}_{\omega, v, X}$ is the set of atoms satisfying at least one of

$$l(x) - \omega \geq X, \quad l(\text{pt}_v(\mathbf{b})) \geq X, \quad |j(x) - \omega/4| \geq \omega^{3/5}, \quad j(\text{pt}_v(\mathbf{b})) \geq \omega^{3/5}.$$

For $p \in \mathcal{P}$ and real S , the *large-triangle event* $\text{Big}_{v, S}^p$ is the set of atoms a with $y_u^p(a) \in \Delta$ for some $u \leq v$ and some $\Delta \in \mathfrak{T}$ with $S_\Delta \geq S$. The *deviation error* is

$$\text{Exc}(v) = 9q_{\text{ex}}^{v+1} + \mathbf{1}[v \geq 1](q_{\text{ex}}(8q_{\text{ex}})^v + e^{-14(v+1)}), \quad q_{\text{ex}} = (25/27)^{136}.$$

Lemma 13.10 (Mass of the deviation event). *Let $p \in \mathcal{P}$, let $\omega \geq 2^{160}$ be an integer, $v \leq J$, $X = 136(v+1)$, and let S be real with $4096X \leq S$ and $S^2 < 2\omega$. Then $\mathbb{P}_{p, \omega}(\text{Dev}_{\omega, v, X}) \leq 1/S + \text{Exc}(v)$.*

Proof. We follow the proof of [10, (7.61)]. Each of the four conditions involves only x or only the first v hold points, so by Lemma 13.2 it suffices to bound four probabilities under F_ω or $\eta^{\otimes v}$. First, the first-passage overshoot tail (Lemma 10.10, valid for $\omega \geq 64$ and integer $X \geq 1$) bounds the probability of $l(x) \geq \omega + X$ by $9(25/27)^X = 9q_{\text{ex}}^{v+1}$. Second, if $v = 0$ then $l(\text{pt}_v(\mathbf{b})) = 0 < X$; otherwise the second bound of Lemma 10.9 bounds the probability of $l(\text{pt}_v(\mathbf{b})) \geq X$ by $8^v(25/27)^X = q_{\text{ex}}(8q_{\text{ex}})^v$. Third, the horizontal tail of the first passage (Lemma 10.15 with threshold $\omega^{3/5}$, where both exponents are at least $\omega^{1/5}/65536$ as $1 + \omega \leq 2\omega$) and Lemma A.10(4) bound the probability of $|j(x) - \omega/4| \geq \omega^{3/5}$ by

$$2^{55}e^{-\omega^{1/5}/65536} < (2\omega)^{-1/2} < 1/S.$$

Fourth, if $v = 0$ then $j(\text{pt}_v(\mathbf{b})) = 0$. Otherwise it is an integer, so the event is $10 \lfloor \omega^{3/5} \rfloor \leq 10j(\text{pt}_v(\mathbf{b}))$, and the first bound of Lemma 10.9 bounds its probability by $e^{v/2 - \omega^{3/5}/16}$. Since $\omega \geq 1$, $S < \sqrt{2\omega} \leq 16\omega^{3/5}$, so $\omega^{3/5}/16 > S/256 \geq 16X = 2176(v+1)$, and the bound is at most $e^{-14(v+1)}$. Adding the four bounds gives $1/S + \text{Exc}(v)$. $\square$

Lemma 13.11 (Outside the deviation event, large triangles sit at anchors). *Let $p \in \Delta \in \mathfrak{T}$, $\omega = \text{gap}(p)$, $v \leq J$, $X \geq 1$ an integer and S real, and assume $2\omega^{3/5} + 2\lambda X + 1 \leq \omega\lambda_0$. If an atom $a = (x, \mathbf{b})$ of positive mass lies in $\text{Big}_{v,S}^p$ but not in $\text{Dev}_{\omega,v,X}$, then some $\zeta \in \text{Anc}(\Delta, S, 2X)$ satisfies $-X < j(x) + j(p) - \zeta < 2\lambda X$.*

Proof. We follow [10, (7.63)–(7.66)], for all $u \leq v$ at once. Choose $u \leq v$ and $\Delta' \in \mathfrak{T}$ with $S_{\Delta'} \geq S$ and $(j, l) = y_u^p(a) \in \Delta'$. By Definition 13.1, $j = j(p) + j(x) + j(\text{pt}_u(\mathbf{b}))$ and $l = l(p) + l(x) + l(\text{pt}_u(\mathbf{b}))$. Since a has positive mass and lies outside the deviation event, Lemmas 10.6(1) and 10.4(3) give

$$0 < l(x) - \omega < X, \quad 0 \leq j(\text{pt}_u(\mathbf{b})) \leq l(\text{pt}_u(\mathbf{b})) \leq l(\text{pt}_v(\mathbf{b})) < X,$$

$$|j(x) - \omega/4| < \omega^{3/5}, \quad j(\text{pt}_u(\mathbf{b})) \leq j(\text{pt}_v(\mathbf{b})) < \omega^{3/5}.$$

Hence $0 < l - l_\Delta = (l(x) - \omega) + l(\text{pt}_u(\mathbf{b})) < 2X$ and $|j - (j(p) + \omega/4)| < 2\omega^{3/5}$. As $l > l_\Delta$, the point is not in Δ , so $\Delta' \neq \Delta$. Lemma 13.8, with $W = 2X$ and $Y = 2\omega^{3/5}$, gives $\text{tip}_{\Delta'} \geq l_\Delta$; and $\text{tip}_{\Delta'} \leq l < l_\Delta + 2X$ by Lemma 13.7. Thus $\zeta = j_{\Delta'} \in \text{Anc}(\Delta, S, 2X)$. Lemma 13.7 also gives $(j - \zeta) \log 9 \leq (l - \text{tip}_{\Delta'}) \log 2 \leq (l - l_\Delta) \log 2 < 2X \log 2$, so $0 \leq j - \zeta < 2\lambda X$. Since $j(x) + j(p) - \zeta = (j - \zeta) - j(\text{pt}_u(\mathbf{b}))$ and $0 \leq j(\text{pt}_u(\mathbf{b})) < X$, the claim follows. $\square$

Lemma 13.12 (Large triangles outside the deviation event). *Let $p \in \Delta \in \mathfrak{T}$, let $\omega = \text{gap}(p) \geq 1024$, $v \leq J$, let $X \geq 1$ be an integer and S real with $4096X \leq S$ and $S^2 < 2\omega$, and assume $2\omega^{3/5} + 2\lambda X + 1 \leq \omega\lambda_0$. Then*

$$\mathbb{P}_{p,\omega}(\text{Big}_{v,S}^p \setminus \text{Dev}_{\omega,v,X}) < Q_{\text{pk}} \frac{X}{S},$$

with Q_{pk} of Lemma A.10(2).

Proof. We follow the proof of [10, (7.62)]. The integers i with $-X < i < 2\lambda X$ number $X - 1 + \lceil 2\lambda X \rceil < (1 + 2\lambda)X$. For each such i put $\mathcal{K}_i = \{\zeta - j(p) + i : \zeta \in \text{Anc}(\Delta, S, 2X)\}$; by Lemma 13.9, distinct elements of $\mathcal{K}_i$ differ by more than $d = \frac{36}{25}\lambda S$. By Lemma 13.11, every atom of positive mass in the event has $j(x) \in \bigcup_i \mathcal{K}_i$, and the law of $j(x)$ is the horizontal first-passage law. The packing bound for the horizontal first-passage coordinate (Lemma 10.14, for $\omega \geq 1024$, separation at least d , applied to the set $\mathcal{K}_i$) and the union bound over these i bound the mass of the event by

$$(1 + 2\lambda)X \left(\frac{17}{10\sqrt{\omega}} + \frac{33}{32d} \right) < \frac{X}{S} \left((1 + 2\lambda) \frac{17}{10} \sqrt{2} + \left(\frac{1}{\lambda} + 2 \right) \frac{33}{32} \cdot \frac{25}{36} \right),$$

using $1/\sqrt{\omega} < \sqrt{2}/S$ and $(1 + 2\lambda)/\lambda = 1/\lambda + 2$. Lemma A.10(2) completes the proof. $\square$

Lemma 13.13 (Packing estimate; explicit form of Tao's Lemma 7.10). *Let $p \in \Delta \in \mathfrak{T}$, let $\omega = \text{gap}(p) \geq 2^{160}$, let $v \leq J$ and $X = 136(v + 1)$, and let S be real with $4096X \leq S$ and $S^2 < 2\omega$. Assume $2\omega^{3/5} + 2\lambda X + 1 \leq \omega\lambda_0$. Then*

$$\mathbb{P}_{p,\omega}(\text{Big}_{v,S}^p) \leq Q_{\text{pk}} \frac{X}{S} + \frac{1}{S} + \text{Exc}(v). \quad (13.1)$$

Proof. Split $\text{Big}_{v,S}^p$ along $\text{Dev}_{\omega,v,X}$ and add Lemmas 13.10 and 13.12. $\square$

Corollary 13.14 (Mass of the oversize event). *Let $p \in \Delta \in \mathfrak{T}$ and $\omega = \text{gap}(p)$. Let $m \geq D_{\text{sc}}$ be real with $P \leq J$ and $m/(\log m)^2 < \omega$. Then the oversize event $E^{*,p}$ of Definition 13.3 satisfies*

$$\mathbb{P}_{p,\omega}(E^{*,p}) < \frac{2 + 272(H_{\text{tr}} + 1)Q_{\text{pk}}}{K_{\text{tr}}} + \frac{1}{3360}.$$

Proof. Put $\hat{u} = P - 1$, which is at least 2 by Lemma A.6, take $v(k)$, $X(k)$ and $S(k)$ from Lemma A.4(3) with this $\hat{u}$, and put $\bar{k} = \lfloor \log_2(\hat{u} + 1) \rfloor$. A union bound over the dyadic ranges $0 \leq k \leq \bar{k}$ of times reduces the oversize event to one packing estimate per range. Indeed, let $u \leq \hat{u}$ and choose k with

$2^k \leq u + 1 < 2^{k+1}$. Then $k \leq \bar{k}$ and $u \leq v(k)$, and since $u \geq 2^k - 1$ and β is nondecreasing (Lemma A.4(1)), $\beta(u) \geq S(k)$. Hence $E^{*,p} \subseteq \bigcup_{k=0}^{\bar{k}} \text{Big}_{v(k), S(k)}^p$.

Fix $0 \leq k \leq \bar{k}$; we check the hypotheses of Lemma 13.13 with $(v, X, S) = (v(k), X(k), S(k))$. First $v(k) \leq P - 1 < J$, and $4096X(k) < S(k)$ by Lemma A.4(3). As $2^k \leq \hat{u} + 1$, monotonicity gives $S(k) = \beta(2^k - 1) \leq \beta(P - 1) = B_{sz}$. By Lemma A.8(2) and Lemma A.6, the integer ω satisfies $\omega > m/(\log m)^2 > \omega_{\text{far}} = 2^{36}P > 2^{365}$, and $S(k)^2 \leq B_{sz}^2 < 2m/(\log m)^2 < 2\omega$. Finally, as $2\lambda X(k) = 272\lambda(v(k) + 1)$, Lemma A.10(3) gives $2\omega^{3/5} + 2\lambda X(k) + 1 \leq \omega\lambda_0$.

Sum (13.1) over $0 \leq k \leq \bar{k}$. The partial sums of (A.3) bound the first two terms by $(272(H_{\text{tr}} + 1)Q_{\text{pk}} + 2)/K_{\text{tr}}$, with strict inequality, and Lemma A.11 bounds $\sum_{k \leq \bar{k}} \text{Exc}(v(k))$ by $1/3360$. $\square$

13.5. The outer horizontal tail and the assembly. We bound the event Bad_m^p of Definition 13.3, continue Q through the first passage, and assemble Case 3.

Lemma 13.15 (Outer horizontal tail). *Let $p \in \mathcal{P}$, $\omega \in \mathbb{N}$, and let $m \geq D_{\text{bad}}$ be an integer with $P \leq J$ and $\omega \log 2 \leq m \log 9$. Then $m^{A_*} \mathbb{P}_{p,\omega}(\text{Bad}_m^p) < 1/8000$.*

Proof. As in [10, proof of (7.54)], we split the advance between the first passage and the next P holds, with explicit thresholds. By Definition 13.1, Bad_m^p is the set of atoms with $400(j(x) + j(\text{pt}_P(\mathbf{b}))) \geq 317m$. Since $317/400 = 158500/200000$, it is contained in the union of $\{j(x) \geq \frac{158499}{200000}m\}$ and $\{j(\text{pt}_P(\mathbf{b})) \geq \frac{m}{200000}\}$.

By Lemma A.10(1), $\omega \leq m/\lambda < \frac{158497}{50000}m$. On the first event, $j(x) - \omega/4 \geq m/100000$. The horizontal tail of the first passage (Lemma 10.15 with threshold $m/100000$) bounds its mass by $2^{54}(e^{-m^2/(10^{10}2^{15}(1+\omega))} + e^{-m/(256 \cdot 10^5)})$. As $m \geq 1$, $1 + \omega \leq \frac{21}{5}m$, and the integer comparison $21 \cdot 2^{15} \cdot 10^{10} \leq 5 \cdot 2^{51}$ makes both exponents at most $-m/2^{51}$. So the first event has mass at most $2^{55}e^{-m/2^{51}}$. On the second event, $10 \lceil m/200000 \rceil \leq 10j(\text{pt}_P(\mathbf{b}))$, and the first bound of Lemma 10.9, applied via Lemma 13.2 with $k = N = P$ and threshold $10 \lceil m/200000 \rceil$, bounds its mass by $e^{P/2 - m/320000} \leq e^{P/2}e^{-m/2^{51}}$. Therefore

$$m^{A_*} \mathbb{P}_{p,\omega}(\text{Bad}_m^p) \leq m^{A_*} (2^{55} + e^{P/2}) e^{-m/2^{51}}.$$

Since $m \geq D_{\text{bad}} = (2^{53}A_*)^2 + 2^{53}(P + 65)$, we have $A_* \log m \leq A_* \sqrt{m} \leq m/2^{53}$ and $P + 56 \leq m/2^{53} - 9$; and $2^{55} + e^{P/2} \leq 2e^{P+55} \leq e^{P+56}$. The logarithm of the right side is therefore at most $2m/2^{53} - 9 - 4m/2^{53} < -9$, and $e^9 > 8000$ (Lemma A.1(4)). $\square$

Lemma 13.16 (Continuation through the first passage). *Let $p \in \mathcal{P}$, $\omega \in \mathbb{N}$ and $1 \leq L \leq J$. Then*

$$Q(p) \leq \sum_a \mathbb{P}_{p,\omega}(a) e^{-z \text{Rw}_{L-1}^p(a)} Q(y_L^p(a)).$$

