

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 every N , at least $\lfloor N/2^{2^{163}} \rfloor$ odd positive integers $n < N$ reach 1 within $10.46 \log n$ ordinary steps and $3.4881 \log n$ 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 $D_A = 2^{2^{163}}$. Then, for every N , at least $\lfloor N/D_A \rfloor$ odd starting values $n < N$ reach 1 within $\lfloor 10.46 \log n \rfloor$ ordinary steps and $\lfloor 3.4881 \log n \rfloor$ 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 $c(t) = t^{-25/11} D_A^{-1}$. Then, for every N , at least $\lfloor c(t)N \rfloor$ odd starting values $n < N$ reach t within $\lfloor 10.46 \log(n/t) \rfloor$ ordinary steps using the Collatz map and reach the odd part of t within $\lfloor 3.4881 \log(n/t) \rfloor$ accelerated steps using the Syracuse map.*

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 .

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

Key words and phrases. Collatz conjecture, arithmetic dynamics.

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 , which implies they 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 [9] 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 [7] 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

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

Mazur also proved [6] 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 [8], 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:

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

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 [9], which states that for large enough d , the residue density ρ_d is, on average over residues, well approximated by ρ_m for much smaller m . More precisely,

$$3^{-d} \sum_{y \in \mathbb{Z}/3^d \mathbb{Z}} |\rho_d(y) - \rho_m(y \bmod 3^m)| \leq 2^{96} m^{-1.45} \log m$$

for any $d \geq m \geq 2^{80}$, and the same average is also at most $2^{365} m^{-1.8693} \log m$. The first bound has the smaller constant and is used for the propagation estimate below, and the second has the larger exponent and is used for the census inequality. They come from two Fourier bounds with different parameters. 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 $0 \leq \text{off}(w) \leq \frac{3}{4} + 2^{-8}$, whenever $j \geq 29$.

We choose $R_j = 29 + \lfloor j/56 \rfloor$ explicitly, but leave δ_j undefined here except that $\delta_0 = 16544$ and it behaves like $16544 \cdot 1.01^j$.

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 $j \geq 9853$, Proposition 6.1 and Corollary 6.2 state that

$$|Z_{j+1}(t) - Z_j(t)| \leq 0.007584 \cdot 0.990945^{j-9853} + 2^{-90} \cdot 0.995^{j-9853}. \quad (2.1)$$

As a consequence, the sum of all these errors from generation 9853 onwards is less than $6/7$ (Lemma 6.3), so it suffices that the main term at that generation exceeds $6/7$ by a margin.

We indeed have a large enough main term. Precisely, let $j_0 := 9853$ and $d_0 := 5 \cdot 2^{160}$. 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} . Because of the filter, the words of L_{j_0} concentrate in one residue class modulo 3^8 , so we average over that class. The estimate we need is that the average value of $Z_{j_0}(t)$ across the residues $t \equiv 6275 \pmod{3^8}$ is greater than $\frac{13}{15}$, which is supplied by Proposition 6.7:

$$\frac{1}{3^{d_0-8}} \sum_{\substack{t \in \mathbb{Z}/3^{d_0}\mathbb{Z} \\ t \equiv 6275 \pmod{3^8}}} Z_{j_0}(t) > \frac{13}{15}. \quad (2.2)$$

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. Proposition 6.9 then tells us the following:

$$\text{Let } t \geq 2^{66176}. \text{ Then every } X \geq 2^{2^{352}}t \text{ is reachable from some } j \geq 23000. \quad (2.3)$$

This requires another hypothesis on t :

$$t \geq 2^{66176}. \quad (\text{H3})$$

We now provide an estimate to connect the weighted density to the terminal weighted density. Precisely, for $j \geq 23000$ and X reachable from j , according to Proposition 6.11,

$$|Z_{j,X}^{\text{term}}(t) - Z_j(t)| \leq 2^{169} \cdot 0.993^j. \quad (2.4)$$

And finally an inequality to connect the terminal weighted density to the total weight: for any j and any X reachable from j , (6.13) provides us with the bound

$$Z_{j,X}^{\text{term}}(t) \leq 2^{177} t^{14/11} \beta_X(t) + 2^{-32}. \quad (2.5)$$

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 residue $y_0 \in \mathbb{Z}/3^{d_0}\mathbb{Z}$ with

$$Z_{j_0}(y_0) > \frac{13}{15},$$

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 $X \geq 2^{2^{352}}t$, with $j \geq 23000$ 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 $\frac{6}{7}$, so that we can obtain

$$Z_j(t) > 2^{-28},$$

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

$$Z_{j,X}^{\text{term}}(t) > 2^{-29}.$$

Now the inequality linking terminal weighted density with total weight (2.5) gives us

$$\beta_X(t) > 2^{-207} t^{-14/11},$$

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

$$\#\{s \in [X, 32X) \text{ maps to } t\} > 2^{-207} t^{-25/11} X,$$

subject to the hypotheses (H1), (H2), (H3), (H4). This is Proposition 7.1.

2.5.3. Discharging the hypotheses: choosing a seed. Recall the hypotheses require that $t \geq 2^{66176}$ 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 7.3. 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 7.2). 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 7.3, 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 $X \geq 2^{2352} t_1$, we have

$$\#\{s \in [X, 32X) \text{ maps to } t_1\} > 2^{-207} t_1^{-25/11} X.$$

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

$$\#\{s \in [X, 32X) \text{ maps to } t\} \geq t^{-25/11} D_A^{-1} X.$$

The exact version of this bound is Proposition 7.4. 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 $< X$. For $X < D_A t$ the count $\left\lfloor t^{-25/11} D_A^{-1} X \right\rfloor$ in the main theorem is 0, so no lower bound on X is needed.

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

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

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. [7, Section 2], [9, 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 $j \geq 1$ (that is, y lies on a cycle of S), 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 [7, Section 2]. For a word w , the map $\sigma_w : G_k \rightarrow G_{k+|w|}$, $\sigma_w(y) = \text{wt}(w)y + \text{off}(w)$, 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 $\sigma_w(y)$ and zero off the image. By [7, 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 σ_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

This section proves the mixing lemma of Section 2 (Proposition 4.1) from the Fourier bound (4.2), and derives the transfer estimate used by propagation and by the terminal error (Lemma 4.8). Here n denotes a level.

The Fourier bound is proved for two parameter sets, indexed by $\iota \in \{\text{o}, \text{b}\}$ (Table 1): the original set o , whose larger exponent serves the terminal error and the census, and the bootstrap set b , whose much smaller threshold serves propagation. Put

$$A_{\text{o}} = \frac{21193}{10000}, \quad A_{\text{b}} = \frac{17}{10}, \quad a_{\iota} = A_{\iota} - \frac{1}{4}, \quad n_0 = 2^{80}, \quad (4.1)$$

so $a_{\text{o}} = \frac{18693}{10000}$ and $a_{\text{b}} = \frac{29}{20}$. Theorem 14.6, an explicit form of [9, Proposition 1.17] proved in Sections 9–14, gives for each ι , every $n \geq 1$ and every frequency $\xi \in G_n$ prime to 3 the *Fourier bound*

$$|\hat{\mu}_n(\xi)| \leq \left(\frac{D_F(\iota)}{n} \right)^{A_{\iota}}, \quad \hat{\mu}_n(\xi) = \sum_{y \in G_n} \mu_n(y) e^{-2\pi i \xi y / 3^n}, \quad (4.2)$$

with $D_F(\iota)$ and $B_F(\iota)$ as in Definition A.5; by Lemma A.6(3), $\log_2 B_F(o) < 325.7939550288$ and $\log_2 B_F(b) < 88.1188851177$. Mixing is proved first between levels whose ratio is at most $\frac{27}{25}$ (Lemma 4.7) and then along a geometric chain of levels in ratio $c_{\text{ch}} = \frac{1079}{1000}$. The chain is summed exactly by the *mixing profile*

$$\mathfrak{M}_\iota(x) = 18B_F(\iota)x^{-a_\iota}(\mathfrak{c}_\iota \log x + \mathfrak{d}_\iota) + \mathfrak{s}x^{-2} \quad (x \geq n_0), \quad (4.3)$$

where, with $\mathfrak{z}_\iota = c_{\text{ch}}^{a_\iota}$,

$$\mathfrak{c}_\iota = \frac{1}{\mathfrak{z}_\iota - 1}, \quad \mathfrak{d}_\iota = \frac{\mathfrak{z}_\iota \log c_{\text{ch}}}{(\mathfrak{z}_\iota - 1)^2}, \quad \mathfrak{s} = \frac{15}{c_{\text{ch}}^2 - 1} = \frac{15000000}{164241}.$$

Proposition 4.1 (Fine-scale mixing). *Let $\iota \in \{o, b\}$. For integers $n_0 \leq m \leq d$,*

$$\langle |\rho_d - \rho_m \circ \text{pr}_{d,m}| \rangle_d \leq \mathfrak{M}_\iota(m). \quad (4.4)$$

The profile $\mathfrak{M}_\iota$ is nonincreasing on $[n_0, \infty)$, and for real $x \geq n_0$

$$\mathfrak{M}_o(x) \leq 126B_F(o)x^{-a_o} \log x + 105x^{-2} \leq 2^{365}x^{-a_o} \log x, \quad \mathfrak{M}_b(x) \leq 2^{96}x^{-a_b} \log x. \quad (4.5)$$

4.1. Tail estimates. Mixing discards the random words whose partial degrees or offsets are atypical; two tail bounds control them. The first is a Chernoff bound for partial degrees.

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

$$\mathbb{P}(|\deg_d(W) - 2d| \geq x) \leq 2 \exp\left(-\min\{x^2/(32d), x/8\}\right). \quad (4.6)$$

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 $2e^{\mathfrak{t}}/(2 - e^{\mathfrak{t}})^2$ is nonnegative and increasing, at most 6 on $|\mathfrak{t}| \leq 1/4$ (as $e^{1/4} \leq 4/3$) and at most $\frac{496}{225} < \frac{9}{4}$ on $|\mathfrak{t}| \leq 1/32$ (as $e^{1/32} \leq 32/31$); Taylor's formula gives the bounds. Since $\mathbb{E}e^{\mathfrak{t}(\deg_d(W)-2d)} = \mathbf{M}_{\text{geo}}(\mathfrak{t})^d$, the Cramér–Chernoff method [1, Section 2.2] for $\pm(\deg_d(W) - 2d)$ with $\mathfrak{t} = \min(x/(16d), 1/4)$ gives (4.6) (cf. [7, Section 3.1]). $\square$

The second bounds the offset, which has a $1/x$ tail.

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 $x \geq 5$,

$$3 \sum_{a \geq 1} \frac{2^{-a}}{2^a x + 5} \leq \frac{1}{x+2}. \quad (4.7)$$

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 (4.7), the probability is at most $\sum_a 2^{-a} \frac{7}{x_a + 2} = 21 \sum_a \frac{2^{-a}}{2^a x + 5} \leq \frac{7}{x+2}$. $\square$

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 $\text{Osc}_{n,m}(\mu_n) = \langle |\rho_n - \rho_m \circ \text{pr}_{n,m}| \rangle_n$. For $m \leq r \leq n$, $\text{Av}_{n,m} = \text{Av}_{n,m} \text{Av}_{n,r}$ and projectivity identifies $\text{Osc}_{n,m}(\text{Av}_{n,r} \mu_n)$ with $\text{Osc}_{r,m}(\mu_r)$, so the triangle inequality gives

$$\text{Osc}_{n,m}(\mu_n) \leq \text{Osc}_{n,r}(\mu_n) + \text{Osc}_{r,m}(\mu_r). \quad (4.8)$$

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 4.4 (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 [9, (6.10)–(6.11)]. Fiber averaging keeps exactly the Fourier coefficients at frequencies factoring through G_m , so Cauchy–Schwarz and Parseval give

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

□

4.3. A fixed binary clock. Following [9, Section 6], we cut a random word of length n where its partial degree first exceeds a height N with 2^{N+1} just above 3^n . The part before the cut, the *head*, has offset determined up to few collisions, and the part after it is a fresh random word, whose offset law inherits the Fourier bound. Throughout this subsection $n \geq n_0$ and

$$N = \lceil n \log_2 3 \rceil - 1, \quad h = \lceil \sqrt{128n \log n} \rceil, \quad k_- = \lfloor \frac{N+1}{2} - h \rfloor, \quad k_+ = \lceil \frac{N+1}{2} + h \rceil, \quad j = \lceil 2 \log_2 n \rceil. \quad (4.9)$$