Proof. We adapt [10, (7.45), (7.53)]. Lemma 11.9 gives $Q(p) \leq \sum_{x \in \mathcal{P}} F_\omega(x) Q(p + x)$. Fix x with $F_\omega(x) \neq 0$ and put $\bar{o} = p + x \in \mathcal{P}$. Induction on K , substituting Lemma 11.7(2) at the point $\bar{o} + \text{pt}_K(\mathbf{b})$ in the inductive step, gives with $K = L$ the continuation identity

$$Q(\bar{o}) = \sum_{\mathbf{b} \in \mathfrak{B}^L} \mathbf{q}(\mathbf{b}) \prod_{k=0}^{L-1} \mathbf{w}(\bar{o} + \text{pt}_k(\mathbf{b})) \prod_{k=1}^L \mathbf{l}(\bar{o} + \text{pt}_{k-1}(\mathbf{b}); b^k) Q(\bar{o} + \text{pt}_L(\mathbf{b})),$$

with all factors in $[0, 1]$. Drop $\mathbf{w}(\bar{o})$ and $\mathbf{l}(\bar{o} + \text{pt}_{L-1}(\mathbf{b}); b^L)$. Since $\text{pt}_k(\mathbf{b}) = \text{pt}_{k-1}(\mathbf{b}) + \text{pt}(b^k)$, the identity recorded after Definition 12.1 gives

$$\mathbf{w}(\bar{o} + \text{pt}_k(\mathbf{b})) \mathbf{l}(\bar{o} + \text{pt}_{k-1}(\mathbf{b}); b^k) = e^{-z \text{rw}(\bar{o} + \text{pt}_{k-1}(\mathbf{b}), b^k)} \quad \text{for } 1 \leq k \leq L - 1,$$

and the product of these is $e^{-z \text{Rw}(\bar{o}, \mathbf{b}; L-1)}$. Extending $\mathbf{b}$ by $J - L$ further holds changes neither $\text{Rw}(\bar{o}, \mathbf{b}; L - 1)$ nor $\bar{o} + \text{pt}_L(\mathbf{b})$ (Lemma 12.2(2)), and the extensions of $\mathbf{b}$ have total mass $\mathbf{q}(\mathbf{b})$ (Lemmas 12.2(1) and 10.4(5)). Multiplying by $F_\omega(x)$ and summing over x gives the claim, by Definition 13.1. $\square$

Lemma 13.17 (Case 3 of the descent step). *Let m be an integer with $m \geq D_{\text{sc}}$, $m \geq D_{\text{bad}}$ and $m \leq J$. Let $p \in \mathcal{P}$ with $j(p) + m = J$, and suppose $p \in \Delta$ for some $\Delta \in \mathfrak{T}$ with $\omega = \text{gap}(p) > m/(\log m)^2$. Then $\mathbf{Q}(p) \leq m^{-A_*} \mathbf{Q}_{m-1}$.*

Proof. We follow [10, (7.52)–(7.56)], with $\frac{317}{400}m$ in place of $0.9m$ and an explicit error budget. By Lemma A.8(2), $P \leq m \leq J$, and ω is a positive integer. The right-end bound (9.2) and the membership of p give

$$\omega \log 2 \leq S_\Delta - (j(p) - j_\Delta) \log 9 < (J - j(p)) \log 9 = m \log 9.$$

Lemma 13.16 with $L = P$ gives $\mathbf{Q}(p) \leq \sum_a \mathbb{P}_{p,\omega}(a) e^{-z \text{Rw}_{P-1}^p(a)} \mathbf{Q}(y_P^p(a))$. Let a have positive mass and put $\text{Adv}(a) = j(y_P^p(a)) - j(p) = j(x) + j(\text{pt}_P(\mathbf{b})) \geq 1$. Then $j(y_P^p(a)) + m = J + \text{Adv}(a)$, so the boundary reduction (Lemma 11.12(1) with $r = \text{Adv}(a)$) gives

$$e^{-z \text{Rw}_{P-1}^p(a)} \mathbf{Q}(y_P^p(a)) \leq e^{-z \text{Rw}_{P-1}^p(a)} \max(m - \text{Adv}(a), 1)^{-A_*} \mathbf{Q}_{m-1}.$$

Since $\mathbf{Q}_{m-1} \geq 0$, the right side is at most $m^{A_*} \cdot m^{-A_*} \mathbf{Q}_{m-1}$. If $a \notin \text{Bad}_m^p$, then $m - \text{Adv}(a) > \frac{83}{400}m$, so the right side is at most $(400/83)^{A_*} m^{-A_*} \mathbf{Q}_{m-1} \leq c_{\text{cont}} m^{-A_*} \mathbf{Q}_{m-1}$, with $c_{\text{cont}} = 511/10$ of Definition A.15. If moreover $a \notin \text{Low}^p$, then $\text{Rw}_{P-1}^p(a) > T_{\text{tr}} = 101$, and the right side is at most $e^{-z T_{\text{tr}}} (400/83)^{A_*} m^{-A_*} \mathbf{Q}_{m-1} \leq m^{-A_*} \mathbf{Q}_{m-1} / P_{\text{wc}}$, with P_{wc} of the same definition. The two numerical bounds are Lemma A.16(1),(2). Splitting the atoms into Bad_m^p , $\text{Low}^p \setminus \text{Bad}_m^p$ and the rest, and using that $\mathbb{P}_{p,\omega}$ has total mass 1,

$$\mathbf{Q}(p) \leq m^{-A_*} \mathbf{Q}_{m-1} \left(m^{A_*} \mathbb{P}_{p,\omega}(\text{Bad}_m^p) + c_{\text{cont}} \mathbb{P}_{p,\omega}(\text{Low}^p \setminus \text{Bad}_m^p) + \frac{1}{P_{\text{wc}}} \right).$$

Lemma 13.15 bounds the first term by $1/8000$. The Case 3 low-reward bound (Lemma 13.6, which uses $j(p) + m = J$ and $P \leq J$) gives $\mathbb{P}_{p,\omega}(\text{Low}^p \setminus \text{Bad}_m^p) \leq \mathbb{P}_{p,\omega}(\mathbf{E}^{*,p}) + 1/P_{\text{tr}} + 2^{-43}$, and Corollary 13.14 bounds $\mathbb{P}_{p,\omega}(\mathbf{E}^{*,p})$. Since $c_{\text{cont}}/3360 < 1/60$ (Lemma A.16(1)), the bracket is less than the *error budget*

$$\mathcal{B} = \frac{1}{8000} + \frac{c_{\text{cont}}(2 + 272(H_{\text{tr}} + 1)Q_{\text{pk}})}{K_{\text{tr}}} + \frac{c_{\text{cont}}}{P_{\text{tr}}} + \frac{1}{60} + \frac{1}{P_{\text{wc}}} + c_{\text{cont}} 2^{-43}, \quad (13.2)$$

and $\mathcal{B} < 1$ (Lemma A.16(3)). $\square$

14. THE WHITE AND NEAR-TOP CASES AND THE FOURIER BOUND

This section proves Cases 1 and 2 of the descent step (Section 11), where the boundary point p for m is white or black with $\text{gap}(p) \leq m/(\log m)^2$, following Cases 1 and 2 of [10, §7.4] with the constants of Appendix A.1. With Case 3 (Lemma 13.17) this gives the descent step for the weighted maxima beyond the threshold D_* , from which (4.2) follows.

14.1. Envelope factors. To average the bound of Lemma 11.12(1) over the hold law or a first-passage law, we dominate $\max(m - r, 1)^{-A_*}$ by m^{-A_*} times the *envelope factor* $e^{\mathbf{t}_m r}$ (compare [10, (7.42)]), whose average is an exponential moment; here, for an integer $m \geq 1$,

$$\mathbf{t}_m = \frac{4A_*(1 + \log m)}{m}. \quad (14.1)$$

Lemma 14.1 (Envelope bound). *Let $m \geq 2$ and $r \geq 1$ be integers. Then*

$$\max(m - r, 1)^{-A_*} \leq m^{-A_*} e^{\mathbf{t}_m r}. \quad (14.2)$$

If $m \geq D_{\text{wc}}$, then $\mathbf{t}_m \leq \chi/32 \leq 1/1024$, and

$$\sum_{r \geq 1} \nu_{45}(r) \max(m - r, 1)^{-A_*} \leq M_{45}(\mathbf{t}_m) m^{-A_*} \leq e^{\chi/2} m^{-A_*}. \quad (14.3)$$

Proof. If $2r \leq m$, then $(1 - r/m)(1 + 2r/m) \geq 1$ gives $(1 - r/m)^{-1} \leq 1 + 2r/m \leq e^{2r/m}$, so $(m - r)^{-A_*} \leq m^{-A_*} e^{2A_* r/m}$, and $2A_*/m \leq t_m$. If $2r > m$, then $\max(m - r, 1)^{-A_*} \leq 1 = m^{-A_*} e^{A_* \log m}$ and $A_* \log m \leq A_*(1 + \log m) \cdot 4r/m = t_m r$.

Let $m \geq D_{wc}$, where $D_{wc} = (256A_*/\chi)^2 + 2$ by Definition A.7. Since $\log y \leq \sqrt{y}$ for $y > 0$, $t_m \leq 8A_*/\sqrt{m} \leq 8A_* \cdot \chi/(256A_*) = \chi/32 = 63/80000 < 1/1024$. Summing (14.2) against ν_{45} gives the first inequality of (14.3), and Lemma 10.8(2) gives $M_{45}(t_m) \leq e^{13t_m/4} \leq e^{13\chi/128} \leq e^{\chi/2}$. $\square$

14.2. The white case. At a white boundary point the closing letter at p carries the factor e^{-z} , which outweighs the cost $e^{\chi/2}$ of averaging the envelope factor over one hold.

Lemma 14.2 (White boundary points). *Let $m \geq D_{wc}$ be an integer, and let $p \in \mathcal{P}$ be white with $j(p) + m = J$. Then*

$$Q(p) \leq e^{-z+\chi/2} m^{-A_*} Q_{m-1} \leq m^{-A_*} Q_{m-1}.$$

Proof. We follow [10, (7.43)]. The second inequality holds as $\chi/2 = 63/5000 < z$. By Lemma 11.7(3) and $w(p) = e^{-z}$, $Q(p) \leq e^{-z} \sum_y \eta(y) Q(p+y)$. Every $y \in \mathcal{P}$ has $j(y) \geq 1$ and $j(p+y) + m = J + j(y)$, so Lemma 11.12(1) gives $Q(p+y) \leq \max(m - j(y), 1)^{-A_*} Q_{m-1}$. Grouping by $r = j(y)$ and using (10.2),

$$\sum_y \eta(y) Q(p+y) \leq Q_{m-1} \sum_{r \geq 1} \nu_{45}(r) \max(m - r, 1)^{-A_*} \leq e^{\chi/2} m^{-A_*} Q_{m-1}$$

by (14.3), since $m \geq D_{wc} \geq 2$ and $Q_{m-1} \geq 0$. $\square$

14.3. The near-top case. Following Case 2 of [10, §7.4], we halt the hold walk from a black boundary point p near the top of its triangle Δ when its height first exceeds l_Δ . With probability at least $\frac{3}{16}$ it halts in the exit window, at a white point where Case 1 applies; there, after paying $\chi/2$, it still gains a factor $1 - \frac{3}{5}z$ (Lemma A.9), which more than pays the cost $\chi/8$ of the envelope factors along the first passage (Lemma 14.3).

Lemma 14.3 (Moment threshold). *Let m be an integer with $m \geq \max(D_{wc}, D_{exp})$, and let $\omega \in \mathbb{N}$ with $\omega \leq m/(\log m)^2$. Then, with t_m of (14.1), $M_{45}(t_m)^{\omega+1} \leq 1 + \chi/8$.*

Proof. Part (3) of Lemma A.8 gives $\chi \log m > 234A_*$; as $A_* \geq 1 \geq \chi$, also $\log m > 234$. First, because $\log m \geq 32$,

$$8t_m m/(\log m)^2 = 32A_*(1 + \log m)/(\log m)^2 \leq A_*/\log m + 32A_*/\log m.$$

Second, $m \geq (\log m)^4/24$ and $(\log m)^3 \geq 1024 \log m \geq 256(1 + \log m)$, so $3m \geq (\log m)^4/8 \geq 32 \log m (1 + \log m)$, that is $8t_m \leq 3A_*/\log m$. Adding, $8t_m(\omega + 1) \leq 8t_m(m/(\log m)^2 + 1) \leq 36A_*/\log m$. By Lemma 14.1, $t_m \leq 1/1024$, so Lemma 10.8(2) gives

$$M_{45}(t_m)^{\omega+1} \leq \exp\left(\frac{13t_m(\omega+1)}{4}\right) \leq \exp\left(\frac{13}{32} \cdot \frac{36A_*}{\log m}\right) = \exp\left(\frac{117A_*}{8 \log m}\right) < e^{\chi/16}.$$

Finally $e^y \leq (1 - y)^{-1} \leq 1 + 2y$ for $0 \leq y \leq \frac{1}{2}$; take $y = \chi/16$. $\square$

Lemma 14.4 (Black boundary points near the top). *Let m be an integer with $m \geq \max(D_{pt}, D_{exp})$, and let $p \in \mathcal{P}$ with $j(p) + m = J$ lie in a triangle $\Delta \in \mathfrak{T}$ with $\text{gap}(p) \leq m/(\log m)^2$. Then*

$$Q(p) \leq m^{-A_*} Q_{m-1}.$$

Proof. Put $\omega = \text{gap}(p) \in \mathbb{N}$; every bound below is a multiple of $m^{-A_*} Q_{m-1} \geq 0$. Since $m \geq D_{pt} \geq D_{wc} \geq 2$, Lemma 14.1 gives $0 \leq t_m \leq 1/1024$, so $11e^{t_m} < 16$ and Lemma 10.9 applies with $t = t_m$. Let

$$\mathcal{X}_\omega = \{x \in \mathcal{P} : 1 \leq j(x) \leq f_\omega, \omega < l(x) \leq \omega + 4\}.$$

Let $x \in \mathcal{P}$. Then $j(x) \geq 1$ and $j(p+x) + m = J + j(x)$, so Lemma 11.12(1) and (14.2) give $Q(p+x) \leq m^{-A_*} Q_{m-1} e^{t_m j(x)}$.

Let moreover $x = (r, \omega + o) \in \mathcal{X}_\omega$, so $1 \leq r \leq f_\omega$ and $1 \leq o \leq 4$. Since $m \geq D_{\text{pt}} \geq 344 > e$, $\log m \geq 1$ and $\omega \leq m$. Hence $r \leq f_\omega \leq 5\omega/16 + 1 \leq 5m/16 + 1$ and, as $D_{\text{wc}} = (D_{\text{pt}} - 344)/2 \leq (m - 344)/2$ by Definition A.7,

$$r + D_{\text{wc}} \leq \frac{13m}{16} - 171 < m.$$

Put $m' = m - r$, so that $m' > D_{\text{wc}} \geq 2$ and $m' = \max(m - r, 1)$. The point $p' = p + x = (j(p) + r, l_\Delta + o)$ is white by Lemma 9.14, and $j(p') + m' = J$. So Lemma 14.2 at (p', m') , $\mathbf{Q}_{m'-1} \leq \mathbf{Q}_{m-1}$ (Lemma 11.11(3)), (14.2) and Lemma A.9 give

$$\mathbf{Q}(p + x) \leq e^{-z+\chi/2} m'^{-A_*} \mathbf{Q}_{m-1} \leq (1 - \frac{3}{5}z) m^{-A_*} \mathbf{Q}_{m-1} e^{\mathbf{t}_m j(x)}.$$

Since $e^{\mathbf{t}_m j(x)} \geq 1$, Lemmas 10.9 and 14.3 and the exit mass bound $\sum_{x \in \mathcal{X}_\omega} \mathbf{F}_\omega(x) = \mathbf{q}_{45}(\omega) \geq \frac{3}{16}$ (Lemmas 10.17 and 10.18, as $\mathbf{F}_\omega$ vanishes off $\mathcal{P}$ by Lemma 10.6(1)) yield

$$\begin{aligned} \sum_{x \in \mathcal{P}} \mathbf{F}_\omega(x) \mathbf{Q}(p + x) &\leq m^{-A_*} \mathbf{Q}_{m-1} \left(\sum_{x \in \mathcal{P}} \mathbf{F}_\omega(x) e^{\mathbf{t}_m j(x)} - \frac{3}{5}z \sum_{x \in \mathcal{X}_\omega} \mathbf{F}_\omega(x) \right) \\ &\leq m^{-A_*} \mathbf{Q}_{m-1} \left(1 + \frac{\chi}{8} - \frac{3}{5}z \cdot \frac{3}{16} \right) \leq m^{-A_*} \mathbf{Q}_{m-1}, \end{aligned}$$

the last step by Lemma A.9. Finally Lemma 11.9 gives $\mathbf{Q}(p) \leq \sum_{x \in \mathcal{P}} \mathbf{F}_\omega(x) \mathbf{Q}(p + x)$. $\square$

14.4. The descent step and the Fourier bound.

Proposition 14.5 (Descent). *For every integer m with $D_* \leq m \leq J$, $\mathbf{Q}_m \leq \mathbf{Q}_{m-1}$.*

Proof. By its definition D_* is at least each of D_{wc} , D_{pt} , D_{exp} , D_{bad} , D_{sc} and 2, hence so is m . By Lemma 11.12(2), it suffices to prove $\mathbf{Q}(p) \leq m^{-A_*} \mathbf{Q}_{m-1}$ for every $p \in \mathcal{P}$ with $j(p) + m = J$. Fix such a point.

Case 1. p is white. Apply Lemma 14.2.

Otherwise p is black, because $j(p) \leq J$, so it lies in some $\Delta \in \mathfrak{T}$ by Proposition 9.13.

Case 2. $\text{gap}(p) \leq m/(\log m)^2$. Apply Lemma 14.4.

Case 3. $\text{gap}(p) > m/(\log m)^2$. Apply Lemma 13.17. $\square$

Theorem 14.6 (The Fourier bound). *For every integer $n \geq 1$ and every frequency $\xi \in G_n$ prime to 3,*

$$|\hat{\mu}_n(\xi)| \leq (160D_*)^{A_*} n^{-A_*}.$$

This is (4.2).

Proof. Since ξ is prime to 3, it is a unit of G_n ; form $\mathbf{Q}$ and $\mathbf{Q}_m$ for (n, ξ, ε) and the exponent A_* . Proposition 11.8 gives $|\hat{\mu}_n(\xi)| \leq \sum_y \eta(y) \mathbf{Q}(y)$. We bound the right side in three steps, following [10, (7.36)–(7.40)].

First, $\mathbf{Q}_m \leq D_*^{A_*}$ for $0 \leq m \leq J$, by induction on m : if $m \leq D_*$, then $\mathbf{Q}_m \leq \max(m, 1)^{A_*} \leq D_*^{A_*}$ by Lemma 11.11(1), as $D_* \geq 2$; if $m > D_*$, then $\mathbf{Q}_m \leq \mathbf{Q}_{m-1}$ by Proposition 14.5. So Lemma 11.11(2) with $m = J$ gives $\mathbf{Q}(y) \leq D_*^{A_*} d_J(y)^{-A_*}$ for every $y \in \mathcal{P}$.

Second, let $y \in \mathcal{P}$ and $r = j(y) \geq 1$; then $n \leq 8r d_J(y)$. Indeed $n \leq 2J + 1$. If $r < J$, then $J \geq 2$ and $r(J - r) - (J - 1) = (r - 1)(J - r - 1) \geq 0$, so $n \leq 2J + 1 \leq 8(J - 1) \leq 8r(J - r)$. If $r \geq J$, then $d_J(y) = 1$ and $n \leq 2r + 1 \leq 8r$. Hence $\mathbf{Q}(y) \leq (8D_*)^{A_*} r^{A_*} n^{-A_*}$.

Third, $\sum_{r \geq 1} \nu_{45}(r) r^{A_*} \leq 20^{A_*}$. For $r \geq 1$, $r^5 \leq r(r + 1) \cdots (r + 4) = 5! \binom{r+4}{5}$, and the negative binomial series $\sum_{k \geq 0} \binom{k+5}{5} x^k = (1 - x)^{-6}$ at $x = 11/16$ gives

$$\sum_{r \geq 1} \nu_{45}(r) r^5 \leq \frac{5}{16} \cdot 120 \cdot \left(\frac{16}{5} \right)^6 = \frac{25165824}{625} < 20^5.$$

Since ν_{45} is a probability law, $1 \leq A_* \leq 5$ and $u \mapsto u^{A_*/5}$ is concave, Jensen's inequality gives $\sum_r \nu_{45}(r)r^{A_*} \leq (\sum_r \nu_{45}(r)r^5)^{A_*/5} \leq 20^{A_*}$. Summing the second step against η and grouping by $r = j(y)$ with (10.2) gives $\sum_y \eta(y)Q(y) \leq (8D_*)^{A_*} 20^{A_*} n^{-A_*}$. $\square$

15. ESTIMATES FOR THE FIRST-CROSSING FAMILIES

This section proves the facts about scales and first-crossing families used in Sections 5–8: the growth of scales and levels (Lemma 15.1), prefix-disjointness (Lemma 15.2), the lower bounds on products of masses of block families used in Lemma 5.3, the deficit bound in (6.1), and the bounds on the terminal deficit and mask failure used in Proposition 6.9.

15.1. Scales and levels. The finite computations below need the first scales exactly, the deficit bounds need geometric growth of the scales, and propagation needs the levels $k_j = \lfloor \delta_j/4 \rfloor$ to reach n_0 , where Proposition 4.1 applies.

Lemma 15.1 (Scales and levels).

- (1) For $13 \leq \delta \leq 255$ the set defining e_δ contains $\lceil \delta/100 \rceil$; hence e_δ is well defined and $9 \cdot 3^\delta \cdot 16^{e_\delta} < 4^{\delta+1}$. For every $\delta \geq 9$, $e_\delta \geq \max(1, \delta/100)$, and $e_\delta = \lceil \delta/100 \rceil \leq (\delta + 99)/100$ for $\delta \geq 256$.
- (2) The scales increase; $(\delta_0, \dots, \delta_6) = (9, \dots, 15)$, $\delta_j \leq 255$ for $j \leq 41$, $\delta_j \geq 256$ for $j \geq 42$, $\delta_{42} = 256$, $\delta_{444} = 16544$ and $2^{40} < \delta_{2254} = 1100658881854 < 2^{41}$.
- (3) The scales satisfy (5.1); $\delta_j > 190g^j$ for $j \geq 444$; and $\delta_j \geq 2^{40}j$ and $\delta_j \geq 8000j^2$ for $j \geq 20 \cdot 10^9$.
- (4) For $j \geq 444$, $k_j \geq 47g^j$; for $j \geq 6000000$, $k_j \geq \max(n_0, 47g^j)$.

Proof. Let $i = \lceil \delta/100 \rceil$, so $i \leq \delta$. The inequality $9 \cdot 3^\delta \cdot 16^i < 4^{\delta+1}$ is equivalent to $(4/3)^\delta > 36 \cdot 16^{i-1}$. If $i = 1$, then $(4/3)^\delta \geq (4/3)^{13} > 36$, by the integer comparison $4^{13} > 36 \cdot 3^{13}$. If $i \geq 2$, then $\delta \geq 100(i-1) + 1 \geq 13 + 88(i-1)$, and $(4/3)^{88} > 16$ gives the claim. Thus the maximum is over a nonempty set, all of whose elements are at least $\lceil \delta/100 \rceil \geq \max(1, \delta/100)$; for $\delta \leq 12$, $e_\delta = 1$.

Consequently $\delta_{m+1} \geq (1 + 1/100)\delta_m = g\delta_m$ for every m , and induction on j gives $g^j\delta_m \leq \delta_{m+j}$. The scales increase. Running the integer recurrence to $j = 2254$ gives the listed values and $\delta_{41} = 233$. With these values, the following are exact integer comparisons:

$$100^j\delta_j < 201 \cdot 101^j \quad (0 \leq j \leq 444), \quad 16544 \cdot 100^{444} < 201 \cdot 101^{444}, \quad 16544 \cdot 100^{444} > 190 \cdot 101^{444}.$$

The first proves $\delta_j < 201g^j$ for $j \leq 444$. For $j \geq 444$ we show $\delta_j + 100 < 201g^j$ by induction, the case $j = 444$ being the second comparison. Since $\delta_j \geq 256$,

$$\delta_{j+1} + 100 \leq \delta_j + \frac{\delta_j + 99}{100} + 100 < g(\delta_j + 100) < 201g^{j+1}.$$

Next $\delta_j \geq g^{j-444}\delta_{444} > 190g^j$ for $j \geq 444$, by the third comparison. Finally, for $j \geq 20 \cdot 10^9$, (5.1) and the binomial theorem give $\delta_j \geq 9(1 + 1/100)^j \geq 9\binom{j}{4} \cdot 100^{-4}$. Each of $j, \dots, j-3$ is at least $j/2$, so $\binom{j}{4} \geq j^4/384$ and $\delta_j \geq 9j^4/(384 \cdot 10^8)$. This is at least $2^{40}j$ because $9j^3 \geq 72 \cdot 10^{30} > 2^{40} \cdot 384 \cdot 10^8 = 42221246506598400000000$, and at least $8000j^2$ because $9j^2 \geq 8000 \cdot 384 \cdot 10^8$.

For $j \geq 444$, $k_j \geq \delta_j/4 - 1 > 47.5g^j - 1 \geq 47g^j$, because $g^j \geq g^{444} > 2$ (the integer comparison $101^{444} > 2 \cdot 100^{444}$). For $j \geq 6000000$, moreover $47g^j > g^j = 2^{j \log_2 g} > 2^{6000000/70} > n_0$, since $\log_2 g > 1/70$ (Section 5). $\square$

15.2. The line and the families. The transfer of Section 4 needs finite prefix-disjoint families; the first lemma shows that the first-crossing families are such.

Lemma 15.2 (Structure of the families). *Let $\delta \geq 9$, u and $R \geq 0$ be integers. The line is nondecreasing: $H_{\delta,u}(d) \leq H_{\delta,u}(d+1)$ for integers $d \geq 0$. The family $\text{FC}_u(\delta, R)$ is finite and prefix-disjoint, hence so is each of its subfamilies, in particular $\text{FC}(\delta, R) \subseteq \text{FC}_{r_\delta}(\delta, R)$.*

Proof. The line is nondecreasing since B_0 and $(d - \delta)_+$ are nondecreasing and $(\delta - d)_+$ is nonincreasing in d . Words of $\text{FC}_u(\delta, R)$ have length at most d_δ^{hi} and letters at most $\max_{d \leq d_\delta^{\text{hi}}} H_{\delta, u}(d) + R$, so the family is finite. Suppose $w \in \text{FC}_u(\delta, R)$ of length d is a proper prefix of $w' \in \text{FC}_u(\delta, R)$. Since $d_\delta^{\text{lo}} < d < |w'|$, the definition for w' gives $\deg(w) = \deg_d(w') < H_{\delta, u}(d)$, while the definition for w gives $\deg(w) \geq H_{\delta, u}(d)$. $\square$

All deficit estimates below rest on one decomposition of the random word $W = (a_1, a_2, \dots)$ of Section 3 and its partial degrees $\deg_i(W)$. For integers $\delta \geq 9$ and u , provided $\deg_{d_\delta^{\text{lo}}}(W) < H_{\delta, u}(d_\delta^{\text{lo}})$ and some d with $d_\delta^{\text{lo}} < d \leq d_\delta^{\text{hi}}$ satisfies $\deg_d(W) \geq H_{\delta, u}(d)$, call the least such d the first crossing $\hat{d}$ of W . Then for every $R \in \mathbb{N}$ and every length d , $W_{\leq d} \in \text{FC}_u(\delta, R)$ if and only if $d = \hat{d}$ and $\deg_{\hat{d}}(W) \leq H_{\delta, u}(\hat{d}) + R$.

Lemma 15.3 (Covering the deficit). *Let $\delta \geq 9$, $u, R \geq 0$ be integers, and let $d_\delta^{\text{lo}} < d' \leq d'' \leq d_\delta^{\text{hi}}$. Let $\mathcal{V}$ consist of the words of $\text{FC}_u(\delta, R)$ with length in $[d', d'']$. Then*

$$1 - \text{mass}(\mathcal{V}) \leq \mathbb{P}(\deg_i(W) \geq H_{\delta, u}(i) \text{ for some } d_\delta^{\text{lo}} \leq i < d') + \mathbb{P}(\deg_{d''}(W) < H_{\delta, u}(d'')) + 2^{-(R+1)}.$$

Proof. We follow [8, proof of (3.6)]. As $\mathcal{V}$ is prefix-disjoint, the events that W begins with $w \in \mathcal{V}$ are disjoint (Section 4), so $\text{mass}(\mathcal{V})$ is the probability that some prefix of W lies in $\mathcal{V}$. If neither of the first two events occurs, the first crossing exists and lies in $[d', d'']$, so W has a prefix in $\mathcal{V}$ unless the event $\mathcal{O}$ occurs that $\deg_{\hat{d}}(W) > H_{\delta, u}(\hat{d}) + R$. Fix $d_\delta^{\text{lo}} < d \leq d_\delta^{\text{hi}}$ and a word $(a_1, \dots, a_{d-1})$ whose partial degrees at lengths $d_\delta^{\text{lo}}, \dots, d-1$ lie below the line; its degree l satisfies $l < H_{\delta, u}(d-1) \leq H_{\delta, u}(d)$ by Lemma 15.2. Given that W begins with this word, $\hat{d} = d$ is the event $a_d \geq H_{\delta, u}(d) - l \geq 1$, and $\mathcal{O}$ adds $a_d \geq H_{\delta, u}(d) - l + R + 1$. For the geometric law, $\mathbb{P}(a_1 \geq c + R + 1) = 2^{-(R+1)}\mathbb{P}(a_1 \geq c)$ for $c \geq 1$. Summing over the disjoint pairs (d, prefix) gives $\mathbb{P}(\mathcal{O}) = 2^{-(R+1)}\mathbb{P}(\hat{d} \text{ exists}) \leq 2^{-(R+1)}$. $\square$

15.3. Tail bounds for partial degrees. The sub-Gaussian maximal bound serves the deficits at large scales (Lemma 15.6); the exact binomial form with Lévy's reflection serves the finite evaluation at the scales $256 \leq \delta < 16544$ (Lemma 15.8).