From $1.58 < \log_2 3 < 1.585$, $\log n \geq 80 \log 2 > 55$ and the monotonicity of $(\log n)/n$ on $[n_0, \infty)$ one gets

$$64 \log n \leq h \leq \frac{n}{21600}, \quad 0 < k_-, \quad 21600 k_+ < 17119 n + 32400, \quad j + 1 < 3 \log n. \quad (4.10)$$

Realize the letters $a_1, a_2, \dots$ as the waiting times between the successes of independent fair bits, so that the partial degrees are the positions of the successes. Expose the first N bits, and let D be their number of successes, at positions $s_1 < \dots < s_D$. Each pattern of these bits has probability 2^{-N} , D has the binomial law with parameters N and $\frac{1}{2}$, and given $D = d$ the set $\{s_1, \dots, s_d\}$ is a uniform d -subset of $\{1, \dots, N\}$. For a set $\mathcal{C} = \{s_1 < \dots < s_d\}$ put $O(\mathcal{C}) = \sum_{i \leq d} 3^{i-1} 2^{-s_i}$; so $O(\{s_1, \dots, s_D\}) = \text{off}(W_{\leq D})$. After the last exposed success the remaining waiting times are fresh geometric variables, by memorylessness, so on $D < n$, by (3.2),

$$\text{off}(W_{\leq n}) = O(\{s_1, \dots, s_D\}) + 3^D 2^{-N} \text{off}(W'_{\leq n-D}), \quad (4.11)$$

where W' is a random word independent of the exposed bits. Retaining $k_- \leq D \leq k_+ - 1$ means that the partial degrees first exceed N at a letter in $(k_-, k_+]$; then $D < n$ by (4.10). Split the head offset into the *shells* $I_0 = [0, 1)$ and $I_j = [2^{j-1}, 2^j)$, $1 \leq j \leq j$, of lengths $|I_0| = 1$ and $|I_j| = 2^{j-1}$. The *slices* are the events $D = d$, $O(\{s_1, \dots, s_D\}) \in I_j$, for $k_- \leq d \leq k_+ - 1$ and $0 \leq j \leq j$; they are disjoint. The rest has probability at most

$$4n^{-16} + 7n^{-2}. \quad (4.12)$$

Indeed $D < k_-$ forces $\deg_{k_-}(W) \geq N + 1 \geq 2k_- + 2h$, and $D \geq k_+$ forces $\deg_{k_+}(W) \leq N \leq 2k_+ - 2h - 1$; by (4.6) with $x = 2h$ and $k_+ < n$, whose exponent is at least $\min(h^2/(8n), h/4) \geq 16 \log n$ by (4.10), each has probability at most $2n^{-16}$. And $O(\{s_1, \dots, s_D\}) \geq 2^j \geq n^2$ forces $\text{off}(W_{\leq n}) \geq n^2$, of probability at most $7n^{-2}$ by Lemma 4.3.

For a slice let $\Gamma_{d,j}$ be the measure on G_n of the residue of the head offset, $\Gamma_{d,j}(y) = 2^{-N} \#\{\mathcal{C} : |\mathcal{C}| = d, O(\mathcal{C}) \in I_j, [O(\mathcal{C})]_n = y\}$, and let $p_{d,j} = \|\Gamma_{d,j}\|_1$ and $p_d = \mathbb{P}(D = d)$. By (4.11), the restriction of μ_n to the slice is $\Gamma_{d,j} * \mathcal{L}_d$, where the *tail law* $\mathcal{L}_d$ is the law of $3^d 2^{-N} [\text{off}(W'_{\leq n-D})]_n$.

Lemma 4.5 (Head packing). *For every slice, $3^n \|\Gamma_{d,j}\|_\infty < |I_j| + 2$ and $3^n \sum_y \Gamma_{d,j}(y)^2 \leq (|I_j| + 2)p_{d,j}$.*

Proof. Distinct d -subsets $\mathcal{C}$ have distinct offsets $O(\mathcal{C})$, since offsets determine words of a fixed length [9, Lemma 6.2]. So the integers $2^N O(\mathcal{C})$ in one slice are distinct and lie in an interval of length $2^N |I_j|$, and at most $2^N |I_j| 3^{-n} + 1$ of them lie in one residue class modulo 3^n . Hence $3^n \Gamma_{d,j}(y) \leq |I_j| + 3^n 2^{-N} < |I_j| + 2$, as $3^n < 2^{N+1}$. The second bound follows by multiplying by $\Gamma_{d,j}(y)$ and summing. $\square$

The shell of the head offset is controlled by a tail bound for the offset of a uniform random subset, the conditional form of Lemma 4.3.

Lemma 4.6 (Offsets of random subsets). *Let $\mathcal{R}$ be a uniform d -subset of $\{1, \dots, N'\}$.*

- (1) *If $N' \geq 2$ and $100(d-1) \leq 51(N'-1)$, then $\mathbb{E}\sqrt{O(\mathcal{R})} \leq 40$.*
- (2) *If $3n \leq 2N'$, $|2d - N'| \leq 2h + 4$ and $0 < x \leq 2n^2$, then $\mathbb{P}(O(\mathcal{R}) \geq x) < 8/x$.*

Proof. For (1), let $p = (d-1)/(N'-1)$. For $s \leq N'$, given $s \in \mathcal{R}$, the number X_s of elements of $\mathcal{R}$ below s has factorial moments $\mathbb{E}\binom{X_s}{k} = \binom{s-1}{k} \frac{(d-1)\dots(d-k)}{(N'-1)\dots(N'-k)} \leq \binom{s-1}{k} p^k$, so $\mathbb{E}3^{X_s/2} \leq (1 + (\sqrt{3} - 1)p)^{s-1}$. By subadditivity of the square root, $\sqrt{O(\mathcal{R})} \leq \sum_{s \in \mathcal{R}} 3^{X_s/2} 2^{-s/2}$, and $\mathbb{P}(s \in \mathcal{R}) = d/N' \leq 1$, so $\mathbb{E}\sqrt{O(\mathcal{R})} \leq 2^{-1/2} \sum_{s \geq 1} \mathfrak{x}^{s-1}$ with $\mathfrak{x} = 2^{-1/2}(1 + (\sqrt{3} - 1)p) \leq \frac{49+51\sqrt{3}}{100\sqrt{2}} < \frac{39}{40}$.

For (2), put $q = \lceil 2048 \log n \rceil$, and expose the membership of $1, \dots, q$ in $\mathcal{R}$ one by one. Each conditional probability is at most $(N'/2 + h + 2)/(N' - q)$, so every pattern of these q memberships has probability at most $2^{-q}(1 + \frac{q+2h+4}{N'-q})^q < \frac{100}{99} 2^{-q}$, the exponent $q(q+2h+4)/(N'-q)$ being below $\frac{1}{100}$ on $n \geq n_0$. Write $O(\mathcal{R}) = O_e + 3^{\mathfrak{k}} 2^{-q} O_f$, where O_e collects the $\mathfrak{k}$ elements of $\mathcal{R}$ at most q and O_f is the offset of the rest, shifted by q . For fair bits O_e is the offset of at most q letters of a random word, so the pattern bound and Lemma 4.3 give $\mathbb{P}(O_e \geq \frac{99}{100}x) \leq \frac{100}{99} \cdot \frac{7}{99x/100} = \frac{70000}{9801x}$. Given the exposed memberships, the rest of $\mathcal{R}$ is a uniform subset of $q+1, \dots, N'$ satisfying the hypothesis of (1), so $\mathbb{E}\sqrt{O_f} \leq 40$; the factorial moments of $\mathfrak{k}$ give, as in (1), $\mathbb{E}(3^{\mathfrak{k}/2} 2^{-q/2}) \leq (39/40)^q \leq n^{-256/5}$. By Markov's inequality and $x \leq 2n^2$, $\mathbb{P}(3^{\mathfrak{k}} 2^{-q} O_f \geq \frac{x}{100}) \leq 400n^{-256/5} x^{-1/2} < \frac{1}{100x}$. Finally $\frac{70000}{9801} + \frac{1}{100} < 8$. $\square$

Lemma 4.7 (Local mixing). *Let $\iota \in \{o, b\}$ and let $n_0 \leq m \leq n$ be integers with $25n \leq 27m$. Then*

$$\text{Osc}_{n,m}(\mu_n) \leq \mathfrak{M}_\iota^{\text{loc}}(n), \quad \mathfrak{M}_\iota^{\text{loc}}(x) = 18B_F(\iota)x^{-a_\iota} \log x + 15x^{-2}.$$

Proof. Let $\xi \in G_n$ not factor through G_m , of order 3^s in G_n , so $s \geq m+1$. The Fourier coefficient of $\mathcal{L}_d$ at ξ is that of μ_{n-d} at $2^{-N}\xi$ reduced into G_{n-d} , of order 3^{s-d} ; by projectivity it is $\hat{\mu}_{s-d}(\xi')$ for some ξ' prime to 3. For $d \leq k_+ - 1$, (4.10) and $m \geq \frac{25}{27}n$ give $s-d > \frac{2n}{15}$, so (4.2) bounds it by $(15D_F(\iota)/(2n))^{A_\iota} = B_F(\iota)n^{-A_\iota}$. Lemmas 4.4 and 4.5 give

$$\text{Osc}_{n,m}(\Gamma_{d,j} * \mathcal{L}_d) \leq B_F(\iota)n^{-A_\iota} \sqrt{(|I_j| + 2)p_{d,j}}.$$

For $k_- \leq d \leq k_+ - 1$, $N \geq \frac{3n}{2}$ and $|2d - N| \leq 2h + 4$, so Lemma 4.6(2) with $x = 2^{j-1} \leq 2n^2$ gives $p_{d,j} \leq p_d q_j$ with $q_0 = 1$ and $q_j = \min(1, 2^{4-j})$. The values $\sqrt{(|I_j| + 2)p_{d,j}}$ for $j \leq 4$ are $\sqrt{3}, \sqrt{3}, 2, \sqrt{6}, \sqrt{10}$, with sum below 12, and for $j \geq 5$ they are at most 3; as $j \geq 4$, their sum over $j \leq j$ is below $3j$. With the weights $\varpi_d = 1 + 4(d - N/2)^2/N$, whose mean under the binomial law is 2, and $\sum_{d \in \mathbb{Z}} \varpi_d^{-1} \leq 2 + \frac{\pi}{2}\sqrt{N}$ (comparison with the integral), Cauchy-Schwarz gives $(\sum_d \sqrt{p_d})^2 \leq 4 + \pi\sqrt{N} < 4\sqrt{n}$. Summing over the slices, the retained part of μ_n has oscillation at most

$$B_F(\iota)n^{-A_\iota} \cdot 2n^{1/4} \cdot 3j < 18B_F(\iota)n^{-a_\iota} \log n,$$

by (4.10). Fiber averaging is an ℓ^1 contraction, so the rest adds at most twice its probability, $8n^{-16} + 14n^{-2} \leq 15n^{-2}$, by (4.12). $\square$

Proof of Proposition 4.1. Fix ι and write $a = a_\iota$, $c = c_{\text{ch}}$, $\mathfrak{z} = \mathfrak{z}_\iota$, $\mathfrak{c} = \mathfrak{c}_\iota$, $\mathfrak{d} = \mathfrak{d}_\iota$, $\mathfrak{M} = \mathfrak{M}_\iota$ and $\mathfrak{M}^{\text{loc}} = \mathfrak{M}_\iota^{\text{loc}}$. As $1 \leq a \leq 2$, $1 < \mathfrak{z} \leq c^2 < 2$, so $\mathfrak{c} \geq 1$, $\mathfrak{d} > 0$ and $\mathfrak{s} \geq 15$; hence $\mathfrak{M}^{\text{loc}} \leq \mathfrak{M}$ on $[1, \infty)$. For $x \geq n_0$, $a \log x \geq 1$, and then $x^{-a} \log x$ is nonincreasing (its logarithmic derivative is $(1 - a \log x)/x$); so $\mathfrak{M}$ and $\mathfrak{M}^{\text{loc}}$ are nonincreasing on $[n_0, \infty)$. Since $(cx)^{-a} = x^{-a}/\mathfrak{z}$, $\log(cx) = \log x + \log c$, $(\mathfrak{c} + 1)/\mathfrak{z} = \mathfrak{c}$, $((\mathfrak{c} + 1) \log c + \mathfrak{d})/\mathfrak{z} = \mathfrak{d}$ and $(\mathfrak{s} + 15)/c^2 = \mathfrak{s}$, the profile satisfies the exact identity

$$\mathfrak{M}(cx) + \mathfrak{M}^{\text{loc}}(cx) = \mathfrak{M}(x) \quad (x > 0).$$