Lemma 15.4 (Maximal inequality). *Let $d \geq 1$ be an integer and $x > 0$ real with $4x/(9d) \leq 1/32$. Then*

$$\mathbb{P}\left(\max_{1 \leq i \leq d} (\deg_i(W) - 2i) \geq x\right) \leq e^{-2x^2/(9d)}, \quad \mathbb{P}(\deg_d(W) - 2d \leq -x) \leq e^{-2x^2/(9d)}.$$

Proof. Put $t = 4x/(9d) \leq 1/32$, so that $1 \leq M_{\text{geo}}(\pm t) \leq e^{9t^2/8}$ by Lemma 4.2. As $M_{\text{geo}}(t) \geq 1$, the process $(e^{t(\deg_i(W) - 2i)})_{i \leq d}$ is a nonnegative submartingale, and Doob's maximal inequality [2, Theorem 4.4.2] gives

$$e^{tx} \mathbb{P}\left(\max_{1 \leq i \leq d} (\deg_i(W) - 2i) \geq x\right) \leq \mathbb{E}e^{t(\deg_d(W) - 2d)} = M_{\text{geo}}(t)^d \leq e^{9dt^2/8},$$

and $-tx + 9dt^2/8 = -2x^2/(9d)$. The second bound is the exponential Markov inequality, as in Lemma 4.2, with $M_{\text{geo}}(-t)^d \leq e^{9dt^2/8}$. $\square$

Lemma 15.5 (Binomial form and reflection). *For integers $d, N \geq 1$,*

$$\mathbb{P}(\deg_d(W) \geq N) = 2^{-(N-1)} \sum_{c=0}^{d-1} \binom{N-1}{c}.$$

For an integer $d \geq 1$ and real x ,

$$\mathbb{P}\left(\max_{1 \leq i \leq d} (\deg_i(W) - 2i) \geq x\right) \leq 2\mathbb{P}(\deg_d(W) - 2d \geq x).$$

Proof. The first claim is the duality between the negative binomial and binomial laws [3, Section VI.8]: the letters are the waiting times between successes in independent fair Bernoulli trials, and $\deg_d(W) \geq N$ if and only if the first $N - 1$ trials show fewer than d successes. By the symmetry of $\binom{2d'-1}{c}$ it gives $\mathbb{P}(\deg_{d'}(W) \geq 2d') = \frac{1}{2}$ for $d' \geq 1$, so 0 is a median of $\deg_{d'}(W) - 2d'$. The variables $a_i - 2$ are independent, and for $i < d$ the difference $(\deg_d(W) - 2d) - (\deg_i(W) - 2i)$ has the law of $\deg_{d-i}(W) - 2(d-i)$, hence median 0. Lévy's inequality [6, Section 18.1] therefore gives the second claim. $\square$

15.4. Deficits of first-crossing families. Lemma 15.3 is applied in four regimes: large scales (Lemmas 15.6 and 15.7), the scales $256 \leq \delta < 16544$ (Lemma 15.8), the scales $\delta \leq 255$, where the masses are computed exactly by counting (Lemma 15.9), and every shift $0 \leq u \leq 2r_\delta$, as the terminal extension needs (Lemma 15.10).

Lemma 15.6 (Central length ranges). *Let $\delta, R \geq 0$ and y be integers with $1000 \leq y \leq \delta/8$, and let $\mathcal{V}_y$ consist of the words of $\text{FC}_{r_\delta}(\delta, R)$ with length in $[\delta - y, \delta + y]$. Then*

$$1 - \text{mass}(\mathcal{V}_y) \leq 2 \exp\left(-\frac{2(207y/500)^2}{9(\delta + y)}\right) + 2^{-(R+1)}.$$

Proof. Put $x = 207y/500$. Since $y \leq \delta/8 < 3\delta/5 - 1 \leq \ell_\delta$, Lemma 15.3 applies with $d' = \delta - y$, $d'' = \delta + y$. The shift cancels: $H_{\delta, r_\delta}(i) = 2\delta + B_0((i - \delta)_+) - B_0((\delta - i)_+)$. If $d_\delta^0 \leq i < \delta - y$ and $\deg_i(W) \geq H_{\delta, r_\delta}(i)$, then

$$\deg_i(W) - 2i \geq 2(\delta - i) - B_0(\delta - i) > (2 - \log_2 3)(\delta - i) - 1 > \frac{83}{200}y - 1 \geq x,$$

by $B_0(n) < n \log_2 3 + 1$ (Section 5), $2 - \log_2 3 > \frac{83}{200}$ (Lemma A.1(1)) and $(\frac{83}{200} - \frac{207}{500})y = y/1000 \geq 1$. If $\deg_{\delta+y}(W) < H_{\delta, r_\delta}(\delta + y) = 2\delta + B_0(y)$, then $\deg_{\delta+y}(W) - 2(\delta + y) < B_0(y) - 2y < 1 - (2 - \log_2 3)y < -x$. Also $4x/(9(\delta + y)) \leq \frac{4 \cdot 207}{9 \cdot 500 \cdot 8} < \frac{1}{32}$. Lemma 15.4 with $d = \delta + y$ bounds the two events of Lemma 15.3 by $e^{-2x^2/(9(\delta+y))}$ each. $\square$

Lemma 15.7 (Deficit of the block family at large scales). *Let $\delta \geq 16384$ and $R \geq 0$ be integers. Then*

$$1 - \text{mass}(\text{FC}(\delta, R)) \leq 2 \exp(-\min(\lceil \log_2 \delta \rceil^2, 32 \lceil \log_2 \delta \rceil)/30) + 2^{-(R+1)}.$$

Proof. For $\delta \geq 256$, $\text{FC}(\delta, R) = \mathcal{V}_{\ell_\delta^c}$ in the notation of Lemma 15.6, and $\mathcal{V}_y \subseteq \mathcal{V}_{y'}$ for $y \leq y'$. Note $\frac{16}{81}(\frac{207}{500})^2 = \frac{529}{15625} > \frac{1}{30}$. We choose $1000 \leq y \leq \min(\delta/8, \ell_\delta^c)$ with $y^2 \geq \delta \min(\lceil \log_2 \delta \rceil^2, 32 \lceil \log_2 \delta \rceil)$; then $\delta + y \leq 9\delta/8$, and Lemma 15.6 gives the bound $2e^{-\frac{529}{15625} \min(\lceil \log_2 \delta \rceil^2, 32 \lceil \log_2 \delta \rceil)} + 2^{-(R+1)}$, which suffices.

Case $\lceil \log_2 \delta \rceil \leq 32$. Take $y = \lceil \sqrt{\delta} \lceil \log_2 \delta \rceil \rceil$. Since $\log_2 \delta \geq 14$, $\lceil \log_2 \delta \rceil \leq \log_2 \delta + 1$ and $(\log_2 x + 1)/\sqrt{x}$ decreases for real $x \geq 4$, $\lceil \log_2 \delta \rceil / \sqrt{\delta} \leq 15/2^7$, so $y \leq \frac{15}{128}\delta + 1 \leq \delta/8$. Also $y \geq 128 \cdot 14 > 1000$, and $y \leq \ell_\delta^c$ because $\sqrt{\delta} \lceil \log_2 \delta \rceil \leq \sqrt{32\delta \lceil \log_2 \delta \rceil}$ and $y \leq \delta/8 \leq \ell_\delta - 1$.

Case $\lceil \log_2 \delta \rceil \geq 33$. Take $y = \ell_\delta^c$. As $\log_2 \delta \geq 32$, $2^{\log_2 \delta - 13} \geq 2^{19}(1 + (\log_2 \delta - 32)/2) \geq \log_2 \delta + 1 > \lceil \log_2 \delta \rceil$, so $\delta > 8192 \lceil \log_2 \delta \rceil$ and $\sqrt{32\delta \lceil \log_2 \delta \rceil} < \delta/16$. Hence $\lceil \sqrt{32\delta \lceil \log_2 \delta \rceil} \rceil \leq \delta/16 + 1 \leq \ell_\delta - 1$, so $\ell_\delta^c = \lceil \sqrt{32\delta \lceil \log_2 \delta \rceil} \rceil \leq \delta/8$, $(\ell_\delta^c)^2 \geq 32\delta \lceil \log_2 \delta \rceil$, and $\ell_\delta^c \geq \sqrt{32 \cdot 2^{32} \cdot 33} > 1000$. $\square$

Lemma 15.8 (Binomial lower bound). *Let $\delta \geq 256$ and $R \geq 0$ be integers, $y = \ell_\delta^c$, $\bar{y} = \lfloor 83y/200 \rfloor$, $d' = \delta - y$, $d'' = \delta + y$, $N' = 2d' + \bar{y} - 1$ and $N'' = 2d'' - \bar{y} - 1$. Then*

$$\text{mass}(\text{FC}(\delta, R)) \geq 1 - \frac{2}{2^{N'}} \sum_{c=0}^{d'-1} \binom{N'}{c} - \frac{1}{2^{N''}} \sum_{c=0}^{d''-\bar{y}-1} \binom{N''}{c} - 2^{-(R+1)}.$$

Proof. Apply Lemma 15.3 with these d', d'' ; this is legitimate since $1 \leq y \leq \ell_\delta - 1 < \ell_\delta$. If $d_\delta^{\text{lo}} \leq i < d'$ and $\deg_i(W) \geq H_{\delta, r_\delta}(i)$, the integer $\deg_i(W) - 2i$ is at least $2(\delta - i) - B_0(\delta - i) > (2 - \log_2 3)(y + 1) - 1 > 83y/200 - 1$, hence at least $\bar{y}$. Lemma 15.5 bounds the first event of Lemma 15.3 by $2\mathbb{P}(\deg_{d'}(W) \geq N' + 1)$, which is the first sum. If $\deg_{d''}(W) < H_{\delta, r_\delta}(d'')$, the integer $\deg_{d''}(W) - 2d'' \leq B_0(y) - 2y - 1$ is less than $-(2 - \log_2 3)y < -83y/200$, hence at most $-(\bar{y} + 1)$, that is, $\deg_{d''}(W) \leq N''$. By the first claim of Lemma 15.5 and the symmetry of binomial coefficients, this has probability $1 - 2^{-N''} \sum_{c < d''} \binom{N''}{c} = 2^{-N''} \sum_{c \leq N'' - d''} \binom{N''}{c}$, and $N'' - d'' = d'' - \bar{y} - 1$. $\square$

For $9 \leq \delta \leq 255$ define counts $\mathcal{N}_i(l)$ for $d_\delta^{\text{lo}} \leq i < d_\delta^{\text{hi}}$ and $l \in \mathbb{N}$ by

$$\begin{aligned} \mathcal{N}_{d_\delta^{\text{lo}}}(l) &= \binom{l-1}{d_\delta^{\text{lo}}-1} \mathbf{1}[d_\delta^{\text{lo}} \leq l < H_{\delta, r_\delta}(d_\delta^{\text{lo}})], \\ \mathcal{N}_i(l) &= \mathbf{1}[l < H_{\delta, r_\delta}(i)] \sum_{l' < l} \mathcal{N}_{i-1}(l') \quad (i > d_\delta^{\text{lo}}). \end{aligned} \tag{15.1}$$

By induction, $\mathcal{N}_i(l)$ is the number of words of length i and degree l whose partial degrees at lengths $d_\delta^{\text{lo}}, \dots, i$ lie strictly below the line H_{δ, r_δ} : there are $\binom{l-1}{d_\delta^{\text{lo}}-1}$ words of length d_δ^{lo} and degree $l \geq d_\delta^{\text{lo}}$, and a word of length i is a word of length $i-1$ followed by one positive letter.

Lemma 15.9 (Exact masses at small scales). *For $9 \leq \delta \leq 255$ and an integer $R \geq 1$,*

$$\text{mass}(\text{FC}(\delta, R)) = \sum_{d=d_\delta^{\text{lo}}+1}^{d_\delta^{\text{hi}}} \left(\sum_l \mathcal{N}_{d-1}(l) \right) \frac{2^{R+1-o_\delta(d)} - 1}{2^{H_{\delta, r_\delta}(d)+R}}.$$

Proof. For $\delta < 256$ the family $\text{FC}(\delta, R)$ has no length condition beyond that of $\text{FC}_{r_\delta}(\delta, R)$. A word of $\text{FC}(\delta, R)$ of length d is a word of length $d-1$ and degree l counted by $\mathcal{N}_{d-1}(l)$, followed by a letter a with $H_{\delta, r_\delta}(d) + o_\delta(d) \leq l + a \leq H_{\delta, r_\delta}(d) + R$. Since $l < H_{\delta, r_\delta}(d-1) \leq H_{\delta, r_\delta}(d)$, each total c in this range arises from exactly one positive a , so every such word of length $d-1$ contributes $\sum_{c=H_{\delta, r_\delta}(d)+o_\delta(d)}^{H_{\delta, r_\delta}(d)+R} 2^{-c}$. The values of o_δ satisfy $o_\delta(d) \leq 2 \leq R+1$ (Appendix A.4 lists them). $\square$

Lemma 15.10 (First-crossing families and the mask). *Let $\delta \geq 9$, $R \geq 0$ and $0 \leq u \leq 2r_\delta$ be integers. Then*

$$\begin{aligned} 1 - \text{mass}(\text{FC}_u(\delta, R)) &\leq 2^{116}\delta^{-6} + 2^{-(R+1)}, \\ \text{mass}(\{w \in \text{FC}_u(\delta, R) : 2|w| > H_{\delta, u}(|w|) + \lfloor \delta/8 \rfloor\}) &\leq 2^{113}\delta^{-6}; \end{aligned}$$

the words of the second set are those failing the mask (5.7) at scale δ and shift u .

Proof. We follow [8, proofs of Proposition 5.2 and Lemma 5.1], with smaller constants. Apply Lemma 15.3 with $d' = d_\delta^{\text{lo}} + 1$, $d'' = d_\delta^{\text{hi}}$. From $\delta - d_\delta^{\text{lo}} = d_\delta^{\text{hi}} - \delta = \ell_\delta$, the definition of r_δ and $0 \leq u \leq 2r_\delta$,

$$H_{\delta, u}(d_\delta^{\text{lo}}) = 2\delta - B_0(\ell_\delta) + u - r_\delta \geq 2d_\delta^{\text{lo}} + 2e_\delta, \quad H_{\delta, u}(d_\delta^{\text{hi}}) = 2\delta + B_0(\ell_\delta) + u - r_\delta \leq 2d_\delta^{\text{hi}} - 2e_\delta.$$

So both events of Lemma 15.3 lie in $\{|\deg_i(W) - 2i| \geq 2e_\delta\}$ for $i = d_\delta^{\text{lo}}$ and $i = d_\delta^{\text{hi}}$, respectively. As $e_\delta \geq \delta/100$ and $i \leq 2\delta$, the exponent in (4.8) (Lemma 4.2) is at least $\delta/160000$. The bound $e^{-\delta/160000} \leq 6!(\delta/160000)^{-6}$ then gives $1 - \text{mass}(\text{FC}_u(\delta, R)) \leq 4 \cdot 6! \cdot 160000^6 \delta^{-6} + 2^{-(R+1)}$, with $4 \cdot 6! \cdot 160000^6 < 2^{116}$.

For the mask, let $\mathcal{V}$ be the set of failing words; it is prefix-disjoint. Since $2^{113} > 13^6$, assume $\delta \geq 14$. If $w \in \mathcal{V}$ has length d , its prefix of length $d-1 \geq d_\delta^{\text{lo}}$ has degree at most $H_{\delta, u}(d-1) - 1 \leq H_{\delta, u}(d) - 1 \leq 2d - \lfloor \delta/8 \rfloor - 2$, so $\deg_{d-1}(W) - 2(d-1) \leq -\lfloor \delta/8 \rfloor \leq -\delta/16$ if W begins with w . Hence $\text{mass}(\mathcal{V}) \leq \sum_{i=d_\delta^{\text{lo}}}^{d_\delta^{\text{hi}}-1} \mathbb{P}(|\deg_i(W) - 2i| \geq \delta/16)$. There are at most 2δ terms, each at most $2e^{-\delta/16384}$ by (4.8), and $e^{-\delta/16384} \leq 7!(\delta/16384)^{-7}$ gives $\text{mass}(\mathcal{V}) \leq 4 \cdot 7! \cdot 16384^7 \delta^{-6} < 2^{113}\delta^{-6}$. $\square$

15.5. Deficits along the scale sequence. At the scales δ_j these bounds give the decay of the block deficits used for propagation, the lower bounds on products of block masses used in Lemma 5.3, and the terminal bounds used in Proposition 6.9.