We prove $\text{Osc}_{n,m}(\mu_n) \leq \mathfrak{M}(m)$ for $n_0 \leq m \leq n$ by induction on $n - m$; this is (4.4). Let $q = \lfloor 27m/25 \rfloor$; as $m \geq 1000$, $q \geq cm > m$. If $n \leq q$, Lemma 4.7 gives $\text{Osc}_{n,m}(\mu_n) \leq \mathfrak{M}^{\text{loc}}(n) \leq \mathfrak{M}^{\text{loc}}(m) \leq \mathfrak{M}(m)$. Otherwise (4.8) with $r = q$, the induction hypothesis for (n, q) , Lemma 4.7 for (q, m) and monotonicity give

$$\text{Osc}_{n,m}(\mu_n) \leq \mathfrak{M}(q) + \mathfrak{M}^{\text{loc}}(q) \leq \mathfrak{M}(cm) + \mathfrak{M}^{\text{loc}}(cm) = \mathfrak{M}(m).$$

For (4.5) let $x \geq n_0$, so $\log x \geq 80 \log 2 > \frac{160}{3}$. For $\iota = \text{o}$: the series $\log \frac{1+y}{1-y} = 2 \sum_{k \geq 0} y^{2k+1}/(2k+1)$ at $y = \frac{79}{2079}$ and $\log(1+y) \leq y$ give $\frac{158}{2079} \leq \log c \leq \frac{79}{1000}$, so $a_o \log c \geq \frac{71}{500}$ and $\mathfrak{z}_o \geq 1 + \frac{71}{500} + \frac{1}{2}(\frac{71}{500})^2 > \frac{23}{20}$. Hence $\mathfrak{c}_o \leq \frac{20}{3}$, and since $\mathfrak{z}/(\mathfrak{z}-1)^2$ decreases for $\mathfrak{z} > 1$, $\mathfrak{d}_o \leq \frac{79}{1000} \cdot \frac{460}{9} = \frac{1817}{450}$. So

$$18(\mathfrak{c}_o \log x + \mathfrak{d}_o) \leq 18\left(\frac{20}{3} + \frac{3}{160} \cdot \frac{1817}{450}\right) \log x = \left(120 + \frac{49059}{36000}\right) \log x < 126 \log x,$$

and $\mathfrak{s} \leq 105$ as $15 \cdot 10^6 \leq 105 \cdot 164241$. Then $x^{-2} \leq x^{-a_o} \leq x^{-a_o} \log x$ and $126B_F(\text{o}) + 105 < 2^{333} \leq 2^{365}$. For $\iota = \text{b}$: the same series and the partial sum of order 8 of the exponential series give $0.076034686275 < \log c < 0.076034686276$ and $\mathfrak{z}_b > 1.116557504$, hence $\mathfrak{c}_b < 8.5794562$ and $\mathfrak{d}_b < 6.24902664$; so $\mathfrak{c}_b \log x + \mathfrak{d}_b \leq (\mathfrak{c}_b + \frac{3}{160}\mathfrak{d}_b) \log x < 8.7 \log x$. With $B_F(\text{b}) < 2^{88.12} < 1.09 \cdot 2^{88}$ and $\mathfrak{s} < 92$, $\mathfrak{M}_b(x) \leq (18 \cdot 8.7 \cdot 1.09 \cdot 2^{88} + 92)x^{-a_b} \log x < 2^{96}x^{-a_b} \log x$. $\square$

4.4. 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 $\iota \in \{\text{o}, \text{b}\}$ and $d, k, m \in \mathbb{N}$, 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*

$$\left\langle \left| \sum_{w \in \mathcal{V}} \mathcal{T}_{w,Q} \rho_k - \rho_m \circ \text{pr}_{Q,m} \right| \right\rangle_Q \leq \mathfrak{M}_\iota(k) + \mathfrak{M}_\iota(m) + 1 - \text{mass}(\mathcal{V}).$$

Proof. This is [7, (3.7)] in our normalization. Put $A^* = \sum_{w \in \mathcal{V}} \mathcal{T}_w \rho_{Q-|w|}$. Since $Q - |w| \geq k$, (4.4) bounds $\langle |\rho_k \circ \text{pr}_{Q-|w|,k} - \rho_{Q-|w|}| \rangle_{Q-|w|}$ by $\mathfrak{M}_\iota(k)$, and (3.6) multiplies it by $2^{-\deg(w)}$; summing, the first sum differs from A^* by at most $\text{mass}(\mathcal{V})\mathfrak{M}_\iota(k) \leq \mathfrak{M}_\iota(k)$ in mean. By (3.2) and the disjointness of the events, A^* is 3^Q times the restriction of μ_Q to the event that W begins with a word of $\mathcal{V}$, so $0 \leq A^* \leq \rho_Q$ and $\langle \rho_Q - A^* \rangle_Q = 1 - \text{mass}(\mathcal{V})$. Finally (4.4) bounds $\langle |\rho_Q - \rho_m \circ \text{pr}_{Q,m}| \rangle_Q$ by $\mathfrak{M}_\iota(m)$. $\square$

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 . With $e_\delta = \lceil \delta/100 \rceil$, the scales are $\delta_0 = 16544$ and $\delta_{j+1} = \delta_j + e_{\delta_j}$, and the caps are

$$R_j = 29 + \lfloor j/56 \rfloor.$$

The first scale is large enough that the first block family is nearly complete (Lemma 15.8); the increments make the sources of generation j exceed 16^{δ_j} (Lemma 5.3); and the caps grow so that the losses $2^{-(R_j+1)}$ decay geometrically (15.1). Put $g = 101/100$; the scales grow geometrically and linearly (Lemma 15.1(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 *levels* are $k_j = \lfloor \delta_j/4 \rfloor$. Put $R_\Sigma = \sum_{i < j} (R_i + 3)$, suppressing j ; it bounds the spread of the sources of a generation (Lemma 16.1). Since $R_i + 3 \leq 32 + i/56$, $R_\Sigma + 1 \leq 1 + 32j + j(j-1)/112$, and $j^2/82$ exceeds the right side by $j(\frac{15}{4592}j - \frac{3583}{112}) - 1$, which is positive for $j \geq 9853$ (at $j = 9853$ the bracket is $\frac{223}{1148}$). Hence

$$R_\Sigma + 1 < \frac{j^2}{82} \quad (j \geq 9853). \quad (5.1)$$

For an integer scale $\delta \geq 256$, 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 $\delta \geq 256$, 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 15.7); blocks use $u = r_\delta$, the terminal extension all $u \in [0, 2r_\delta]$.

Lemma 5.1 (Shift radius). *For $\delta \geq 256$, $\frac{229}{1000}\delta - 4 < r_\delta \leq \frac{23}{100}\delta$.*

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

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

□

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* $\text{FC}(\delta, R)$, from which the blocks are drawn, is the subfamily of the first crossings of Section 2, that is of $\text{FC}_{r_\delta}(\delta, R)$, with length in $[\delta - \ell_\delta^c, \delta + \ell_\delta^c]$ (the *length restriction*), where

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

it narrows the length spread for the histogram estimate (5.8) and the step bounds at a deficit of order $\delta^{-3/2}$ (Lemma 15.6). For $\delta \geq 2^{40}$,

$$20000 \ell_\delta^c \leq \delta. \quad (5.3)$$

Indeed, with $k = \lfloor \log_2 \delta \rfloor \geq 40$, $\lceil \log_2 \delta \rceil \leq k + 1$, $2^k \leq \delta$ and $(k + 1)/2^k$ nonincreasing,

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

the last being the rational comparison $1312/2^{40} \leq (1/20000 - 2^{-40})^2$.

Lemma 5.2 (Weight, offset and source of a first-crossing word). *Let $\delta \geq 256$, 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 $u = r_\delta$, then $16^{e_\delta} \text{wt}(w) < \frac{1}{2}$;
- (4) for every odd $n \geq 16^\delta$,

$$2^{u-r_\delta-2}(4/3)^\delta n < \text{src}(w, n) < 2^{R+1+u-r_\delta}(4/3)^\delta n. \quad (5.4)$$

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 [7, 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), $e_\delta \geq 3$ 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 [7, (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 block, since later blocks enter with tiny weights. The *selector* bounds the offset of the first block, which sharpens the histogram estimate (5.8), and the offsets of the next 28 blocks, which are otherwise only bounded by 2^{δ_i} : it retains w if

$$\text{off}(v_0) \leq \frac{3}{4} \quad (j \geq 1), \quad \text{off}(v_i) \leq 2^{600} \quad (0 < i < \min(j, 29)). \quad (5.5)$$

These words form the set $\bar{L}_j$ of *selected words*; $\bar{L}_0$ consists of the empty word. The selector involves only the first 29 blocks, so for $j \geq 29$ 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

$$0 \leq \text{off}(w) \leq \frac{3}{4} + 2^{-8} \quad (5.6)$$

(Lemma 16.2). Lemma 6.6 shows that the selector keeps a large mass on a residue class of the first eight letters.

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)

$$s_{w_{<i+1}} = \frac{s_{w_{<i}} - \text{off}(v_i)}{\text{wt}(v_i)} \quad (i < j). \quad (5.7)$$

Lemma 5.3 (Growth of sources). *Let $t \geq 16^{16544}$ 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 $t \geq 16^{\delta_0}$. If $s_{w_{<i}} \geq 16^{\delta_i}$, then Lemma 5.2(2) gives $0 < \text{off}(v_i) < 2^{\delta_i} \leq s_{w_{<i}}/2$, and by (5.7) and Lemma 5.2(3),

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

The third claim is (5.4) with $u = r_{\delta_i}$ and $R = R_i$. $\square$

The *weighted density* is

$$Z_j(t) = \sum_{w \in L_j} \text{wt}(w) \rho_{k_j}(s_w).$$

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 $j \geq 9853$, $0 \leq q \leq 2\delta_j$, any $f : G_q \rightarrow \mathbb{R}$, and any subset of L_j ,

$$\sum_w \text{wt}(w) |f(s_w)| \leq 344 j^{5/2} g^{j/2} \langle |f| \rangle_q. \quad (5.8)$$

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 $\delta \geq 256$ and a real $n > 0$, set, as in [7, 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 (5.4), 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*

$$2|v| \leq H_{\delta_j, u_w}(|v|) + \lfloor \delta_j/8 \rfloor, \quad (5.9)$$

the terminal restriction of [7, (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 $t \geq 16^{16544}$ and X is reachable at j , these sources lie in $[X, 32X)$ [7, Lemma 6.1], since $s_w \geq 16^{\delta_j}$ (Lemma 5.3): the factor 2 comes from the rounding of u_w , and the factor 16 from overshoot at most 1 and the slack 3 above. The words wv are valid at t and distinct, since $v_0 \cdots v_{j-1}v$ determines $v_0, \dots, v_{j-1}$ and v as for the blocks of L_j ; so $\beta_{j,X}^{\text{term}}(t) \leq \beta_X(t)$. If moreover t is not S -periodic, their sources are distinct (Lemma 3.3) and $\text{wt}(wv) \leq t/X$, so the argument of (3.3) gives

$$X \beta_{j,X}^{\text{term}}(t) \leq t \# \text{Src}_{j,X}(t). \quad (5.10)$$

6. KEY ESTIMATES

This section proves the key estimates of Section 2: propagation (Proposition 6.1, with the geometric form (6.5) and the tail Lemma 6.3), the average at generation j_0 (Proposition 6.7), reachability (Proposition 6.9), the terminal error (Proposition 6.11) and the census inequality (6.12) with its short form (6.13). Apart from the average and the census, the arguments are Mazur's [7, Proposition 3.1, Sections 4–6 and 8]. Throughout, t is a positive odd integer, and

$$j_0 = 9853, \quad J_c = 23000 : \quad (6.1)$$

propagation runs from j_0 , and the count uses the generations $j \geq J_c$.

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. For $j \geq j_0$ the level is large enough for Proposition 4.1 and the block family has small deficit:

$$k_j \geq \max(n_0, 2^{12}g^j), \quad 1 - \text{mass}(\text{FC}(\delta_j, R_j)) \leq \frac{1}{2}g^{-3j/2} + 2^{-29-j/56}. \quad (6.2)$$

The first bound is Lemma 15.1(2) and the second is Lemma 15.9. Define, with the bootstrap profile $\mathfrak{M}_b$ of (4.3),

$$E_j = 344 j^{5/2} g^{j/2} \left(2\mathfrak{M}_b(2^{12}g^j) + \frac{1}{2}g^{-3j/2} + 2^{-29-j/56} \right). \quad (6.3)$$

Proposition 6.1 (Propagation). *For every $j \geq j_0$,*

$$|Z_{j+1}(t) - Z_j(t)| \leq E_j. \quad (6.4)$$

Proof. This is [7, Proposition 3.1] with explicit constants. Put $Q = d_{\delta_j}^{\text{hi}} + k_{j+1}$. Since $\delta_{j+1} \leq (51/50)\delta_j$, $Q \leq (8/5 + 51/200)\delta_j < 2\delta_j$. 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 $j \geq 29$, 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 with $\iota = b$ at the levels $k_{j+1}, k_j \geq 2^{12}g^j \geq n_0$, the monotonicity of $\mathfrak{M}_b$ and (6.2) bound $\langle |\mathcal{D}_j| \rangle_Q$ by the bracket in (6.3). Apply (5.8). $\square$

The majorant decays geometrically from j_0 on.

Corollary 6.2 (Geometric form of propagation). *For every $j \geq j_0$,*

$$E_j \leq 0.007584 \cdot 0.990945^{j-j_0} + 2^{-90} \left(\frac{199}{200} \right)^{j-j_0}. \quad (6.5)$$

Proof. Write $a = a_b = \frac{29}{20}$ and $\mathfrak{y} = g^{1/2}2^{-1/56}$. Inserting (4.3) with $x = 2^{12}g^j$, $\log x = 12 \log 2 + j \log g$, into (6.3) splits E_j into three parts. The first is

$$E_j^{\text{mx}} = 36 \cdot 344 B_F(b) 2^{-12a} j^{5/2} g^{-(a-\frac{1}{2})j} (\mathfrak{c}_b(12 \log 2 + j \log g) + \mathfrak{d}_b).$$

The second, from $\mathfrak{s}x^{-2}$ and $\frac{1}{2}g^{-3j/2}$, is at most $344(\mathfrak{s}/2^{23} + \frac{1}{2})j^{5/2}g^{-j} < 2^8 j^{5/2}g^{-j}$. The third is at most $2^{-20}j^{5/2}\mathfrak{y}^j$, as $344 < 2^9$. We bound each by a geometric sequence from j_0 ; the bounds below for the ratios of consecutive terms decrease in j , so it suffices to check them at j_0 .

The ratio $E_{j+1}^{\text{mx}}/E_j^{\text{mx}}$ is at most $(1 + \frac{1}{j})^{7/2}g^{-19/20}$, because the affine factor grows by at most the factor $1 + \frac{1}{j}$. As $\frac{7}{2} \log(1 + \frac{1}{j_0}) \leq \frac{7}{19706}$ and $\log g > 0.009950330853168$, this is at most $e^{-9097/10^6} < 0.990945$. Evaluating $E_{j_0}^{\text{mx}}$ with $\log_2 B_F(b) < 88.1188851177$, $\mathfrak{c}_b < 8.5794562$, $\mathfrak{d}_b < 6.24902664$

(proof of Proposition 4.1) and rational enclosures of $\log 2$, $\log g$ and e^x from partial sums of their series gives $E_{j_0}^{\max} < 0.007583$.

For the second part the ratio is $(1 + \frac{1}{j})^{5/2} \frac{100}{101} \leq 0.990945$, by the integer comparison $(j_0 + 1)^5 \cdot 10^{16} \leq j_0^5 \cdot 100085445^2$; and $2^8 j_0^{5/2} g^{-j_0} < 2^{8+34-140} < 10^{-6}$, by $j_0^5 < 2^{68}$ and $g^{j_0} > 2^{140}$.

For the third part the ratio is $(1 + \frac{1}{j})^{5/2} \eta \leq \frac{199}{200}$, by the comparison of 56th powers $(j_0 + 1)^{140} \cdot 101^{28} \cdot 200^{56} \leq 2 \cdot j_0^{140} \cdot 100^{28} \cdot 199^{56}$. As $\eta < 1$, $\eta^{56} = g^{28}/2$ and $j_0 \geq 56 \cdot 175$, $\eta^{j_0} \leq (g^{28}/2)^{175} \leq 2^{-104}$, by $101^{4900} \cdot 2^{104} \leq 2^{175} \cdot 100^{4900}$; so $2^{-20} j_0^{5/2} \eta^{j_0} < 2^{-20+34-104} = 2^{-90}$. Adding the three geometric bounds gives (6.5). $\square$

Lemma 6.3 (Propagation tail). $\sum_{j \geq j_0} E_j < \frac{7584}{9055} + 2^{-82} < 0.8376 < \frac{6}{7}$. Consequently, if $Z_{j_0}(t) > \frac{6}{7} + 2^{-28}$, then $Z_j(t) > 2^{-28}$ for every $j \geq j_0$.

Proof. Sum the two geometric series of (6.5): $0.007584/(1 - 0.990945) = \frac{7584}{9055}$ and $2^{-90} \cdot 200 < 2^{-82}$. The consequence follows by summing (6.4) from j_0 to $j - 1$. $\square$

6.2. 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

$$d_0 = 5 \cdot 2^{160} \geq 1668728 g^{j_0}, \quad (6.6)$$

by the integer comparison $1668728 \cdot 101^{j_0} \leq 5 \cdot 2^{160} \cdot 100^{j_0}$. So $2^{162} < d_0$ and, as $317 \cdot 5 < 1600$,

$$4 + \frac{317}{200} d_0 < 2^{163}. \quad (6.7)$$

The mean of Z_{j_0} on a residue class modulo 3^8 is the selected mass with a prescribed offset residue of the first eight letters, which is computed through the joint law of the offset and that residue.

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 i has length at most $\delta_i + \ell_{\delta_i}^c \leq \frac{8}{5} \delta_i$, and at most $\frac{20001}{20000} \delta_i$ if $\delta_i \geq 2^{40}$, by (5.3). By Lemma 15.1(1), $\delta_i \leq g^{-(i'-i)} \delta_{i'}$ for $i \leq i'$, so the scales below 2^{40} sum to less than $2^{40} \sum_{k \geq 0} g^{-k} = 101 \cdot 2^{40}$, and all scales δ_i , $i < j_0$, sum to less than $100 \delta_{j_0}$. As $\delta_{j_0} \geq 16544 g^{j_0} > 2^{55} > 200 \cdot \frac{8}{5} \cdot 101 \cdot 2^{40}$ and $k_{j_0} \leq \delta_{j_0}/4$, the length plus k_{j_0} is less than

$$\left(\frac{1}{4} + 100 \cdot \frac{20001}{20000} + \frac{1}{200} \right) \delta_{j_0} = \frac{20052}{200} \delta_{j_0} < \frac{20052}{200} \cdot 16644 g^{j_0} \leq 1668728 g^{j_0} \leq d_0.$$

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 $Z_{j_0}(t)$. $\square$

For integers $D \geq q \geq 0$, $r \in G_q$ and real x , the joint law of the offset and the offset residue of the first q letters is

$$F_{D,q}(r, x) = \mathbb{P}(\text{off}(W_{\leq D}) \leq x, [\text{off}(W_{\leq q})]_q = r). \quad (6.8)$$

Lemma 6.5 (Joint law). For every $D \geq 8$, $F_{D,8}([6275]_8, \frac{3}{4}) \geq 617014680840434 \cdot 2^{-62}$.

Proof. Conditioning on the first letter a and writing $W = (a)W'$, (3.2) gives $\text{off}(W_{\leq D}) = 2^{-a}(1 + 3 \text{off}(W'_{\leq D-1}))$, and for $q \geq 1$, $[\text{off}(W_{\leq q})]_q = r$ holds if and only if $3 \mid 2^a \tilde{r} - 1$ and $[\text{off}(W'_{\leq q-1})]_{q-1}$ is the residue of $(2^a \tilde{r} - 1)/3$, where $\tilde{r}$ is any integer representing r . Hence

$$F_{D,q}(r, x) = \sum_{a \geq 1, 3 \mid 2^a \tilde{r} - 1} 2^{-a} F_{D-1,q-1}\left(\left[\frac{2^a \tilde{r} - 1}{3}\right]_{q-1}, \frac{2^a x - 1}{3}\right),$$

and the level $q = 0$ satisfies the same recursion without the residue condition and $F_{D,0}(0, x) \geq \max(0, 1 - 7/(x + 2))$ by Lemma 4.3. Lemma A.10 gives tables of integers, read as nonnegative step functions of the offset on a finite grid, times 2^{-62} : a nondecreasing one at level 0, at most the maximum of $1 - 7/(x + 2)$ and the right side of the recursion evaluated on it, and ones at levels $1 \leq q \leq 8$, at most the right side of the recursion evaluated on the tables of level $q - 1$. Induction on D , using the monotonicity of $F_{D,q}$ in x , shows that the tables are below the joint laws for every D ; the entry of the level-8 table at $([6275]_8, \frac{3}{4})$ is 617014680840434. $\square$

Lemma 6.6 (Selected mass in a residue class). *For every $j \geq 29$, the mass of the words of $\bar{L}_j$ whose first eight letters have offset residue $[6275]_8$ exceeds*

$$\mathbf{m}_8 = \left(\frac{617014680840434}{2^{62}} - \frac{1}{20000000} \right) \left(\frac{199}{200} - 2^{-592} \right).$$

Proof. Every block of scale δ_0 has length greater than $d_{16544}^{\text{lo}} = 6618 \geq 8$, so the offset residue of the first eight letters is that of v_0 . The degree is additive under concatenation and the conditions of (5.5) concern separate blocks, so the mass in question is $m'_0 \prod_{0 < i < j} m_i$, where m'_0 is the mass of the $v \in \text{FC}(\delta_0, R_0)$ with $\text{off}(v) \leq \frac{3}{4}$ and $[\text{off}(v_{\leq 8})]_8 = [6275]_8$, m_i is the mass of the $v \in \text{FC}(\delta_i, R_i)$ with $\text{off}(v) \leq 2^{600}$ for $0 < i < 29$, and $m_i = \text{mass}(\text{FC}(\delta_i, R_i))$ for $i \geq 29$. Offsets of prefixes increase, and the events that W begins with a word of a prefix-disjoint family are disjoint. So if $D = d_{\delta_0}^{\text{hi}}$, $\text{off}(W_{\leq D}) \leq \frac{3}{4}$, $[\text{off}(W_{\leq 8})]_8 = [6275]_8$ and W begins with some $v \in \text{FC}(\delta_0, R_0)$, then v is counted in m'_0 ; by Lemmas 6.5 and 15.8,

$$m'_0 \geq F_{D,8}([6275]_8, \frac{3}{4}) - (1 - \text{mass}(\text{FC}(\delta_0, R_0))) > \frac{617014680840434}{2^{62}} - \frac{1}{20000000}.$$

Likewise Lemma 4.3 gives $m_i \geq \bar{m}_i - c_i$ with $\bar{m}_i = \text{mass}(\text{FC}(\delta_i, R_i))$, $c_i = 7/(2^{600} + 2)$ for $0 < i < 29$ and $c_i = 0$ otherwise. As $0 \leq m_i \leq \bar{m}_i \leq 1$, telescoping gives $\prod_{0 < i < j} \bar{m}_i - \prod_{0 < i < j} m_i \leq \sum_i c_i = 196/(2^{600} + 2) < 2^{-592}$, and $\prod_{0 < i < j} \bar{m}_i \geq \prod_{i < j} \bar{m}_i > \frac{199}{200}$ by Lemma 15.10. $\square$

Proposition 6.7 (Average at generation j_0). *The mean of Z_{j_0} over the residues $y \in G_{d_0}$ with $y \equiv 6275 \pmod{3^8}$ is at least $6561\mathbf{m}_8 > \frac{13}{15} + 2^{-28}$. In particular some $y_0 \in G_{d_0}$ has $Z_{j_0}(y_0) > \frac{13}{15}$.*

Proof. Let $w \in \bar{L}_{j_0}$, so $|w| \geq 8$. By (3.2), $\text{wt}(w) \equiv 0$ and $\text{off}(w) \equiv \text{off}(w_{\leq 8}) \pmod{3^8}$ in $\mathbb{Z}[1/2]$, so the image of σ_w lies in the residue class of $\text{off}(w_{\leq 8})$ modulo 3^8 , and $\mathcal{T}_{w,d_0}\rho_{k_{j_0}}$ vanishes off it. By (3.6) and $\langle \rho_k \rangle_k = 1$, its sum over G_{d_0} is $3^{d_0} 2^{-\deg(w)}$. By Lemma 6.4, the sum of Z_{j_0} over the class of 6275 is therefore 3^{d_0} times the mass in Lemma 6.6, which exceeds $3^{d_0}\mathbf{m}_8$ as $j_0 \geq 29$. The class has 3^{d_0-8} elements, and $6561\mathbf{m}_8 - \frac{13}{15} - 2^{-28} > 0.0064$ by Lemma A.10. $\square$