Lemma 15.11 (Deficit of the block families). *For every $j \in \mathbb{N}$, $1 - \text{mass}(\text{FC}(\delta_j, R_j)) \leq (1024 + 2^{-16})g^{-3j/2}$.*

Proof. First, $2^{-(R_j+1)} \leq 2^{-16}2^{-j/32} \leq 2^{-16}g^{-3j/2}$, because $g^{48} < 2$ (the integer comparison $101^{48} < 2 \cdot 100^{48}$). Next, $g^{666} \leq 2^{10}$, since $101^{666} \leq 2^{10} \cdot 100^{666}$. For $j < 444$ the claim follows from $1 - \text{mass}(\text{FC}(\delta_j, R_j)) \leq 1 \leq 1024g^{-3j/2}$.

Let $j = 444 + m$ with $m \geq 0$. By Lemma 15.1 and $101^{70} > 2 \cdot 100^{70}$, $\delta_j \geq 16544g^m \geq 16544 \cdot 2^{\lfloor m/70 \rfloor}$, so $x = \lceil \log_2 \delta_j \rceil \geq 15 + \lfloor m/70 \rfloor > y = 14 + m/70$. Then $\min(x^2, 32x)/30 \geq \min(y^2, 32y)/30 \geq 6 + \frac{3m}{200}$: for $y \leq 32$ this is $y^2 - \frac{63}{2}y + 261 \geq 0$, a quadratic with negative discriminant (as $\frac{3m}{200} = \frac{21}{20}(y - 14)$), and for $y \geq 32$ it follows from $\frac{32}{30} > \frac{21}{20}$ and $32 \cdot 14/30 > 6$. So Lemma 15.7 gives

$$1 - \text{mass}(\text{FC}(\delta_j, R_j)) \leq 2e^{-6}e^{-3m/200} + 2^{-(R_j+1)}.$$

Here $2e^{-6} < 1$, and $\log g < 1/100$ gives $e^{-3m/200} \leq g^{-3m/2} = g^{666}g^{-3j/2} \leq 1024g^{-3j/2}$. $\square$

Lemma 15.12 (Startup product). $\prod_{j=0}^{443} \text{mass}(\text{FC}(\delta_j, R_j)) > \frac{72}{5} 2^{-25}$.

Proof. By Lemma 15.1, $9 \leq \delta_j \leq 255$ for $j \leq 41$ and $256 \leq \delta_j < 16544$ for $42 \leq j \leq 443$. All factors lie in $[0, 1]$ by Lemma 15.2, so it suffices to bound each factor below by a nonnegative rational. The following finite computation uses exact integer and rational arithmetic, with δ_j and e_δ computed exactly from their definitions (Lemma 15.1) and $B_0(n) = \text{bitlen}(3^n)$ for $n \geq 1$ (Section 5).

- (1) For $j \leq 41$, evaluate the right side of Lemma 15.9 at $\delta = \delta_j$, $R = R_j$, computing the counts $\mathcal{N}_i(l)$ by their recurrence and, for $\delta \leq 11$, each $o_\delta(d)$ as the least $o \geq 0$ satisfying the integer inequality of Section 5. This is $\text{mass}(\text{FC}(\delta_j, R_j))$ exactly; for $j \leq 4$ these are the masses $\text{mass}(\text{FC}(\delta_j, 16))$ tabulated in Appendix A.4.
- (2) For $42 \leq j \leq 443$, evaluate the right side of Lemma 15.8 at $\delta = \delta_j$, $R = R_j$, with the binomial sums as exact integers, $\lceil \log_2 \delta \rceil = \text{bitlen}(\delta - 1)$, and $\left\lceil \sqrt{32\delta \lceil \log_2 \delta \rceil} \right\rceil = \left\lceil \sqrt{32\delta \lceil \log_2 \delta \rceil - 1} \right\rceil + 1$ (an integer square root). Rounding it down to a multiple of 2^{-64} gives a rational $\text{mass}(\text{FC}(\delta_j, R_j))^{\text{lo}} \leq \text{mass}(\text{FC}(\delta_j, R_j))$.
- (3) Multiply the 42 exact masses of step 1 and the 402 numbers $\text{mass}(\text{FC}(\delta_j, R_j))^{\text{lo}}$, and compare $5 \cdot 2^{25}$ times this product with 72.

All these factors lie in $(0, 1)$ (the least $\text{mass}(\text{FC}(\delta_j, R_j))^{\text{lo}}$ is about 0.98866). The masses of step 1 have product $\approx 5.7117 \cdot 10^{-7}$, the $\text{mass}(\text{FC}(\delta_j, R_j))^{\text{lo}}$ have product ≈ 0.76732 , and 2^{25} times the product of all 444 factors lies in $(14.7058, 14.7059)$; in particular it exceeds $72/5$. $\square$

Lemma 15.13 (Late product). *For every integer $j_1 \geq 444$, $\prod_{j=444}^{j_1-1} \text{mass}(\text{FC}(\delta_j, R_j)) > 5/9$.*

Proof. All factors lie in $[0, 1]$, so the product is at least $1 - \sum_{j=444}^{j_1-1} (1 - \text{mass}(\text{FC}(\delta_j, R_j)))$. Let $j = 444 + m$ and $i = \lfloor m/71 \rfloor$. By Lemma 15.1 and $101^{71} > 2 \cdot 100^{71}$, $\delta_j \geq 16544 \cdot 2^i > 2^{14+i}$, so $\lceil \log_2 \delta_j \rceil \geq 14 + i$, and both $(14 + i)^2$ and $32(14 + i)$ are at least $196 + 29i$. By Lemma 15.7 and $e^{98/15} > 650$, $e^{29/30} > 2$ (Lemma A.1(4) with $k = 11$ and $k = 2$),

$$1 - \text{mass}(\text{FC}(\delta_j, R_j)) \leq \frac{2}{650} 2^{-i} + 2^{-(R_j+1)}.$$

Each i occurs for 71 values of m , so the first terms sum to at most $\frac{142}{650} \cdot 2$. Since $R_{444+m} + 1 \geq 30 + \lfloor m/32 \rfloor$, the cap terms sum to at most $32 \cdot 2^{-30} \cdot 2 < 1/2048$. Hence the product exceeds $1 - \frac{284}{650} - \frac{1}{2048} = \frac{374459}{665600} > \frac{5}{9}$. $\square$

Proposition 15.14 (Terminal deficit and mask failure). *For every $j \in \mathbb{N}$ and every integer $0 \leq u \leq 2r_{\delta_j}$,*

$$1 - \text{mass}(\text{FC}_u(\delta_j, R_j)) \leq (2^{121} + 2^{-16})g^{-3j/2},$$

$$\text{mass}(\{v \in \text{FC}_u(\delta_j, R_j) : v \text{ fails (5.7)}\}) \leq 2^{95}g^{-3j/2},$$

where the mask is taken at scale δ_j and shift u , as in Lemma 15.10.

Proof. By (5.1), $\delta_j \geq 9g^j$, so $\delta_j^{-6} \leq 9^{-6}g^{-6j} \leq 9^{-6}g^{-3j/2}$, and $9^6 > 2^{19}$. By Lemma 15.10, the two quantities are at most $2^{97}g^{-3j/2} + 2^{-(R_j+1)}$ and $2^{94}g^{-3j/2}$. Bound the cap term as in the proof of Lemma 15.11. $\square$

16. SOURCES AND THE HISTOGRAM ESTIMATE

This section proves the spread of the sources (Lemma 16.1), used in Lemmas 6.8 and 6.6, and the histogram estimate (5.6) (Proposition 16.5) used in Section 6. The arguments are those of [8, proof of Proposition 3.1], made explicit and extended to the scales $9 \leq \delta < 256$.

16.1. Sources.

Lemma 16.1 (Spread of the sources). *Let $t \geq 16^9$ be odd and $j \in \mathbb{N}$. For all $w, w' \in L_j$,*

$$s_{w'} \leq 2^{R_\Sigma} s_w.$$

Proof. Multiplying the bounds of Lemma 5.4 over $i < j$, starting from $s_{w_{<0}} = s_{w'_{<0}} = t$, gives $s_w \geq t \prod_{i < j} \frac{1}{4}(4/3)^{\delta_i}$ and $s_{w'} \leq t \prod_{i < j} 2^{R_i+1}(4/3)^{\delta_i}$. Divide. $\square$

16.2. Offsets of selected words. The five-block selector pins down the offset of every selected word of at least five blocks; this places the sources of the words of one length and degree in a short interval (Lemma 16.3).

Lemma 16.2 (Offsets of selected words). *Let $j \geq 5$ and $w = v_0 \cdots v_{j-1} \in \bar{L}_j$. Then*

$$\frac{85}{256} \leq \text{off}(w) \leq \frac{637}{256}.$$

Proof. For every word v and integer $\mathfrak{r}$, rounding each of the $|v|$ summands up to a multiple of $2^{-\mathfrak{r}}$ gives

$$\text{rd}_{\mathfrak{r}}(v) - |v|2^{-\mathfrak{r}} \leq \text{off}(v) \leq \text{rd}_{\mathfrak{r}}(v). \quad (16.1)$$

By Lemma 15.1, $(\delta_0, \dots, \delta_5) = (9, \dots, 14)$, and Lemma 5.2(3) gives $0 < \text{wt}(v_i) < 1/16$ for $i < 5$. Moreover $|v_0| \leq d_9^{\text{hi}} = 14$ and $|v_1| \leq d_{10}^{\text{hi}} = 16$. For $5 \leq i \leq j$ put $\mathfrak{o}_i = \text{off}(v_i \cdots v_{j-1})$, so $\mathfrak{o}_j = 0$ and $\mathfrak{o}_i = \text{off}(v_i) + \text{wt}(v_i)\mathfrak{o}_{i+1} \geq 0$ for $i < j$ by (3.2). If $\mathfrak{o}_{i+1} \leq 2^{\delta_{i+1}+1}$, then Lemma 5.2(2),(3), with $e_{\delta_i} \geq 1$ and $\delta_{i+1} = \delta_i + e_{\delta_i}$, give

$$\mathfrak{o}_i \leq 2^{\delta_i} + 16^{-e_{\delta_i}} \cdot 2^{\delta_i + e_{\delta_i} + 1} = 2^{\delta_i} + 2^{\delta_i + 1 - 3e_{\delta_i}} \leq 2^{\delta_i + 1}.$$

Downward induction from $i = j$, as in [8, proof of Proposition 3.1], gives $0 \leq \text{off}(v_5 \cdots v_{j-1}) \leq 2^{\delta_5+1} = 2^{15}$. Applying (3.2) repeatedly,

$$\begin{aligned} \text{off}(w) &= \text{off}(v_0) + \text{wt}(v_0) \text{off}(v_1) \\ &\quad + \text{wt}(v_0 v_1) \left[\text{off}(v_2) + \text{wt}(v_2) (\text{off}(v_3) + \text{wt}(v_3) (\text{off}(v_4) + \text{wt}(v_4) \text{off}(v_5 \cdots v_{j-1}))) \right], \end{aligned}$$

where all offsets are nonnegative and all weights positive.

For the lower bound, drop the bracket and apply (16.1) with $\mathfrak{r} = 9$ and $\mathfrak{r} = 6$:

$$\text{off}(w) \geq \text{rd}_9(v_0) + \text{wt}(v_0) \text{rd}_6(v_1) - \frac{14}{512} - \frac{\text{wt}(v_0) \cdot 16}{64} \geq \frac{3}{8} - \frac{14}{512} - \frac{1}{64} = \frac{85}{256}.$$

For the upper bound, (16.1) and the five-block selector give $\text{off}(v_0) + \text{wt}(v_0) \text{off}(v_1) \leq 19/8$, $\text{off}(v_2) \leq 16$, $\text{off}(v_3) \leq 64$ and $\text{off}(v_4) \leq 256$. With the weight bounds, the bracketed term multiplied by $\text{wt}(v_0 v_1)$ is at most

$$\frac{16}{16^2} + \frac{64}{16^3} + \frac{256}{16^4} + \frac{2^{15}}{16^5} = \frac{29}{256},$$

so $\text{off}(w) \leq 19/8 + 29/256 = 637/256$. $\square$

16.3. The histogram estimate. The histogram estimate (5.6) follows from a uniform bound on the *histogram* of $L_j = L_j(t)$ at level q , defined for a positive odd integer t , integers $j, q \geq 0$ and $y \in G_q$ by

$$\Xi_{j,q}(t; y) = \sum_{\substack{w \in L_j \\ [s_w]_q = y}} \text{wt}(w).$$

It is bounded by splitting according to the length and the degree of w . Words with the same pair share their weight, their sources are distinct, and by Lemma 16.2 those sources lie in a short interval.

Lemma 16.3 (Words of one length and degree). *Let t be a positive odd integer, $j \geq 42$, $0 \leq q \leq 2\delta_j$, $y \in G_q$, and $d, l \in \mathbb{Z}$. Then*

$$3^q \sum_{\substack{w \in L_j, |w|=d, \deg(w)=l \\ [s_w]_q = y}} \text{wt}(w) < \frac{277}{128}.$$

Proof. All words w in the sum have the same weight $\text{wt}(w) = 3^d \cdot 2^{-l}$, and $t = \text{wt}(w)s_w + \text{off}(w)$ by (3.1). Since $j \geq 5$, Lemma 16.2 places $s_w = (t - \text{off}(w))/\text{wt}(w)$ in a closed interval of length $\frac{637-85}{256} \text{wt}(w)^{-1} = \frac{69}{32} \text{wt}(w)^{-1}$. A word valid at t is determined by its length and its source (proof of Lemma 3.3), so the sources in the sum are distinct integers congruent to y modulo 3^q in this interval; there are at most $\frac{69}{32} \text{wt}(w)^{-1} 3^{-q} + 1$ of them. The left side is therefore at most $\frac{69}{32} + 3^q \text{wt}(w)$.