6.3. Reachability. 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), which (5.1) bounds.

Lemma 6.8 (Consecutive generations overlap). *Let $t \geq 16^{16544}$ be odd and $j \geq J_c$. 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. Let $k = \lfloor j/70 \rfloor \geq 328$. By Lemma 15.1(1) and $g^{70} > 2$, $\delta_j \geq 16544 \cdot 2^k \geq 16544 \binom{k}{3} \geq 148896k^2$, as $(k-1)(k-2) \geq 54k$; and $j < 71k$, so $\frac{37}{1000}\delta_j \geq 5041k^2 > j^2$. Also $\delta_j \leq \delta_{j+1} \leq \frac{101}{100}\delta_j + 1$. Lemmas 5.3 and 16.1 give $s_w < 2^{R_j+1}(4/3)^{\delta_j} s_{w_{<j}} \leq 2^{R_j+1+R_\Sigma}(4/3)^{\delta_j} s_{w'}$. 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, by (5.1), $R_j + R_\Sigma + 10 < 38 + \frac{j}{56} + \frac{j^2}{82} \leq j^2$. $\square$

Proposition 6.9 (Reachability). *Let $t \geq 16^{16544}$ be odd. Every real $X \geq 2^{2^{352}}t$ is reachable at some generation $j \geq J_c$.*

Proof. We follow [7, Lemma 6.2] in the effective form of [7, Section 8.5]. Put $\text{gr}(\delta) = 2\delta - \lfloor 19\delta/12 \rfloor$, $\mathfrak{l}_0 = 0$ and $\mathfrak{l}_{j+1} = \mathfrak{l}_j + R_j + 1 + \text{gr}(\delta_j)$ (cf. [7, Section 8.2]). The inequality $2^{19} \leq 3^{12}$ gives $(4/3)^\delta \leq 2^{\text{gr}(\delta)}$, so Lemma 5.3 and induction give $s_w \leq 2^{\mathfrak{l}_j}t$ for $w \in L_j$. By Lemma 5.1 and $\text{gr}(\delta) < \frac{5\delta}{12} + 1$, $X_\delta^{\text{hi}}(n) = 2^{r_\delta}(4/3)^\delta n/4 \leq 2^{r_\delta + \text{gr}(\delta) - 2}n$ with $r_\delta + \text{gr}(\delta) - 2 < \frac{97\delta}{150} - 1 < \lfloor 2\delta/3 \rfloor$. Hence every upper end at generation J_c is at most $2^B t$, where $B = \mathfrak{l}_{J_c} + \lfloor 2\delta_{J_c}/3 \rfloor$. Using $R_j + 1 \leq 30 + j/56$, Lemma 15.1(1) and $\sum_{j < J} g^j < 100g^J$, with $J = J_c$,

$$B < 31J + \frac{J^2}{112} + \frac{5}{12} \cdot 1664400 g^J + 11096 g^J < 2^{30} + 2^{20} g^J.$$

As $101^{348} < 2^5 \cdot 100^{348}$, $g^{69} < 2$ and $J = 348 \cdot 66 + 32$, $g^J < 2^{330} \cdot 2$, so $B < 2^{30} + 2^{351} \leq 2^{352}$, and every upper end at generation J_c is below X .

For $i \geq J_c$ and $w \in L_i$, Lemma 5.3 gives $s_w \geq 16^{\delta_i}$, so $X_{\delta_i}^{\text{hi}}(s_w) \geq 2^{\delta_i} \rightarrow \infty$. Let $j \geq J_c$ be least such that every upper end at generation j is at least X . If $j = J_c$, then L_{J_c} is empty and X is reachable vacuously. Otherwise some upper end at generation $j - 1 \geq J_c$ is below X , and by Lemma 6.8 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 6.10 (Few terminal shifts). *Let $t \geq 16^{16544}$ be odd, $j \geq j_0$ and $X > 0$. Then the terminal shifts u_w , $w \in L_j$, take fewer than $j^2/82$ values.*

Proof. We follow [7, 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 $R_\Sigma + 1 < j^2/82$ integers, by (5.1). $\square$

Proposition 6.11 (Terminal error). *Let $t \geq 16^{16544}$ be odd. For $j \geq j_0$ and every X reachable at j ,*

$$|Z_{j,X}^{\text{term}}(t) - Z_j(t)| \leq T_j = 5j^{9/2}g^{j/2} \left(2\mathfrak{M}_o(2^{12}g^j) + (2^{32} + 2^{29})g^{-6j} + 2^{-29-j/56} \right),$$

and for $j \geq J_c$

$$T_j \leq 2^{169} \cdot 0.993^j; \tag{6.9}$$

in particular, for $j \geq J_c$,

$$|Z_{j,X}^{\text{term}}(t) - Z_j(t)| < 2^{-29}. \tag{6.10}$$

Proof. We follow [7, 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 6.10 there are fewer than $j^2/82$ groups. For each shift $0 \leq u \leq 2r_{\delta_j}$ define on G_Q

$$\mathcal{D}_u^{\text{term}} = \sum_{v \in \text{FC}_u(\delta_j, R_j)} \mathcal{T}_{v,Q} \rho_{k_j} - \rho_{k_j} \circ \text{pr}_{Q,k_j}, \quad \tilde{\mathcal{D}}_u^{\text{term}} = \sum_{\substack{v \in \text{FC}_u(\delta_j, R_j) \\ v \text{ fails (5.9)}}} \mathcal{T}_{v,Q} \rho_{k_j}.$$

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 with $\iota = o$ at the levels k_j, k_j and Lemma 15.7 give $\langle |\mathcal{D}_u^{\text{term}}| \rangle_Q \leq 2\mathfrak{M}_o(k_j) + 2^{116}\delta_j^{-6} + 2^{-(R_j+1)}$; also $\tilde{\mathcal{D}}_u^{\text{term}} \geq 0$ and, by (3.6), its mean is the mass of the words failing the mask, at most $2^{113}\delta_j^{-6}$ by the same lemma. Apply (5.8) to each group and sum over fewer than $j^2/82$ shifts; as $344/82 < 5$,

$$|Z_{j,X}^{\text{term}}(t) - Z_j(t)| \leq 5j^{9/2}g^{j/2} \left(2\mathfrak{M}_o(k_j) + (2^{116} + 2^{113})\delta_j^{-6} + 2^{-(R_j+1)} \right).$$

This is at most T_j : $\mathfrak{M}_0(k_j) \leq \mathfrak{M}_0(2^{12}g^j)$ by (6.2) and monotonicity, $\delta_j \geq 2^{14}g^j$ by Lemma 15.1(1), and $R_j + 1 \geq 29 + j/56$.

Let $j \geq J_c$ and $x_j = 2^{12}g^j$. By (4.5), $T_j \leq \tau_1(j) + \tau_2(j) + \tau_3(j) + \tau_4(j)$ with

$$\begin{aligned}\tau_1 &= 1260 j^{9/2} g^{j/2} B_F(o) x_j^{-a_0} \log x_j, & \tau_2 &= 1050 j^{9/2} g^{j/2} x_j^{-2}, \\ \tau_3 &= 5(2^{32} + 2^{29}) j^{9/2} g^{-11j/2}, & \tau_4 &= 5 \cdot 2^{-29} j^{9/2} g^{j/2} 2^{-j/56}.\end{aligned}$$

Each $\tau_i(j+1)/\tau_i(j)$ is below 0.993. Indeed $(1+1/j)^{9/2} \leq (23001/23000)^{9/2}$; $\log x_{j+1}/\log x_j \leq 1+1/j$ since $\log x_j \geq j \log g$; and $g^{1/2-a_0} < \frac{100}{101}$, $g^{-3/2} < \frac{100}{101}$. So the four ratios are at most $\frac{100}{101}(\frac{23001}{23000})^{11/2}$, $\frac{100}{101}(\frac{23001}{23000})^{9/2}$, $(\frac{100}{101})^{11/2}(\frac{23001}{23000})^{9/2}$ and $(\frac{101}{100})^{1/2}2^{-1/56}(\frac{23001}{23000})^{9/2}$, and each is below 0.993; after raising to the powers 2, 2, 2 and 56 these are comparisons of rationals. At $j = J_c$, with $\log_2 B_F(o) < 325.7939550288$ (Lemma A.6(3)), interval arithmetic gives $\tau_1 < 2^{-65.35}$, $\tau_2 < 2^{-444}$, $\tau_3 < 2^{-1716}$ and $\tau_4 < 2^{-207}$, so $T_{J_c} < 2^{-65.3} < 2^{169} \cdot 0.993^{J_c}$, the right side being above $2^{-64.1}$. Hence (6.9) holds for every $j \geq J_c$, and (6.10) follows from $2^{169} \cdot 0.993^{23000} < 2^{-64}$. $\square$

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 a bounded part at a lower level and a remainder Φ of small average, on which distinct sources cannot concentrate. As in [7, Proposition 5.3], it is applied through the maps $\text{lift}_i(s) = 4^i s + (4^i - 1)/3$, $i \geq 0$.

Proposition 6.12 (Census). *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*

$$\sum_{s \in \mathcal{S}} v_s \sum_{i \geq 0} 4^{-i} \mathfrak{w}(\text{lift}_i(s)) \leq \frac{4}{3} U \sum_{s \in \mathcal{S}} v_s + 44t \langle \Phi \rangle_q. \quad (6.11)$$

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 [7, Proposition 5.3]. $\square$

The bounded part is obtained by clipping the residue density: discarding the words of atypical degree or large offset leaves a subdensity bounded by packing, at a cost that is small on average.

Lemma 6.13 (Finite clipping). *Let $n \geq 3$ and $1 \leq m \leq n$ be integers and $B > 0$ real. There is $\underline{\rho}_m : G_m \rightarrow \mathbb{R}$ with*

$$0 \leq \underline{\rho}_m \leq \rho_m, \quad \underline{\rho}_m \leq \frac{B}{2}, \quad \langle \rho_m - \underline{\rho}_m \rangle_m \leq 2n^{-8} + \frac{2^{13} \sqrt{n \log n}}{B}.$$

Proof. Put $\ell_{\text{cl}} = \lceil \sqrt{256m \log n} \rceil + \lceil 64 \log n \rceil$, $l_{\text{cl}} = \lceil \log_2(8 \cdot 3^m/B) \rceil$ and $x_{\text{cl}} = B/(4(2\ell_{\text{cl}} + 1))$, and let $\mathcal{V}$ be the set of words u of length m with $|\deg(u) - 2m| \leq \ell_{\text{cl}}$, $\deg(u) \geq l_{\text{cl}}$ and $\text{off}(u) \leq x_{\text{cl}}$. Let $\underline{\rho}_m(y) = 3^m \mathbb{P}(W_{\leq m} \in \mathcal{V}, [\text{off}(W_{\leq m})]_m = y)$, so $0 \leq \underline{\rho}_m \leq \rho_m$. Fix y and a degree $l \geq l_{\text{cl}}$. The integers $2^l \text{off}(u)$, for the words $u \in \mathcal{V}$ of degree l with offset residue y , are distinct, since offsets determine words of a fixed length [9, Lemma 6.2], lie in $[0, 2^l x_{\text{cl}}]$ and are congruent modulo 3^m ; so there are at most $2^l x_{\text{cl}} 3^{-m} + 1$ of them, each of probability 2^{-l} . Summing over the at most $2\ell_{\text{cl}} + 1$ degrees $l \geq l_{\text{cl}}$,

$$\underline{\rho}_m(y) \leq (2\ell_{\text{cl}} + 1)x_{\text{cl}} + 3^m \sum_{l \geq l_{\text{cl}}} 2^{-l} = \frac{B}{4} + 2 \cdot 3^m 2^{-l_{\text{cl}}} \leq \frac{B}{2}.$$

The mean $\langle \rho_m - \underline{\rho}_m \rangle_m$ is the probability that $W_{\leq m} \notin \mathcal{V}$. By (4.6), whose exponent is at least $8 \log n$ at $x = \ell_{\text{cl}}$, $\mathbb{P}(|\deg_m(W) - 2m| > \ell_{\text{cl}}) \leq 2n^{-8}$. Next $\mathbb{P}(\deg_m(W) < \ell_{\text{cl}}) \leq 2^{\ell_{\text{cl}}-1} \mathbb{E} 2^{-\deg_m(W)} = 2^{\ell_{\text{cl}}-1} 3^{-m} < 8/B$. By Lemma 4.3, $\mathbb{P}(\text{off}(W_{\leq m}) > x_{\text{cl}}) < 7/x_{\text{cl}} = 28(2\ell_{\text{cl}} + 1)/B$, and $2\ell_{\text{cl}} + 1 < 32\sqrt{n \log n} + 128 \log n + 5 \leq 165\sqrt{n \log n}$, as $\log n \geq 1$ and $\log n \leq \sqrt{n \log n}$. Finally $8 + 28 \cdot 165 < 2^{13}$. $\square$

Let $t \geq 16^{16544}$ be odd and not S -periodic, let X be reachable at j , and let $w \in L_j$. 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 $4s' + 1$ [7, proof of Proposition 5.3]; and (v', i) determines v . So $Z_{j,X}^{\text{term}}(t)$ is at most the left side of (6.11) for $\mathfrak{w} = \rho_{k_j}$, $\mathcal{S} = \text{Src}_{j,X}(t)$ and the weights $v_s = \text{wt}(wv') \leq t/X$ 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 (6.12) below is trivial; otherwise reachability, $s_w \geq 16^{\delta_j}$ and $r_{\delta_j} \leq 3\delta_j/10$ give $X > 16^{\delta_j}/(4 \cdot 3^{\delta_j}) > 3^{k_j}$. Let $n \geq n_0$ be an integer and $B > 0$, and put $m = \min(k_j, n)$. Apply Proposition 6.12 with $q = k_j$, $\underline{\mathfrak{w}} = \underline{\rho}_m$ of Lemma 6.13, $U = B/2$ and $\Phi = (\rho_m - \underline{\rho}_m) \circ \text{pr}_{k_j,m} + |\rho_{k_j} - \rho_m \circ \text{pr}_{k_j,m}|$, which is nonnegative with $\rho_{k_j} \leq \underline{\rho}_m \circ \text{pr}_{k_j,m} + \Phi$. The second term of Φ vanishes if $k_j \leq n$, and otherwise has mean at most $\mathfrak{M}_0(n) \leq 2^{365} n^{-a_0} \log n$ by (4.4) and (4.5). Hence, for every integer $n \geq n_0$ and real $B > 0$,

$$Z_{j,X}^{\text{term}}(t) \leq \frac{2}{3} B \beta_{j,X}^{\text{term}}(t) + 44t \left(2n^{-8} + \frac{2^{13} \sqrt{n \log n}}{B} + 2^{365} n^{-a_0} \log n \right). \quad (6.12)$$

With n and B chosen from t , the error term becomes negligible:

$$Z_{j,X}^{\text{term}}(t) \leq 2^{177} t^{14/11} \beta_{j,X}^{\text{term}}(t) + 2^{-32}. \quad (6.13)$$

Indeed, write $a = a_0$ and $\mathfrak{f} = 1 + \log t$, and take $n = \lceil x_t \rceil$ with $x_t = (2^{420} t \mathfrak{f})^{1/a} \geq 2^{420/a} > n_0$, and $B = 2^{52} t \sqrt{n \log n}$. Then $n < 2x_t$ and $\log n \leq \log 2 + \frac{1}{a}(420 \log 2 + \log t + \log \mathfrak{f}) \leq 256\mathfrak{f}$, using $\log \mathfrak{f} \leq \log t$ and $1 + \frac{420}{a} < 256$. So $2^{365} n^{-a} \log n \leq 2^{365+8-420}/t$, $2n^{-8} \leq 2/x_t^8 \leq 2^{-419}/t$ and $2^{13} \sqrt{n \log n}/B = 2^{-39}/t$; and $44(2^{-47} + 2^{-419} + 2^{-39}) < 2^{-32}$. Moreover

$$\sqrt{n \log n} \leq 2^{\frac{9}{2} + \frac{210}{a}} t^{\frac{1}{2a}} \mathfrak{f}^{\frac{1}{2} + \frac{1}{2a}} < 2^{117} t^{\frac{5000}{18693}} \mathfrak{f}^{\frac{28693}{37386}},$$

and, as $t^{1079/205623} \geq 1 + \frac{1079}{205623} \log t$, $\mathfrak{f}^{28693/37386} \leq \mathfrak{f} \leq \frac{205623}{1079} t^{1079/205623} < 2^8 t^{1079/205623}$. Since $1 + \frac{5000}{18693} + \frac{1079}{205623} = \frac{14}{11}$, $B < 2^{52+117+8} t^{14/11}$, and (6.12) gives (6.13). Since $\beta_{j,X}^{\text{term}}(t) \leq \beta_X(t)$, the same bound holds with $\beta_X(t)$.

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 7.1), chooses a seed that satisfies them (Proposition 7.3), and deduces Theorem 1.3, using the step bounds of Section 8.

7.1. The count under the hypotheses.

Proposition 7.1 (Count under the hypotheses). *Let $t \geq 16^{16544}$ be odd and not S -periodic, with $Z_{j_0}(t) > \frac{6}{7} + 2^{-28}$. Then every real $X \geq 2^{2^{352}} t$ is reachable at some generation $j \geq J_c$, and at every such j*

$$\beta_{j,X}^{\text{term}}(t) > 2^{-207} t^{-14/11}, \quad \# \text{Src}_{j,X}(t) > 2^{-207} t^{-25/11} X. \quad (7.1)$$

Proof. Proposition 6.9 gives j . As $j \geq J_c > j_0$, Lemma 6.3 gives $Z_j(t) > 2^{-28}$, and (6.10) gives $Z_{j,X}^{\text{term}}(t) > 2^{-28} - 2^{-29} = 2^{-29}$. By (6.13),

$$\beta_{j,X}^{\text{term}}(t) > (2^{-29} - 2^{-32}) 2^{-177} t^{-14/11} = 7 \cdot 2^{-209} t^{-14/11} > 2^{-207} t^{-14/11},$$

and the second bound follows by (5.10). $\square$

These hypotheses follow from (H1)–(H4): (H3) is $t \geq 2^{66176} = 16^{16544}$, not being S -periodic is (H2), and (H4) gives $Z_{j_0}(t) = Z_{j_0}(y_0) > \frac{13}{15} > \frac{6}{7} + 2^{-28}$ by Lemma 6.4, as $\frac{13}{15} - \frac{6}{7} = \frac{1}{105}$.

7.2. Choosing a seed. A general target need not satisfy the hypotheses of Proposition 7.1, 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 $2^{66178} = 4^{33089} \equiv 1 \pmod{3}$, exactly one $a^\circ \in \{66178, 66179\}$ 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 7.2 (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 [7, 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 . $\square$

With D_A as in Section 1, the integer comparison $3^{200} < 2^{317}$ and (6.7) give

$$\log_2 D_A = 2^{163} > 2^{4 + \frac{317}{200} d_0} > 16 \cdot 3^{d_0}. \quad (7.2)$$

Proposition 7.3 (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 $a < 66180 + 4 \cdot 3^{d_0}$ that is odd, not S -periodic, satisfies*

$$2^{66176} t < t_1 < 2^{5 \cdot 3^{d_0}} t < D_A t \quad \text{and} \quad Z_{j_0}(t_1) > \frac{13}{15},$$

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

Proof. By Proposition 6.7 there is $y_0 \in G_{d_0}$ with $Z_{j_0}(y_0) > \frac{13}{15}$. Apply Lemma 7.2 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 $a = a^\circ + 2i \leq 66177 + 4 \cdot 3^{d_0}$. 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 $(2^{66178} t - 1)/3 > 2^{66176} t$ and $2^a t/3 < 2^{5 \cdot 3^{d_0}} t < D_A t$, as $66177 < 3^{11} \leq 3^{d_0}$ and by (7.2). 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 7.1 for the seed counts integers reaching t . Let t be a positive integer with $3 \nmid t$, fix t_1 as in Proposition 7.3.

Proposition 7.4. *For every real $X \geq D_A t$, at least $X/(t^{25/11} D_A)$ 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. As $d_0 \geq 352$, $2^{352} \leq 2^{d_0} < 3^{d_0}$, so (7.2) gives $5 + 2^{352} + 5 \cdot 3^{d_0} < 16 \cdot 3^{d_0} < \log_2 D_A$ and $212 + \frac{125}{11} 3^{d_0} < \log_2 D_A$. Hence $X/32 \geq 2^{2^{352} + 5 \cdot 3^{d_0}} t > 2^{2^{352}} t_1$. The seed t_1 is odd, not S -periodic, at least $2^{66176} = 16^{16544}$, and has $Z_{j_0}(t_1) > \frac{13}{15} > \frac{6}{7} + 2^{-28}$, so Proposition 7.1 applies to it at $X/32$: at some $j \geq J_c$ with $X/32$ reachable,

$$\#\text{Src}_{j, X/32}(t_1) > 2^{-212} t_1^{-25/11} X > 2^{-212 - \frac{125}{11} 3^{d_0}} t^{-25/11} X \geq \frac{X}{t^{25/11} D_A}.$$

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. $\square$

Proof of Theorem 1.3. Let $c(t)$ be as in Section 1 and $X > 0$. If $X \geq D_A t$, Proposition 7.4 gives at least $c(t)X \geq \lfloor c(t)X \rfloor$ odd $n < X$ with both step bounds. If $X < D_A t$, then $c(t)X < t^{-14/11} \leq 1$, so $\lfloor c(t)X \rfloor = 0$. $\square$

8. THE NUMBER OF STEPS

This section proves the step bounds used in Proposition 7.4 (Proposition 8.4), following [7, 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, the length restriction 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; only t_1 odd, $3t_1 + 1 = 2^a t$ and $t_1 > 2^{66176} t$ are used. For a generation j define, suppressing j ,

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

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 $j \geq J_c$. Then (1) $\delta_j \geq 2^{40} j$, and hence $2^{22}(j+1) \leq \delta_\Sigma$; (2) $\delta_j \leq 16544 + \delta_\Sigma/100 + j$; and (3) $9000 \ell_\Sigma^c \leq \delta_\Sigma$.*

Proof. For (1), let $k = \lfloor j/70 \rfloor \geq 328$; by Lemma 15.1(1) and $g^{70} > 2$, $\delta_j \geq 16544 \cdot 2^k \geq 2^{14+k}$, and $2^{k-26} \geq 70(k+1) > j$ by induction on $k \geq 328$; so $\delta_j \geq 2^{40} j$ and $2^{22}(j+1) \leq 2^{23} j \leq \delta_j \leq \delta_\Sigma$. For (2), $e_\delta < \delta/100 + 1$, so

$$\delta_j = 16544 + \sum_{i < j} e_{\delta_i} \leq 16544 + \frac{\delta_\Sigma}{100} + j.$$

For (3), as in [7, (7.1)], $\ell_\delta^c \leq \ell_\delta \leq \delta$, and $20000 \ell_\delta^c \leq \delta$ for $\delta \geq 2^{40}$ by (5.3). The scales below 2^{40} sum to less than $101 \cdot 2^{40}$ (proof of Lemma 6.4), and the other terms contribute at most $\delta_\Sigma/20000$. By (1), $\delta_\Sigma \geq \delta_j \geq 2^{342}$, hence $9000 \ell_\Sigma^c \leq 9000 \cdot 101 \cdot 2^{40} + \frac{9}{20} \delta_\Sigma \leq \delta_\Sigma$, as $9000 \cdot 101 \cdot 2^{40} < 2^{60}$. $\square$

Let $j \geq J_c$, 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 5.3, as $t_1 > 16^{16544}$).

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 [7, Section 7]. The first value is $t_1 > 2^{66176}$. 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 $y \geq 16^{\delta_i} 2^{-8\delta_i/5} = 2^{12\delta_i/5} > 4096$.

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 [7, (7.4)].