By Lemma 5.2(3), $\text{wt}(v_i) < 16^{-e_{\delta_i}}$ for every block v_i of w , and $\sum_{i < j} e_{\delta_i} = \delta_j - \delta_0$. Thus $\text{wt}(w) < 16^{9-\delta_j}$, and with $3^q \leq 9^{\delta_j}$,

$$3^q \text{wt}(w) < 2^{36} (9/16)^{\delta_j} \leq 2^{36} (9/16)^{256} < 2^{-7},$$

using $\delta_j \geq \delta_{42} = 256$ and the exact integer inequality $2^{43} \cdot 9^{256} < 16^{256}$. Since $\frac{69}{32} + \frac{1}{128} = \frac{277}{128}$, the lemma follows. $\square$

Lemma 16.4 (Lengths and degrees of words). *Let t be a positive odd integer and $j \geq 10000$. The pairs $(|w|, \deg(w))$ with $w \in L_j$ take fewer than $\frac{4075}{56} j^{5/2} g^{j/2}$ values. More precisely, the lengths $|w|$ take fewer than $4075\sqrt{j} g^{j/2}$ values, and for each length the degrees take fewer than $j^2/56$ values.*

Proof. First, the lengths. Every $v \in \text{FC}(\delta, R)$ satisfies $\delta - \ell_{\delta}^c \leq |v| \leq \delta + \ell_{\delta}^c$: for $\delta \geq 256$ this is the length restriction, and for $\delta < 256$ it is $d_{\delta}^{\text{lo}} < |v| \leq d_{\delta}^{\text{hi}}$ with $\ell_{\delta}^c = \ell_{\delta}$. With $\ell_{\Sigma}^c = \sum_{i < j} \ell_{\delta_i}^c$ as in Section 8, the length $|w|$ is one of the $2\ell_{\Sigma}^c + 1$ integers within ℓ_{Σ}^c of $\sum_{i < j} \delta_i$. We claim $\ell_{\delta}^c \leq \lceil \sqrt{32\delta} \lceil \log_2 \delta \rceil \rceil$ for every $\delta \geq 9$. For $\delta \geq 256$ this is part of the definition of ℓ_{δ}^c . For $9 \leq \delta \leq 255$, $\lceil \log_2 \delta \rceil \geq 4$ and $\ell_{\delta}^2 \leq \frac{9 \cdot 255}{25} \delta < 128\delta \leq 32\delta \lceil \log_2 \delta \rceil$.

By (5.1), $\delta_i < 201g^i$. The exact inequalities $201 < 2^8$ and $101^{69} < 2 \cdot 100^{69}$ give $\log_2 \delta_i < 8 + i/69$, so $\lceil \log_2 \delta_i \rceil < 9 + i/69$. Since $32 \cdot 201 < 81^2$,

$$\ell_{\delta_i}^c < \sqrt{32 \cdot 201 g^i (9 + i/69)} + 1 < 81 g^{i/2} \sqrt{9 + j/69} + 1 \quad (i < j).$$

The exact inequality $(202/201)^2 < 101/100$ gives $g^{1/2} - 1 > 1/201$, hence $\sum_{i < j} g^{i/2} < 201g^{j/2}$. For $j \geq 10000$, $9/j + 1/69 < 1/64$ (exact), so $\sqrt{9 + j/69} < \sqrt{j}/8$. Moreover $g^j \geq \binom{j}{3} \cdot 100^{-3} \geq j$ by the

binomial theorem, because $(j-1)(j-2) \geq 6 \cdot 10^6$; thus $j \leq \sqrt{j} g^{j/2}$. Therefore

$$\ell_\Sigma^c < \left(\frac{81 \cdot 201}{8} + 1 \right) \sqrt{j} g^{j/2} < 2037 \sqrt{j} g^{j/2}, \quad 2\ell_\Sigma^c + 1 < 4075 \sqrt{j} g^{j/2},$$

using $\sqrt{j} g^{j/2} \geq 1$ in the last step.

Next, the degrees at a fixed length d . For a word $v \in \text{FC}(\delta, R)$, the inequalities $H_{\delta, r_\delta}(|v|) \leq \deg(v) \leq H_{\delta, r_\delta}(|v|) + R$ and (5.2) give

$$2\delta + (|v| - \delta) \log_2 3 - 1 < \deg(v) < 2\delta + (|v| - \delta) \log_2 3 + 1 + R.$$

Summing over the blocks $v_i \in \text{FC}(\delta_i, R_i)$ of w , with $\sum_i |v_i| = d$,

$$\begin{aligned} 2 \sum_{i < j} \delta_i + \left(d - \sum_{i < j} \delta_i \right) \log_2 3 - j &< \deg(w) \\ &< 2 \sum_{i < j} \delta_i + \left(d - \sum_{i < j} \delta_i \right) \log_2 3 + j + \sum_{i < j} R_i. \end{aligned}$$

This open interval has integer length $2j + \sum_{i < j} R_i < R_\Sigma + 1$, so it contains fewer than $j^2/56$ integers by (6.6). $\square$

Proposition 16.5 (Histogram estimate). *Let t be a positive odd integer, $j \geq 10000$ and $0 \leq q \leq 2\delta_j$. Then*

$$3^q \Xi_{j,q}(t; y) < \frac{4075}{56} \frac{277}{128} j^{5/2} g^{j/2} \quad (y \in G_q).$$

Consequently (5.6) holds for every $f : G_q \rightarrow \mathbb{R}$ and every subset of L_j .

Proof. Split the sum defining $\Xi_{j,q}(t; y)$ according to the pair $(|w|, \deg(w))$. By Lemma 16.4 fewer than $\frac{4075}{56} j^{5/2} g^{j/2}$ pairs occur, and by Lemma 16.3 (as $j \geq 42$) each contributes less than $\frac{277}{128} 3^{-q}$. For (5.6), group a subset of L_j by the residue $y = [s_w]_q$; since weights are nonnegative,

$$\sum_w \text{wt}(w) |f(s_w)| \leq \sum_{y \in G_q} \Xi_{j,q}(t; y) |f(y)| \leq \max_y (3^q \Xi_{j,q}(t; y)) \langle |f| \rangle_q.$$

$\square$

APPENDIX A. NUMERICAL CONSTANTS AND FINITE CHECKS

This appendix defines the constants of the Fourier argument and records the finite computations behind the numerical claims. Apart from the standard values of $\log_2 3$, $\log 2$, $\log(4/3)$, e , e^9 and e^{10} stated in Lemma A.1, every claim is an exact computation: finite decimals denote exact rational numbers, the integers involved have at most tens of thousands of bits and are computed exactly, a ceiling $\lceil x/y \rceil$ of a positive rational is evaluated as $\lfloor (x+y-1)/y \rfloor$, and bit lengths by comparison with powers of 2.

Lemma A.1 (Elementary bounds).

- (1) $79/50 < \log_2 3 < 317/200$.
- (2) For real $y \geq 1$ and integer $N \geq 1$, with $r = (y-1)/(y+1)$,

$$2 \sum_{j=0}^{N-1} \frac{r^{2j+1}}{2j+1} \leq \log y \leq 2 \sum_{j=0}^{N-1} \frac{r^{2j+1}}{2j+1} + \frac{2r^{2N+1}}{(2N+1)(1-r^2)}. \quad (\text{A.1})$$

- (3) $693147180/10^9 < \log 2 < 347/500$ and $\log(4/3) > 287682/10^6$.
- (4) For real $x \geq 0$ and $k \in \mathbb{N}$, $\sum_{j=0}^k x^j/j! \leq e^x$. In particular $e^9 > 8000$ and $e^{10} > 10^4$.
- (5) $e < 68/25$.

Proof. Parts (1), (3) and (5), and the two numerical claims in (4), follow from the standard values $\log_2 3 = 1.58496\dots$, $\log 2 = 0.6931471805\dots$, $\log(4/3) = 0.2876820724\dots$, $e = 2.7182818\dots$, $e^9 = 8103.08\dots$ and $e^{10} = 22026.46\dots$. Part (2) is the series $\log y = 2 \sum_{j \geq 0} r^{2j+1}/(2j+1)$, valid as $0 \leq r < 1$ and $(1+r)/(1-r) = y$; its tail after N terms lies between 0 and $\frac{2}{2N+1} \sum_{k \geq 0} r^{2N+1+2k} = \frac{2r^{2N+1}}{(2N+1)(1-r^2)}$. Part (4) holds because the exponential series has nonnegative terms. $\square$

A.1. The trace schedule, the threshold D_* , and numerical facts for Cases 2 and 3. The descent step (Proposition 14.5), from which the Fourier bound (4.2) follows, holds above a threshold D_* , which also determines the mixing constant C of (4.3). This subsection defines D_* together with the parameters and the trace schedule from which it is built, proves their numerical properties, and ends with the numerical facts used in Cases 2 and 3 (Sections 14 and 13).

Definition A.2 (Analytic parameters). Set

$$\varepsilon = \frac{8}{217}, \quad \gamma = \frac{87}{200}, \quad \kappa = \frac{4}{25}, \quad z = \frac{21}{500}, \quad \chi = \frac{63}{2500},$$

$$T_{\text{tr}} = 101, \quad r_{\text{tr}} = 580, \quad \mathbf{p}_0 = 1024, \quad L_{\text{tr}} = \frac{7046129}{2^{20}},$$

$$K_{\text{tr}} = 9863028149, \quad H_{\text{tr}} = 17451, \quad a_{\text{tr}} = \lceil 4K_{\text{tr}}/35 \rceil + 11 = 1127203229.$$

Here ε is the threshold on $|\vartheta|$ separating white from black points; γ is the rate of the reward discount $e^{-\gamma R_w}$ of the trace (Section 12); κ weights the raw letter 3, both in $e^{-\kappa z}$ and in $E_8(\kappa\gamma)$ (Definition A.12); z is the cancellation at a white point and χ the allowance for the envelope factors $e^{\mathbf{t}_m r}$ of Lemma 14.1; T_{tr} is the reward bound and $r_{\text{tr}} + 1$ the number of trace times used; and L_{tr} is the *trace reserve*. The value of a_{tr} follows from $35 \cdot 1127203217 < 4K_{\text{tr}} \leq 35 \cdot 1127203218$.

The *size bound* $\beta(u)$ defined next is the triangle size from which a visit at time u counts towards the oversize event (Definition 13.3). It is linear in u along the positions $\mathbf{p}_0, \dots, \mathbf{p}_{r_{\text{tr}}}$ of the trace schedule, and it grows faster afterwards so that the dyadic sums below converge.

Definition A.3 (Trace schedule). For $u \in \mathbb{N}$ define

$$\beta(u) = K_{\text{tr}} \cdot (u+1) \cdot 2^{(\text{bitlen}(u+1) - H_{\text{tr}})_+}, \quad h_{\text{tr}}(u) = \lceil 4\beta(u)/35 \rceil + 8192, \quad \mathbf{g}(u) = u + h_{\text{tr}}(u) + T_{\text{tr}} + 1,$$

and $\mathbf{p}_{i+1} = \mathbf{g}(\mathbf{p}_i)$ for $i \in \mathbb{N}$, starting from the $\mathbf{p}_0$ of Definition A.2.

Lemma A.4 (Linear initial schedule and summable continuation).

- (1) $\beta(u) \geq K_{\text{tr}} \cdot (u+1)$, and β and $\mathbf{g}$ are nondecreasing with $\mathbf{g}(u) > u$. Hence $(\mathbf{p}_i)$ is increasing.
- (2) $1025a_{\text{tr}}^{r_{\text{tr}}} < 2^{H_{\text{tr}}}$, and for $0 \leq i \leq r_{\text{tr}}$ we have $\beta(\mathbf{p}_i) = K_{\text{tr}} \cdot (\mathbf{p}_i + 1)$ and

$$1025 \left(\frac{35 + 4K_{\text{tr}}}{35} \right)^i \leq \mathbf{p}_i + 1 \leq 1025a_{\text{tr}}^i. \quad (\text{A.2})$$

- (3) For every $\hat{u} \in \mathbb{N}$ and $k \in \mathbb{N}$ put $\mathbf{v}(k) = \min(\hat{u}, 2^{k+1} - 2)$, $\mathbf{X}(k) = 136(\mathbf{v}(k) + 1)$ and $\mathbf{S}(k) = \beta(2^k - 1)$; the notation suppresses the dependence of $\mathbf{v}(k)$ and $\mathbf{X}(k)$ on $\hat{u}$. These describe the dyadic ranges of Case 3 (Section 13). Then $4096\mathbf{X}(k) < \mathbf{S}(k)$ and

$$\sum_{k \geq 0} \frac{\mathbf{X}(k)}{\mathbf{S}(k)} < \frac{272(H_{\text{tr}} + 1)}{K_{\text{tr}}}, \quad \sum_{k \geq 0} \frac{1}{\mathbf{S}(k)} \leq \frac{2}{K_{\text{tr}}}. \quad (\text{A.3})$$

Proof. For (1), the power of 2 in $\beta(u)$ is at least 1, and all three factors of β are nondecreasing in u because bitlen is. Hence h_{tr} and $\mathbf{g}$ are nondecreasing, and $\mathbf{g}(u) > u$ is clear.

For (2), the inequality $1025a_{\text{tr}}^{r_{\text{tr}}} < 2^{H_{\text{tr}}}$ is an exact integer comparison, and $\mathbf{p}_0 + 1 = 1025$. Suppose that (A.2) holds at some $i \leq r_{\text{tr}}$. Then $\mathbf{p}_i + 1 \leq 1025a_{\text{tr}}^{r_{\text{tr}}} < 2^{H_{\text{tr}}}$, so $\text{bitlen}(\mathbf{p}_i + 1) \leq H_{\text{tr}}$ and $\beta(\mathbf{p}_i) = K_{\text{tr}} \cdot (\mathbf{p}_i + 1)$. If moreover $i < r_{\text{tr}}$, then, as $h_{\text{tr}}(u) + T_{\text{tr}} + 1 = \lceil 4\beta(u)/35 \rceil + 8294$,

$\mathbf{p}_{i+1} + 1 = \mathbf{p}_i + 1 + \lceil 4K_{\text{tr}} \cdot (\mathbf{p}_i + 1)/35 \rceil + 8294$. This is at least $(1 + 4K_{\text{tr}}/35)(\mathbf{p}_i + 1)$, and, since $\mathbf{p}_i + 1 \geq 1025$ and $8295 < 10 \cdot 1025$, it is at most

$$\left(1 + \frac{4K_{\text{tr}}}{35}\right)(\mathbf{p}_i + 1) + 8295 < \left(11 + \frac{4K_{\text{tr}}}{35}\right)(\mathbf{p}_i + 1) \leq a_{\text{tr}}(\mathbf{p}_i + 1).$$

Induction on i proves (A.2), and with it $\beta(\mathbf{p}_i) = K_{\text{tr}} \cdot (\mathbf{p}_i + 1)$, for $0 \leq i \leq r_{\text{tr}}$.

For (3), $\text{bitlen}(2^k) = k + 1$ gives $S(k) = K_{\text{tr}} \cdot 2^k \cdot 2^{(k+1-H_{\text{tr}})_+}$. Since $\mathbf{v}(k) \leq 2^{k+1} - 2$, we have $X(k) \leq 136(2^{k+1} - 1) < 272 \cdot 2^k$. Since $K_{\text{tr}} > 4096 \cdot 272 = 1114112$, it follows that $4096X(k) < K_{\text{tr}} \cdot 2^k \leq S(k)$. Moreover

$$\sum_{k \geq 0} \frac{X(k)}{S(k)} < \frac{272}{K_{\text{tr}}} \sum_{k \geq 0} 2^{-(k+1-H_{\text{tr}})_+} = \frac{272}{K_{\text{tr}}} \left(H_{\text{tr}} + \sum_{j \geq 1} 2^{-j} \right) = \frac{272(H_{\text{tr}} + 1)}{K_{\text{tr}}},$$

and $S(k) \geq K_{\text{tr}} \cdot 2^k$ gives the second bound by a geometric series. $\square$

Case 3 (Section 13) needs the last position of the schedule, the size bound there, and a lower bound on the gaps it treats.