For (3), which is [7, (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. 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 (5.9) 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 [7, (7.2)] with this paper's scales. Apply (5.4) 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), $r_{\delta_j} \leq \frac{3}{10}\delta_j \leq \frac{24816}{5} + \frac{3}{1000}\delta_\Sigma + \frac{3}{10}j$, and $2j + 2 + \frac{3}{10}j \leq \frac{26}{10}(j + 1)$. Since $\log(t_1/t) > 66176 \log 2 > \frac{24816}{5} \log 2$ 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 $j \geq J_c$, Lemmas 8.1 and 8.3 apply. For the accelerated steps we follow the proof of [7, 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), $j + \delta_j/8 \leq 2068 + \frac{\delta_\Sigma}{800} + \frac{9}{8}j$, and by $\log 2 < 347/500 < 7/10$ and Lemma 8.1(1), $\frac{9}{8}j \log 2 < \delta_\Sigma/10^6$ and $\frac{\delta_\Sigma}{800} \log 2 < \frac{7\delta_\Sigma}{8000}$. Finally $c_{\text{gain}} + 2068 \log 2 < 2069 \log 2 < \log(t_1/t)$. As $1/10^6 + 7/8000 = 219/250000$, we obtain, as in [7, (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 [7, (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

$$\left(c_{\text{src}} c_{\text{acc}} - \frac{1}{9000} \right) \frac{287682}{10^6} - \frac{c_{\text{src}} c_{\text{acc}}}{12288} - c_{\text{src}} - \frac{219}{250000} = \frac{25365007801}{19200000000000000} > 0. \quad (8.3)$$

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 [7, 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 the inequality [7, (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) for both indices $\iota \in \{\text{o}, \text{b}\}$ at once, following Tao's argument [9, Section 7] with explicit constants; the index enters only through the parameters of Table 1. 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, every point bordering a triangle is *white*, and the angles of the points just above a triangle are close to odd dyadic fractions. 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, scales and families 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 $\varepsilon = 37037/10^6$ is the constant of Definition A.2. The objects defined below depend on n and ξ ; this dependence is suppressed from the notation.

The arguments of this section use only $0 < \varepsilon < 1/27$, which holds as $27\varepsilon = 999999/10^6$.

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 [9, (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 [9, (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 [9, (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 [9, Lemma 7.4] assumes $0 < \varepsilon < 1/100$; our ε is larger, which is why we reprove its ingredients with explicit thresholds. The next lemma is a quantitative form of the claims (ii) and (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 (ii) and (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 $2\vartheta(p) = \vartheta(j, l-1)$:*

- (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 $\bar{\vartheta}_2 < 1$.

(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 [9, 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 [9, 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 for $\varepsilon < 1/27$ 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 [9, 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 [9, 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. Exit points above a triangle. In Section 12 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 $r, o \geq 1$. When r is not too large, the exit point lies before the horizon $j = J$ and its angle is an odd dyadic fraction up to a small error, which the packet envelopes of Lemma 11.5 turn into cancellation.

Lemma 9.14 (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, and let $\hat{p} = (j(p), l_\Delta)$.*

- (a) $j(p) + r \leq J$ for every integer $0 \leq r \leq f_\omega$.
- (b) Let $1 \leq r \leq f_\omega$ and $o \geq 1$ be integers and $e = (j(p) + r, l_\Delta + o)$. Then

$$\vartheta(e) = \frac{m + \beta}{2^o}$$

for an odd integer m and a real β with $|\beta| \leq 9^r \varepsilon 2^{-\omega} \leq 9\varepsilon < \frac{1}{3}$.

- (c) $3^n \vartheta(\hat{p})$ is an odd integer and $|\vartheta(\hat{p})| \leq \varepsilon 2^{-\omega}$.

Proof. Since $f_\omega \leq 1 + 5\omega/16$ and $9^5 \leq 2^{16}$, $9^r \leq 9^{f_\omega} \leq 9 \cdot 2^\omega$ for $0 \leq r \leq f_\omega$. Hence $3\varepsilon 9^r \leq 27\varepsilon 2^\omega < 2^\omega$.

(a) 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, $2(j(p) + r) < n + 1$, so $j(p) + r \leq J$.

(c) The point $\hat{p}$ lies in Δ by northwest closure; let $N = 3^n \vartheta(\hat{p})$. By Lemma 9.4, Lemma 9.10(a) and $S_\Delta \geq (j(p) - j_\Delta) \log 9 + \omega \log 2$,

$$|\vartheta(\hat{p})| \leq 9^{j(p)-j_\Delta} \varepsilon e^{-S_\Delta} \leq \varepsilon 2^{-\omega}.$$

The integer N is odd. Otherwise $N/2$ represents $\text{res}(\hat{p} + (0, 1))$, since $\text{res}(\hat{p}) = 2 \text{res}(\hat{p} + (0, 1))$ by Lemma 9.4 and 2 is invertible modulo 3^n ; as $|N/2| < 3^n/2$, (9.1) gives $|\vartheta(\hat{p} + (0, 1))| = |\vartheta(\hat{p})|/2 \leq \varepsilon/2$. But $\hat{p} + (0, 1) \notin \Delta$ is a cardinal neighbour of $\hat{p} \in \Delta$, hence white by Lemma 9.11, a contradiction.

(b) By Lemma 9.4, the point $(j(p) + r, l_\Delta)$ has residue $9^r \text{res}(\hat{p}) = 2^o \text{res}(e)$, so $2^o 3^n \vartheta(e) = 9^r N + m 3^n$ for an integer m . Here $2^o 3^n \vartheta(e)$ is even and $9^r N$ and 3^n are odd, so m is odd. Dividing by $2^o 3^n$ gives $\vartheta(e) = (m + \beta)/2^o$ with $\beta = 9^r \vartheta(\hat{p})$, and $|\beta| \leq 9^r \varepsilon 2^{-\omega} \leq 9\varepsilon < 1/3$ by (c), the first paragraph and $27\varepsilon < 1$. $\square$

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 [9, §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 [9, Lemmas 2.2, 7.6, 7.7], we prove exact laws, exponential moments and explicit tail bounds for it, used in Sections 11–14.

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 real $X, y > 0$ put

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

with values in $[0, \infty]$, and $V(y) = G(1, y) = \mathbb{E} y^{l(h_1)}$.

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) *Let $X > 0$ and $0 < y < 2$, and put $V_{nc}(y) = \frac{y^2}{(2-y)^2} - \frac{3}{16}y^4 - \frac{1}{8}y^5$. If $XV_{nc}(y) < 1$, then*

$$G(X, y) = \frac{X(\frac{3}{16}y^4 + \frac{1}{8}y^5)}{1 - XV_{nc}(y)} < \infty.$$

(5) $V(27/25) < 8$.

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), $X^j y^l = \prod_{i=1}^j X y^{c_i}$ for $b = (c_1, \dots, c_j) \in \mathfrak{B}_{j,l}$, so the factorization in the proof of Lemma 10.4 gives $G(X, y) = \sum_{j \geq 1} (XV_{nc}(y))^{j-1} X(\frac{3}{16}y^4 + \frac{1}{8}y^5)$. 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, by (4) with $X = 1$, $V(27/25) = 36266065281/5016065281 < 8$. $\square$

Lemma 10.9 (Sums of hold points). *Let $N \in \mathbb{N}$.*

- (1) For real $\mathfrak{t}$, $\mathbb{E}e^{\mathfrak{t}(j(\mathbf{h}_1)+\dots+j(\mathbf{h}_N))} = \mathbf{M}_{45}(\mathfrak{t})^N$ in $[0, \infty]$.
(2) For real $X, y > 0$, an integer $0 \leq u \leq N$ and an event E determined by $\mathbf{h}_1, \dots, \mathbf{h}_u$,

$$\mathbb{E}[\mathbf{1}_E X^{j(\mathbf{W}_N)} y^{l(\mathbf{W}_N)}] = \mathbf{G}(X, y)^{N-u} \mathbb{E}[\mathbf{1}_E X^{j(\mathbf{W}_u)} y^{l(\mathbf{W}_u)}]$$

in $[0, \infty]$, with $\mathbf{W}_0 = (0, 0)$ and $0 \cdot \infty = 0$.

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

- (4) 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. Parts (1) and (2) are independence (Lemma 10.4(1)): $\mathbf{W}_N = \mathbf{W}_u + \mathbf{h}_{u+1} + \dots + \mathbf{h}_N$, with $\mathbf{h}_{u+1}, \dots, \mathbf{h}_N$ independent of $\mathbf{h}_1, \dots, \mathbf{h}_u$. Part (3) is the exponential Markov inequality: the probability is at most $e^{-\mathfrak{m}/160} \mathbf{M}_{45}(1/16)^N$; apply Lemma 10.8(3) and $5/4 < 1 + \frac{1}{2} + \frac{1}{8} \leq e^{1/2}$. In (4) 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 (1) and Lemma 10.8(1) apply. $\square$

10.4. First-passage overshoot. The first passage over a height does not overshoot it by much.

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

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

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

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

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. The horizontal first-passage law is a mixture of two binomial rows (Lemma 10.13), whose tails give the horizontal tail of the first passage (Lemma 10.14). We first bound binomial rows.

Lemma 10.11 (Binomial bounds). *For $N \in \mathbb{N}$ and $k \in \mathbb{Z}$, $\binom{N}{k}2^{-N} \leq 2/(3N+2)$; in particular $R(k, l) \leq (2(3l-1))^{-1/2}$ for all $k \in \mathbb{N}$ and integers $l \geq 1$. 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 $1/(3m+4) = 2/(3N+5)$. By Lemma 10.2(2), $R(k, l)$ is 0 or $\frac{1}{2}\binom{l-1}{2k-1}2^{-(l-1)} \leq \frac{1}{2}(2/(3l-1))^{1/2}$.

The last 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 (Horizontal tail of the first passage). *For $\omega \in \mathbb{N}$ and real $u \geq 2$,*

$$\sum_{x: |j(x)-\omega/4| \geq u} F_\omega(x) \leq \begin{cases} 4e^{-u/256} & \text{if } \omega \leq 5, \\ 4 \exp\left(-\frac{u^2}{2(\omega-4)}\right) + 2e^{-u/256} & \text{if } \omega \geq 6. \end{cases}$$

Proof. Case 1. Suppose $\omega \leq 5$. A nonzero term has $j(x) \geq 1$, so $j(x) - \omega/4 \geq -\frac{1}{4} > -u$; hence $j(x) - \omega/4 \geq u$ and $j(x) \geq u$. Markov's inequality, Lemma 10.9(4) with exponent $1/16$ and Lemma 10.8(3) bound the left side by $(5/4)^{\omega+1}e^{-u/16} \leq (5/4)^6 e^{-u/16} \leq 4e^{-u/256}$, as $(5/4)^6 = 15625/4096$.

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 $p_l(i) = \binom{l}{2i} + \binom{l}{2i+1}2^{-l} \geq p_l^{\text{cl}}(i)$. For $i \in \mathcal{J}$ and $i' \in \{2i, 2i+1\}$ we have $|i' - l/2| \geq 2|i - l/4| - 1 \geq u - 1 \geq u/2$, 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

$$\sum_{i \in \mathcal{J}} \mathbf{p}_l(i) + \sum_{k \geq u/2} \nu_{45}(k) \leq 2e^{-u^2/(2l)} + \left(\frac{11}{16}\right)^{u/2-1} \leq 2e^{-u^2/(2l)} + 2e^{-u/256},$$

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

11. THE CHARACTER SUM AND THE WEIGHTED MAXIMA

This section bounds $|\hat{\mu}_n(\xi)|$ by the expectation $\sum_y \eta(y) \mathbf{Q}(y)$ of the expected penalty $\mathbf{Q}$ (Proposition 11.13, used in Theorem 14.6), following [9, §§7.1, 7.3], but with cancellation at every raw letter up to 16 (Lemma 11.2 and the paid rewards of Definition 11.6). It then introduces the weighted maxima $\mathbf{Q}_m$, whose descent step $\mathbf{Q}_m \leq \mathbf{Q}_{m-1}$ is proved in Sections 12–14.

The constants $z = \frac{133}{3125}$ and $\kappa = \frac{3187}{20000}$ are those of Definition A.2; $\kappa z = \frac{423871}{62500000}$. Every series below is indexed by a countable set and has nonnegative terms or converges absolutely, so it may be rearranged and regrouped freely. For $p \in \mathcal{P}$ and $k \in \mathbb{Z}$ write $p - (0, k) = (j(p), l(p) - k)$.

11.1. Pairing. Grouping the letters in pairs, as in [9, §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 $(i, \mathbf{c}_1 + \cdots + \mathbf{c}_i)$. For an integer $t \geq 1$ and real θ , and for $p \in \mathcal{P}$ and an integer c , put

$$\Phi_t(\theta) = \frac{1}{t} \left| \sum_{a=0}^{t-1} e^{2\pi i(2^a-1)\theta} \right|, \quad \varphi_c(p) = \Phi_{c-1}(\vartheta(p)) \quad (c \geq 2), \quad (11.1)$$

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

Lemma 11.1 (Pairing).