Definition A.5 (Schedule-derived quantities). Put

$$P = \mathbf{p}_{r_{\text{tr}}} + 1, \quad B_{\text{sz}} = \beta(\mathbf{p}_{r_{\text{tr}}}), \quad \omega_{\text{far}} = 2^{36}P.$$

Lemma A.6 (Size of P). *We have $B_{\text{sz}} = K_{\text{tr}}P$, $2^{17450} \leq P < 2^{17451}$, $2^{17483} \leq B_{\text{sz}} < 2^{17484}$, and hence $\text{bitlen}(P) = 17451$ and $\text{bitlen}(B_{\text{sz}}) = 17484$. Moreover $\omega_{\text{far}} > 2^{365}$ and $P - 1 \geq 2$.*

Proof. Lemma A.4(2) gives $B_{\text{sz}} = K_{\text{tr}} \cdot (\mathbf{p}_{r_{\text{tr}}} + 1) = K_{\text{tr}}P$ and, with $i = r_{\text{tr}} = 580$, $1025((35 + 4K_{\text{tr}})/35)^{580} \leq P \leq 1025a_{\text{tr}}^{580}$. The claim then reduces to the three exact integer comparisons

$$2^{17450} \cdot 35^{580} \leq 1025(35 + 4K_{\text{tr}})^{580}, \quad 2^{17483} \cdot 35^{580} \leq 1025K_{\text{tr}} \cdot (35 + 4K_{\text{tr}})^{580}, \\ 1025K_{\text{tr}}a_{\text{tr}}^{580} < 2^{17484},$$

together with $K_{\text{tr}} > 2^{33}$, which gives $P = B_{\text{sz}}/K_{\text{tr}} < 2^{17484-33} = 2^{17451}$. The remaining claims follow from $P \geq 2^{17450}$. $\square$

The integer $P - 1 = \mathbf{p}_{r_{\text{tr}}}$ is obtained by iterating the recurrence of Definition A.3, in the form found in the proof of Lemma A.4(2), exactly r_{tr} times from $\mathbf{p}_0$. Both $P - 1$ and P have 5254 decimal digits, beginning 14859748051937132700.

Definition A.7 (Thresholds). Put $D_{\text{cap}} = (B_{\text{sz}} \cdot \text{bitlen}(B_{\text{sz}}))^2$ and

$$D_{\text{wc}} = (256A_*/\chi)^2 + 2, \quad D_{\text{pt}} = 2D_{\text{wc}} + 344, \\ D_{\text{exp}} = 2^{33495}, \quad D_{\text{bad}} = (2^{53}A_*)^2 + 2^{53}(P + 65), \\ D_{\text{sc}} = 2 + P + D_{\text{cap}} + (100\omega_{\text{far}})^2, \\ D_* = \max(D_{\text{wc}}, D_{\text{pt}}, D_{\text{exp}}, D_{\text{bad}}, D_{\text{sc}}, 2).$$

Case 1 of the descent step needs $m \geq D_{\text{wc}}$ and Case 2 needs $m \geq \max(D_{\text{pt}}, D_{\text{exp}})$ (Section 14); Case 3 needs $m \geq \max(D_{\text{bad}}, D_{\text{sc}})$ (Section 13), and D_{cap} enters only through D_{sc} .

Lemma A.8 (The thresholds).

- (1) $D_{\text{wc}} = 836081643217/1296 < 2^{30}$, $D_{\text{pt}} < 2^{31}$, $D_{\text{bad}} < 2^{17506}$, and $D_{\text{sc}} > D_{\text{cap}} \geq 2^{34966}$. Hence $D_* = D_{\text{sc}}$, and

$$\text{bitlen}(3200D_*) = 35008, \quad \left\lfloor 3200D_*/2^{34928} \right\rfloor = 886085405399619011114762. \quad (\text{A.4})$$

- (2) Every real $m \geq D_{\text{sc}}$ satisfies $P \leq m$ and

$$\frac{m}{(\log m)^2} > \omega_{\text{far}}, \quad \frac{m}{(\log m)^2} > \frac{B_{\text{sz}}^2}{2}.$$

(3) Every real $m \geq D_{\text{exp}}$ satisfies $\chi \log m > 234A_*$.

Proof. For (1), $256A_*/\chi = 10241 \cdot 2500/(16 \cdot 63)$, which gives the value of D_{wc} ; its comparison with 2^{30} is exact, and $D_{\text{pt}} < 2^{31}$ follows. Since $A_* < 4$, $(2^{53}A_*)^2 < 2^{110}$. By Lemma A.6, $P+65 < 2^{17452}$, so $D_{\text{bad}} < 2^{110} + 2^{17505} < 2^{17506}$. Also $D_{\text{cap}} \geq B_{\text{sz}}^2 \geq 2^{34966}$ by Lemma A.6, and $D_{\text{exp}} = 2^{33495}$. Thus D_{sc} is the maximum. Equation (A.4) is an exact integer computation from the integer P described above, using $B_{\text{sz}} = K_{\text{tr}}P$ and $\text{bitlen}(B_{\text{sz}}) = 17484$.

For (2), first $P \leq D_{\text{sc}} - 2 \leq m$. The function $x \mapsto x/(\log x)^2$ has derivative $(\log x - 2)/(\log x)^3$, so it increases on $[e^2, \infty)$, and $e^2 < 9$ by Lemma A.1(5). Both D_{cap} and $m' = (100\omega_{\text{far}})^2$ are at least 9 and less than m , so $m/(\log m)^2$ exceeds the values of this function at D_{cap} and at m' .

Since $x \mapsto \log x/x$ decreases for $x > e$ and $\text{bitlen}(B_{\text{sz}}) \geq 4096$, we get $\log \text{bitlen}(B_{\text{sz}})/\text{bitlen}(B_{\text{sz}}) \leq 12 \log 2/4096 < 1/256$. Since $B_{\text{sz}} < 2^{\text{bitlen}(B_{\text{sz}})}$ and $\log 2 < 347/500$,

$$\log D_{\text{cap}} = 2 \log B_{\text{sz}} + 2 \log \text{bitlen}(B_{\text{sz}}) < 2 \text{bitlen}(B_{\text{sz}}) \left(\frac{347}{500} + \frac{1}{256} \right) < \frac{7}{5} \text{bitlen}(B_{\text{sz}}),$$

so $D_{\text{cap}}/(\log D_{\text{cap}})^2 > \frac{25}{49} B_{\text{sz}}^2 > B_{\text{sz}}^2/2$.

For m' , use $\log x < \sqrt{x}$ for $x > 0$ and $\log 10^4 < 10$ (as $e^{10} > 10^4$). Since $\omega_{\text{far}} \geq 1$, we get $\log m' < 10 + 2\sqrt{\omega_{\text{far}}} < 100\sqrt{\omega_{\text{far}}}$, and therefore $m'/(\log m')^2 > 10^4 \omega_{\text{far}}^2 / (10^4 \omega_{\text{far}}) = \omega_{\text{far}}$.

For (3), $\log m \geq 33495 \log 2$. The claim reduces, by Lemma A.1(3), to the exact rational comparison

$$\frac{63}{2500} \cdot 33495 \cdot \frac{693147180}{10^9} = 585.0675 \dots > 585.0571 \dots = 234 \cdot \frac{10241}{4096}.$$

□

Lemma A.9 (Numerical facts for Case 2). $e^{-z+\chi/2} < 1 - \frac{3}{5}z$ and $\frac{\chi}{8} < \frac{3}{16} \cdot \frac{3}{5}z$.

Proof. Since $z - \chi/2 = 147/5000$, the Taylor bound $e^{-y} \leq 1 - y + y^2/2$ ($y \geq 0$) gives $e^{-z+\chi/2} \leq \frac{48551609}{50000000} < \frac{48740000}{50000000} = 1 - \frac{3}{5}z$; and $\frac{\chi}{8} = \frac{126}{40000} < \frac{189}{40000} = \frac{3}{16} \cdot \frac{3}{5}z$. □

Lemma A.10 (Numerical facts for Case 3).

(1) $\frac{5}{16} \leq \lambda < \frac{8}{25}$ and $1/\lambda < 158497/50000 < 127/40$.

(2) $(1+2\lambda)\frac{17}{10}\sqrt{2} + \left(\frac{1}{\lambda} + 2\right)\frac{33}{32} \cdot \frac{25}{36} < Q_{\text{pk}} = \frac{34267893}{4480000}$.

(3) If v is an integer with $0 \leq v \leq P-1$ and $\omega \geq \omega_{\text{far}}$ is real, then $2\omega^{3/5} + 272\lambda(v+1) + 1 \leq \omega(\lambda - \frac{1}{4})$.

(4) If $\omega \geq 2^{160}$ is real, then $2^{55}e^{-\omega^{1/5}/65536} < (2\omega)^{-1/2}$.

Proof. (1) follows from the integer comparisons $9^5 = 59049 \leq 2^{16}$, $2^{25} < 9^8$ and $9^{50000} < 2^{158497}$ (the integer 9^{50000} has 158497 binary digits). For (2), use (1), $1+2\lambda < 41/25$ and $\sqrt{2} < 99/70$ (as $99^2 > 2 \cdot 70^2$): the left side is less than

$$\frac{41}{25} \cdot \frac{17}{10} \cdot \frac{99}{70} + \left(\frac{127}{40} + 2\right)\frac{33}{32} \cdot \frac{25}{36} = \frac{69003}{17500} + \frac{3795}{1024} = \frac{34267893}{4480000},$$

an exact rational identity ($Q_{\text{pk}} \approx 7.6491$). For (3), $\omega \geq \omega_{\text{far}} = 2^{36}P \geq 2^{15}$ gives $\omega^{2/5} \geq 2^6$, so $2\omega^{3/5} \leq \omega/32$; and $272\lambda(v+1) + 1 \leq (2176/25 + 1)P \leq 2^{31}P \leq \omega/32$ by (1). Hence the left side is at most $\omega/16 \leq \omega(\lambda - \frac{1}{4})$. For (4), $\omega^{1/5} \geq 2^{32}$, and it suffices that $\frac{111}{2} \log 2 + \frac{1}{2} \log \omega < 2^{-16}\omega^{1/5}$. Now $\frac{1}{2} \log \omega = 10 \log \omega^{1/20} < 10\omega^{1/20} \leq 2^{-17}\omega^{1/5}$ because $\omega^{3/20} \geq 2^{24} > 10 \cdot 2^{17}$, and $\frac{111}{2} \log 2 < 2^{15} \leq 2^{-17}\omega^{1/5}$. □

Lemma A.11 (Deviation errors). Let $\hat{u} \geq 2$ be an integer, take $\mathbf{v}(k)$ from Lemma A.4(3) with this $\hat{u}$, put $\bar{k} = \lfloor \log_2(\hat{u} + 1) \rfloor$, and let Exc be the deviation error of Section 13. Then $\sum_{k=0}^{\bar{k}} \text{Exc}(\mathbf{v}(k)) < 1/3360$.

For (3), $\mathcal{B}$ is an explicit rational, since P_{tr} and P_{wc} are finite sums of rationals. Its six terms are 0.000125, 0.1881196438..., 0.0616718457..., 0.0166666666..., 0.7334168436... and $5.8 \dots 10^{-12}$. The bounds on $1 - \mathcal{B}$ are exact rational comparisons. $\square$

A.3. Rational logarithms. This subsection proves (4.4), the enclosure of $\log_2 g$, the ceiling $j_0 = 6103914$ in (6.3), and (6.5) by means of (A.1), which is finer than the decimal values of Lemma A.1.

Write $(\log y)^{\text{lo}}, (\log y)^{\text{hi}}$ for the lower and upper sides of (A.1) with $N = 80$, and put $\mathbf{a} = \lfloor 3200D_*/2^{34928} \rfloor$, the integer of (A.4), so that $\mathbf{a}/2^{79} \leq 3200D_*/2^{35007} < (\mathbf{a} + 1)/2^{79}$. By (4.3), $C + 1 = (477/20)(3200D_*)^{A_*} + \frac{169}{10}$, where $\log(477/20) = 4\log 2 + \log(477/320)$ and the additive term changes the natural logarithm by less than 2^{-100} . Hence the rationals

$$\begin{aligned} (\log(3200D_*))^{\text{lo}} &= 35007(\log 2)^{\text{lo}} + (\log(\mathbf{a}/2^{79}))^{\text{lo}}, \\ (\log(3200D_*))^{\text{hi}} &= 35007(\log 2)^{\text{hi}} + (\log((\mathbf{a} + 1)/2^{79}))^{\text{hi}}, \\ (\log_2(C + 1))^{\text{lo}} &= \frac{4(\log 2)^{\text{lo}} + (\log(477/320))^{\text{lo}} + A_*(\log(3200D_*))^{\text{lo}}}{(\log 2)^{\text{hi}}}, \\ (\log_2(C + 1))^{\text{hi}} &= \frac{4(\log 2)^{\text{hi}} + (\log(477/320))^{\text{hi}} + A_*(\log(3200D_*))^{\text{hi}} + 2^{-100}}{(\log 2)^{\text{lo}}}, \\ (\log_2 3)^{\text{lo}} &= \frac{(\log 2)^{\text{lo}} + (\log(3/2))^{\text{lo}}}{(\log 2)^{\text{hi}}}, \quad (\log_2 3)^{\text{hi}} = \frac{(\log 2)^{\text{hi}} + (\log(3/2))^{\text{hi}}}{(\log 2)^{\text{lo}}}, \\ (\log_2 g)^{\text{lo}} &= \frac{(\log(101/100))^{\text{lo}}}{(\log 2)^{\text{hi}}}, \quad (\log_2 g)^{\text{hi}} = \frac{(\log(101/100))^{\text{hi}}}{(\log 2)^{\text{lo}}} \end{aligned}$$

satisfy $(\log_2(C + 1))^{\text{lo}} < \log_2 C < \log_2(C + 1) < (\log_2(C + 1))^{\text{hi}}$, which gives (4.4), and $(\log_2 g)^{\text{lo}} \leq \log_2 g \leq (\log_2 g)^{\text{hi}}$, which gives

$$0.014355292977070041430 < \log_2 g < 0.014355292977070041432.$$

Hence the quantity $(23 + \log_2(3(C + 1)) + 535/8)/\log_2 g$ in (6.3) lies between

$$\frac{23 + (\log_2 3)^{\text{lo}} + (\log_2(C + 1))^{\text{lo}} + 535/8}{(\log_2 g)^{\text{hi}}} \quad \text{and} \quad \frac{23 + (\log_2 3)^{\text{hi}} + (\log_2(C + 1))^{\text{hi}} + 535/8}{(\log_2 g)^{\text{lo}}},$$

and both have ceiling 6103914. Finally, $16 + j_0(\log_2 g)^{\text{lo}}$ and $16 + j_0(\log_2 g)^{\text{hi}}$ give the enclosure (6.5). All divisions in these formulas have positive denominators.

A.4. Exact enumeration of the first five blocks. This subsection computes the retained fraction (5.4) of the five-block selector, used in Lemma 5.3. The masses $\text{mass}(\text{FC}(\delta, 16))$ for $9 \leq \delta \leq 13$ are given by Lemma 15.9; since the selector bounds rounded offsets, we refine the counts (15.1) by the rounded offset. As $B_0(n) = \text{bitlen}(3^n)$ for $n \geq 1$, no logarithm is evaluated. For $\delta \leq 11$ the lower bounds $o_\delta(d)$ on the overshoot, in increasing order of $d = d_\delta^{\text{lo}} + 1, \dots, d_\delta^{\text{hi}}$, are

$$\begin{aligned} \delta = 9 &: (1, 1, 2, 1, 1, 0, 0, 1, 0, 1), \\ \delta = 10 &: (0, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0), \\ \delta = 11 &: (0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0). \end{aligned}$$

Fix δ with $9 \leq \delta \leq 13$ and an integer precision $\mathbf{r}$; the quantities below depend on δ and $\mathbf{r}$, which the notation suppresses. The *bin* of a word v is $2^{\mathbf{r}} \text{rd}_{\mathbf{r}}(v) = \sum_{i=1}^{|v|} I_{\mathbf{r}}(i, \deg_i(v))$, where $I_{\mathbf{r}}(i, l) = \lfloor 2^{\mathbf{r}} 3^{i-1} 2^{-l} \rfloor$ is the contribution to the bin. Put $\bar{H}_i = H_{\delta, r_\delta}(\max(i, d_\delta^{\text{lo}}))$, and let $\mathcal{N}_i(l, y)$ count the words v of length i , degree l and bin y with $\deg_{i'}(v) < \bar{H}_{i'}$ for $1 \leq i' \leq i$. A word of length i is a word of length $i - 1$ followed by one positive letter, which raises the degree from some $l' < l$ to l and the bin by $I_{\mathbf{r}}(i, l)$. Hence $\mathcal{N}_0(l, y) = \mathbf{1}[l = y = 0]$ and

$$\mathcal{N}_i(l, y) = \mathbf{1}[l < \bar{H}_i] \sum_{l' < l} \mathcal{N}_{i-1}(l', y - I_{\mathbf{r}}(i, l)) \quad (i \geq 1), \quad (\text{A.6})$$

where cells of negative bin are zero. Since $\deg_{i'}(v) < \deg_{d_\delta^{\text{lo}}}(v)$ for $i' < d_\delta^{\text{lo}}$, for $d_\delta^{\text{lo}} \leq i < d_\delta^{\text{hi}}$ the sum $\sum_y \mathcal{N}_i(l, y)$ is the count $\mathcal{N}_i(l)$ of (15.1).

As in the proof of Lemma 15.9, a word of $\text{FC}(\delta, 16)$ of length d is a word counted by $\mathcal{N}_{d-1}$ followed by the unique letter that gives degree c , where $d_\delta^{\text{lo}} < d \leq d_\delta^{\text{hi}}$ and $H_{\delta, r_\delta}(d) + o_\delta(d) \leq c \leq H_{\delta, r_\delta}(d) + 16$. For such d and c , the number of words of $\text{FC}(\delta, 16)$ with length d , degree c and bin y is therefore

$$\mathcal{N}_d^{\text{done}}(c, y) = \sum_{l' < c} \mathcal{N}_{d-1}(l', y - I_{\mathbf{r}}(d, c)), \quad (\text{A.7})$$

and we put $\mathcal{N}_d^{\text{done}}(c, y) = 0$ otherwise. For an integer *bin bound* $y_{\max} \geq 0$, the set $\text{FC}^{\text{bin}}(\delta, 16) = \{v \in \text{FC}(\delta, 16) : 2^{\mathbf{r}} \text{rd}_{\mathbf{r}}(v) \leq y_{\max}\}$ has mass

$$\text{mass}(\text{FC}^{\text{bin}}(\delta, 16)) = \sum_{d, c} \sum_{y=0}^{y_{\max}} 2^{-c} \mathcal{N}_d^{\text{done}}(c, y). \quad (\text{A.8})$$

Bins never decrease along prefixes, so only cells with $y \leq y_{\max}$ are needed.

The following table lists, for the $(\mathbf{r}, y_{\max})$ of row δ , the numerators over $2^{\hat{l}_\delta}$, without cancellation, of $\text{mass}(\text{FC}(\delta, 16))$ and $\text{mass}(\text{FC}^{\text{bin}}(\delta, 16))$; here $\hat{l}_\delta = H_{\delta, r_\delta}(d_\delta^{\text{hi}}) + 16$ bounds every degree. For $\delta \geq 10$ the bin bound is the condition of the five-block selector on block $v_{\delta-9}$. For $\delta = 9$, the bin bound $y \leq 1216$, that is $\text{rd}_9(v_0) \leq 19/8$, is implied by the upper bound in the first condition of the selector; the lower bound $\frac{3}{8}$ in that condition is imposed when the first two blocks are joined.

δ	$\hat{l}_\delta$	$\mathbf{r}$	$y_{\max}$	numerator without bin bound	bin-bounded numerator
9	42	9	1216	1121528720843	667669768647
10	46	6	1024	25912441038107	25532579216651
11	48	3	128	110172925800689	108161817817367
12	52	0	64	2156388303073866	2155020659894667
13	54	-3	32	8200094289238498	8199875792177575

Finally we join the first two blocks, which the first condition of the selector couples. Write $\mathcal{N}_{9,d}^{\text{done}}$ and $\mathcal{N}_{10,d}^{\text{done}}$ for the tables (A.7) with $(\delta, \mathbf{r}) = (9, 9)$ and $(10, 6)$, and let $\mathbf{m}(y_1) = \sum_{d,c} 2^{-c} \mathcal{N}_{10,d}^{\text{done}}(c, y_1)$ be the mass of the second blocks $v_1 \in \text{FC}(10, 16)$ with bin $y_1 = 2^6 \text{rd}_6(v_1)$. Call the length, degree and bin (d, l, y) of a first block v_0 a *first-block triple*. With $y = 2^9 \text{rd}_9(v_0)$ and $\text{wt}(v_0) = 3^d 2^{-l}$, multiplying the first condition of the selector by $512 \cdot 2^l$ gives exactly

$$192 \cdot 2^l \leq y \cdot 2^l + 8 \cdot 3^d y_1 \leq 1216 \cdot 2^l.$$

With the condition $y_1 \leq 1024$ on the second block, the allowed interval of y_1 is therefore $y_1^{\text{lo}}(d, l, y) \leq y_1 \leq y_1^{\text{hi}}(d, l, y)$, where

$$y_1^{\text{lo}}(d, l, y) = \max \left(0, \left\lceil \frac{(192 - y)2^l}{8 \cdot 3^d} \right\rceil \right),$$

$$y_1^{\text{hi}}(d, l, y) = \min \left(1024, \left\lfloor \frac{(1216 - y)2^l}{8 \cdot 3^d} \right\rfloor \right)$$

(signed integer ceilings and floors). Hence the joint mass of the first two blocks is

$$P_{\text{jt}} = \sum_{d, l, y} 2^{-l} \mathcal{N}_{9,d}^{\text{done}}(l, y) \sum_{y_1 = y_1^{\text{lo}}(d, l, y)}^{y_1^{\text{hi}}(d, l, y)} \mathbf{m}(y_1), \quad (\text{A.9})$$

an empty inner sum being zero. Exactly 109715 first-block triples give a positive summand, and $2^{88} P_{\text{jt}}$ is the integer 15856838138131582931885054, so

$$P_{\text{jt}} = \frac{7928419069065791465942527}{2^{87}}.$$

Consequently $\log_2 q_{*,M} < 2N_{*,M} < 100000 \log_2 C_M$. Write

$$\mathcal{M}_M = 4^{2b_{0,M}+1+3^{q_{*,M}}}, \quad \log_2 \mathcal{M}_M = 4b_{0,M} + 2 + 2 \cdot 3^{q_{*,M}}, \quad m_{*,M} = \left\lfloor (88\mathcal{M}_M C_M)^{1/6} \right\rfloor + 1.$$

Then

$$1000 \log_2 C_M < \log_2 \mathcal{M}_M < 3 \cdot 3^{q_{*,M}}, \quad m_{*,M} \leq 100 \mathcal{M}_M C_M, \quad \log_2 m_{*,M} < 2 \log_2 \mathcal{M}_M. \quad (\text{B.3})$$

B.1. Mazur's density constant. We bound the iterated logarithms of $1/c_M$. By [8, (8.6)], $c_M = 3/(256\mathcal{M}_M \cdot 2^{m_{*,M}})$. Thus $\log_2(1/c_M) = m_{*,M} + \log_2 \mathcal{M}_M + 8 - \log_2 3 < 128\mathcal{M}_M C_M$. We obtain

$$\begin{aligned} \log_2^{\circ 2}(1/c_M) &< 2 \log_2 \mathcal{M}_M, & \log_2^{\circ 3}(1/c_M) &< 2q_{*,M}, \\ \log_2^{\circ 4}(1/c_M) &< 1 + \log_2 q_{*,M} < 100001 \log_2 C_M. \end{aligned}$$

Here $\log_2(2 \log_2 \mathcal{M}_M) < \log_2 6 + q_{*,M} \log_2 3 < 2q_{*,M}$, since $q_{*,M} \geq N_{*,M} > 100$. Taking another logarithm and using (B.2) gives

$$\log_2^{\circ 5}(1/c_M) < \log_2^{\circ 2}(C_M) + 17 < r_M \log_2 3 + 207. \quad (\text{B.4})$$

B.2. Mazur's cutoff. We bound the iterated logarithms of X_M in the same way. Mazur defines [8, (8.4)–(8.6)]

$$\begin{aligned} J_{*,M} &= \max(N_{*,M} + N_{t,M}, 200(\lceil \log_2 m_{*,M} \rceil - 78)_+, 20 \cdot 10^9), \\ l_{0,M} &= 0, \quad l_{j+1,M} = l_{j,M} + 17 + \lfloor j/100 \rfloor + 2b_{j,M} - \lfloor 19b_{j,M}/12 \rfloor, \\ B_{\text{ht},M} &= l_{J_{*,M},M} + \left\lfloor 2b_{J_{*,M},M}/3 \right\rfloor, \end{aligned}$$

and $X_M = 32(2^{B_{\text{ht},M}} \mathcal{M}_M + 1)$. Equations (B.2)–(B.3) imply $J_{*,M} < 1000 \log_2 \mathcal{M}_M$. Using monotonicity of $j \mapsto b_{j,M}$ and the upper bound $2b_{j,M}$ for the last two terms of the recurrence,

$$B_{\text{ht},M} \leq 4(J_{*,M} + 17)^2 b_{J_{*,M},M},$$

$$\log_2 B_{\text{ht},M} \leq 82 + 2 \log_2(J_{*,M} + 17) + J_{*,M} < 4J_{*,M} < 4000 \log_2 \mathcal{M}_M.$$

For X_M this gives

$$\log_2(X_M) < 16 + B_{\text{ht},M} + \log_2 \mathcal{M}_M < 2^{4096 \log_2 \mathcal{M}_M}, \quad \log_2^{\circ 2}(X_M) < 4096 \log_2 \mathcal{M}_M,$$

$$\log_2^{\circ 3}(X_M) < 12 + \log_2^{\circ 2}(\mathcal{M}_M) < 2q_{*,M}.$$

The last inequality uses $\log_2 \mathcal{M}_M < 3 \cdot 3^{q_{*,M}}$, $\log_2 3 < 8/5$, and $q_{*,M} > 100$. Hence $\log_2^{\circ 4}(X_M) < 100001 \log_2 C_M$, and the same fifth-logarithm bound (B.4) holds for X_M .

B.3. Counting the iterated logarithms. It remains to count the logarithms. For either $Y = 1/c_M$ or $Y = X_M$, the preceding bounds give

$$\log_2^{\circ 5}(Y) < 2r_M + 207 < 2^{4a_M+17}.$$

Indeed $2r_M = 128180 \cdot 2^{4a_M} + 49172 \cdot 2^{a_M}$, and $49172 \cdot 2^{a_M} + 207 < 2892 \cdot 2^{4a_M}$, where $128180 + 2892 = 2^{17}$. Therefore

$$\log_2^{\circ 6}(Y) < 4a_M + 17 = 2^{172} + 17 < 2^{173}, \quad \log_2^{\circ 7}(Y) < 173.$$

Since $Y > \mathcal{M}_M$ and $q_{*,M} > 100$, the iterated logarithms $\log_2^{\circ i}(Y)$ with $i \leq 6$ are positive, so exponentiating seven times gives $1/c_M < D_M$ and $X_M < D_M$. A larger cutoff and a smaller density constant preserve Mazur's theorem, so it holds with the cutoff $D_M > X_M$ and the density constant $1/D_M < c_M$. This proves (1.1).

APPENDIX C. AXIOMPROVER LEAN CERTIFICATE

This paper was written using human-AI collaboration with AxiomProver. AxiomProver is an AI system for mathematical research through formal proof, currently under development at Axiom Math.

The explicit positive lower density established in Theorem 1.3, which is the main result in this paper, has been formalized with AxiomProver in Lean. The repository contains the formal statements, proofs, and instructions for checking the certificate:

<https://github.com/AxiomMath/CollatzPositiveDensity>

The human-written challenge file gives the definitions and hypotheses. The solution proves the stated targets. The repository reports successful verification with the Lean comparator.

APPENDIX D. AI DECLARATION

This project was born out of our goal to formalize Mazur’s paper, with the intent of improving the estimates. With this stated goal, AxiomProver autonomously produced the mathematics in this paper and formalized it in Lean. Although this paper was written with AI assistance, the human author is accountable for the statements and conclusions in the introduction, and summarized the main argument in Section 2 to facilitate human understanding.

REFERENCES

- [1] S. Boucheron, G. Lugosi, and P. Massart, *Concentration inequalities: a nonasymptotic theory of independence*, Oxford University Press, Oxford, 2013.
- [2] R. Durrett, *Probability: theory and examples*, 5th ed., Cambridge University Press, Cambridge, 2019.
- [3] W. Feller, *An introduction to probability theory and its applications*, Vol. I, 3rd ed., Wiley, New York, 1968.
- [4] W. Hoeffding, *Probability inequalities for sums of bounded random variables*, J. Amer. Statist. Assoc. **58** (1963), no. 301, 13–30. doi:10.1080/01621459.1963.10500830.
- [5] I. Krasikov and J. C. Lagarias, *Bounds for the $3x + 1$ problem using difference inequalities*, Acta Arith. **109** (2003), no. 3, 237–258. doi:10.4064/aa109-3-4.
- [6] M. Loève, *Probability theory I*, 4th ed., Graduate Texts in Mathematics, vol. 45, Springer, New York, 1977.
- [7] L. Mazur, *Positive lower density of Collatz predecessors*, manuscript, version 1.0, September 6, 2026. <https://proofatlas.ai/papers/positive-lower-density-collatz-predecessors/paper-v1.pdf>.
- [8] L. Mazur, *Explicit positive-density Collatz convergence in logarithmic time*, manuscript, version 2.1, September 6, 2026. https://proofatlas.ai/papers/positive-density-log-time-collatz/Mazur_Explicit_Positive_Density_Collatz_Convergence_in_Logarithmic_Time_v2.pdf.
- [9] I. A. Shaik, *Positive density of Collatz convergence at every rate above $3/\log(4/3)$* , preprint, version 1.0.0, September 21, 2026. doi:10.5281/zenodo.22871622.
- [10] T. Tao, *Almost all orbits of the Collatz map attain almost bounded values*, Forum Math. Pi **10** (2022), e12. doi:10.1017/fmp.2022.8. Cited by the numbering of the corrected version arXiv:1909.03562v7 (2026), whose Section 6 is the form used here.

AXIOM MATH, 124 UNIVERSITY AVENUE, PALO ALTO, CA 94301

Email address: kenny@axiommath.ai