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

Proof. Let $\mathbf{e}(y) = e^{-2\pi i \xi y / 3^n}$ for $y \in G_n$, evaluated at elements of $\mathbb{Z}[1/2]$ after reduction (Section 3). Then $\hat{\mu}_n(\xi) = \mathbb{E}(\text{off}(W_{\leq n}))$ is the sum of [9, (7.2)] for the character $\mathbf{e}$, since the expression in [9, (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 $p_i = (i, \mathbf{c}_1 + \cdots + \mathbf{c}_i)$ and $x_i = 3^{2(i-1)} 2^{-l(p_i)}$, [9, (7.4)–(7.5)] read

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

where the terms with some $\mathbf{c}_i \leq 1$ vanish. Since $2^k+3 = 5 + 2(2^{k-1}-1)$ and $\mathbf{e}(2x_i) = e^{-2\pi i \vartheta(p_i)}$ (Definition 9.2), we have $\mathbf{e}((2^k+3)x_i) = \mathbf{e}(5x_i) e^{-2\pi i(2^{k-1}-1)\vartheta(p_i)}$. The sum over k of the last factors is the complex conjugate of the sum in (11.1) with $t = \mathbf{c}_i - 1$ and $a = k-1$, 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*

$$\varphi_3(p) < e^{-\kappa z}, \quad \varphi_4(p) < e^{-z}, \quad \varphi_5(p) < e^{-z}.$$

Proof. Put $x = \pi|\vartheta(p)| \in (\pi\varepsilon, \frac{\pi}{2}]$, $\mathbf{a} = e^{2ix}$ and $x_o = 3.1415\varepsilon = \frac{232703471}{2 \cdot 10^9}$. Replacing $\vartheta(p)$ by $-\vartheta(p)$ conjugates the sum in (11.1), so $\varphi_c(p) = \frac{1}{c-1} |\sum_{a=0}^{c-2} \mathbf{a}^{2^a-1}|$. 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 $2^a - 1$ in (11.1) are 0, 1, 3, 7 for $a = 0, 1, 2, 3$.

The raw letter 3. As in [9, Lemma 7.2], $\varphi_3(p) = \frac{1}{2}|1 + \mathbf{a}| = \cos x$. As $\cos$ decreases on $[0, \pi]$, $\cos x < \cos x_o \leq 1 - x_o^2/2 + x_o^4/24$. 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 0.9932387... and 0.9932410...) 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 $\mathbf{u}_o < \frac{1}{4}$, $\frac{18362}{100000} < \mathfrak{h}(\mathbf{u}_o) < \frac{7}{16}$ and

$$9 - 4 \cdot \frac{18362}{100000} < 9(1 - 2z + 2z^2 - \frac{4}{3}z^3) \leq 9e^{-2z}$$

(decimal values 8.26552 and 8.26559...; the last step is the Taylor bound at $y = 2z$). In both cases $\mathfrak{h}(\mathbf{u}) > \frac{18362}{100000}$, 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 0.947. 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 $1 - z = \frac{2992}{3125} \leq e^{-z}$. $\square$

11.3. Packets. A raw letter $c \geq 5$ need not sit at a white point to cancel: a block of consecutive terms of the sum defining $\varphi_c(p)$ is the sum defining φ_4 or φ_5 at a point below p , and cancels if that point is white or if its angle is near an odd dyadic fraction.

Lemma 11.3 (Disjoint packets). *Let $c \geq 2$ be an integer, θ real, and let S be a finite set of pairs $(k, t) \in \mathbb{N} \times \mathbb{Z}_{\geq 1}$ with $k + t \leq c - 1$ whose packets $\{k, k + 1, \dots, k + t - 1\}$ are pairwise disjoint. Then*

$$(c - 1)\Phi_{c-1}(\theta) \leq (c - 1) - \sum_{(k,t) \in S} t + \sum_{(k,t) \in S} t\Phi_t(2^k\theta).$$

In particular $(c - 1)\varphi_c(p) \leq (c - 1) - \sum_{(k,t) \in S} t + \sum_{(k,t) \in S} t\varphi_{t+1}(p - (0, k))$ for $p \in \mathcal{P}$.

Proof. Split the sum $(c - 1)\Phi_{c-1}(\theta) = |\sum_{a=0}^{c-2} e^{2\pi i(2^a-1)\theta}|$ into the packets of S and the remaining $c - 1 - \sum_S t$ terms, each of modulus 1, and apply the triangle inequality. For $(k, t) \in S$, $2^{k+a} - 1 = (2^k - 1) + 2^k(2^a - 1)$, so the packet has modulus $t\Phi_t(2^k\theta)$. For the second claim take $\theta = \vartheta(p)$: $\vartheta(p - (0, k)) - 2^k\vartheta(p) \in \mathbb{Z}$, since $\text{res}(p - (0, k)) = 2^k \text{res}(p)$ (Lemma 9.4), and Φ_t has period 1. $\square$

The angles of exit points are odd dyadic fractions up to an error at most $1/3$ (Lemma 9.14(b)). At such angles Φ_t is bounded away from 1.

Definition 11.4 (Packet envelope table). Let T_U be the set of the 21 pairs (s, t) for which the following table has an entry $U_{s,t}$; each entry is an exact terminating decimal.

$s \backslash t$	2	3	4	5	6	7
1	0.5	0.590	0.504			
2	0.867	0.579	0.626			
3	0.96593	0.800	0.535			
4	0.992	0.94741	0.775			
5	0.99786	0.98672	0.93963			
6	0.999465	0.996672	0.984682	0.939		
7					0.940	
8						0.943

Lemma 11.5 (Packet envelopes). *For every $(s, t) \in T_U$, every odd integer m and every real β with $|\beta| \leq \frac{1}{3}$, $\Phi_t((m + \beta)/2^s) \leq U_{s,t}$.*

Proof. Expanding the square,

$$t^2 \Phi_t(\theta)^2 = g_t(\theta) := t + 2 \sum_{0 \leq i < j \leq t-1} \cos(2\pi(2^j - 2^i)\theta).$$

For $(s, t) = (1, 2)$, $\Phi_2(\theta) = |\cos(\pi\theta)|$ and $\theta = (m + \beta)/2$ lies within $\frac{1}{6}$ of an odd multiple of $\frac{1}{2}$, so $\Phi_2(\theta) \leq \cos \frac{\pi}{3} = \frac{1}{2}$. For the other 20 pairs we show $g_t < t^2 U_{s,t}^2$ on the set of θ in question by a finite computation in exact integer arithmetic. The function g_t is even, 1-periodic, and Lipschitz with constant $L_t = \frac{88}{7} \sum_{i < j < t} (2^j - 2^i)$, as $\pi < \frac{22}{7}$. Reducing m modulo 2^s and replacing θ by $-\theta$ if necessary, θ lies in a cell $[(m' - \frac{1}{3})/2^s, (m' + \frac{1}{3})/2^s]$ with m' odd and $1 \leq m' \leq 2^{s-1}$. All points used below are multiples of $1/(3 \cdot 2^{s+20})$. At such a point the turn $\phi = (2^j - 2^i)\theta$ of each cosine is reduced to $\alpha = |\phi - \lfloor \phi + \frac{1}{2} \rfloor| \in [0, \frac{1}{2}]$; if $\alpha \leq \frac{1}{4}$, $\cos(2\pi\phi)$ is at most the Taylor polynomial of degree 14 of $\cos$ at $2\pi_0\alpha$, $\pi_0 = 3.141592654$, plus 10^{-9} , and if $\alpha > \frac{1}{4}$, $\cos(2\pi\phi) = -\cos(2\pi(\frac{1}{2} - \alpha))$ is bounded in the same way; here 10^{-9} exceeds $2(8/5)^{16}/16! + \frac{1}{2}|\pi - \pi_0|$, a bound for the Lagrange remainder of Taylor's formula plus the error from π_0 , since $2\pi_0\alpha \leq 8/5$ and $|\pi - \pi_0| < 5 \cdot 10^{-10}$. These upper bounds, rounded up to multiples of 2^{-64} , give an upper bound for g_t at the point. A segment $[\theta_0, \theta_1]$ is accepted if the larger of the bounds at its ends plus $L_t(\theta_1 - \theta_0)/2$ is less than $t^2 U_{s,t}^2$, which bounds g_t on the segment; otherwise it is split at $\lfloor (N_0 + N_1)/2 \rfloor / Q$, where $\theta_i = N_i/Q$ and $Q = 3 \cdot 2^{s+20}$, and each part is treated in the same way. Starting from the cells and splitting at most 20 times, every segment is accepted, for each of the 20 pairs; at most 831 segments are tested per pair. $\square$

11.4. Paid rewards and the penalty. The descent step needs the cancellation of every raw letter, not only of the raw letters 3, 4, 5 at white points. The paid reward $R_c(p)$ below records it: $e^{-zR_c(p)}$ bounds $\varphi_c(p)$, with the bounds of Lemmas 11.2 and 11.3 as a floor and the actual value of $\varphi_c(p)$, capped at 2^{-16} , for $3 \leq c \leq 16$.

Definition 11.6 (Paid rewards). Let c be an integer and $p \in \mathcal{P}$. A pair $(k, t) \in \mathbb{N} \times \{3, 4\}$ with $k + t \leq c - 1$ is *admissible* if $p - (0, k)$ is white, and the *coverage* $M_c(p)$ is the largest value of $\sum_{(k,t) \in S} t$ over the sets S of admissible pairs with pairwise disjoint packets (with $M_c(p) = 0$ for $S = \emptyset$). The *baseline reward* is

$$R_c^{\text{old}}(p) = \begin{cases} \kappa \mathbf{1}[p \text{ white}], & c = 3, \\ \mathbf{1}[p \text{ white}], & c = 4, \\ \mathbf{1}[p \text{ white}] + \frac{18351}{25000} \mathbf{1}[p \text{ not white, } p - (0, 1) \text{ white}], & c = 5, \\ \frac{M_c(p)}{c-1} (1 - \frac{z}{2}), & 6 \leq c \leq 16, \\ 0 & \text{otherwise.} \end{cases}$$

The *paid reward* is $R_c(p) = \max\{R_c^{\text{old}}(p), -z^{-1} \log \max(\varphi_c(p), 2^{-16})\}$ if $3 \leq c \leq 16$ and $j(p) \leq J$, and $R_c(p) = R_c^{\text{old}}(p)$ otherwise. The *residual reward* is $R_c^{\text{res}}(p) = R_c(p) - \mathbf{1}[c \in \{4, 5\}] \mathbf{1}[p \text{ white}]$.

Here $\frac{18351}{25000} = \frac{3}{4}(1 - \frac{z}{2})$.

Lemma 11.7 (Paid rewards). *Let c be an integer and $p \in \mathcal{P}$.*

- (1) $R_c(p) \geq R_c^{\text{old}}(p) \geq 0$ and $R_c^{\text{res}}(p) \geq 0$; if $j(p) > J$, then $R_c(p) = 0$.
- (2) $\varphi_c(p) \leq e^{-zR_c^{\text{old}}(p)}$ and $\varphi_c(p) \leq e^{-zR_c(p)}$.
- (3) If $3 \leq c \leq 16$ and $j(p) \leq J$, then $e^{-zR_c(p)} \leq \max(\varphi_c(p), 2^{-16})$.

Proof. (1) The first inequality holds by definition, and each case of R_c^{old} is nonnegative. If $c \in \{4, 5\}$ and p is white, then $R_c^{\text{old}}(p) = 1$, so $R_c^{\text{res}}(p) \geq 0$; otherwise $R_c^{\text{res}}(p) = R_c(p)$. If $j(p) > J$, no point $p - (0, k)$ is white, so $M_c(p) = 0$ and $R_c(p) = R_c^{\text{old}}(p) = 0$.

(2) The second bound follows from the first, since also $\varphi_c(p) \leq \max(\varphi_c(p), 2^{-16})$ and $y \mapsto e^{-zy}$ decreases. For the first we may assume $R_c^{\text{old}}(p) > 0$, as $\varphi_c \leq 1$. If $c \in \{3, 4, 5\}$ and p is white, apply Lemma 11.2. If $c = 5$, p is not white and $p - (0, 1)$ is white, Lemma 11.3 with $S = \{(1, 3)\}$ and Lemma 11.2 give $\varphi_5(p) \leq \frac{1}{4}(1 + 3e^{-z}) \leq 1 - \frac{3}{4}z(1 - \frac{z}{2}) \leq e^{-zR_5^{\text{old}}(p)}$, using $e^{-z} \leq 1 - z + z^2/2$ and $1 - y \leq e^{-y}$. If $6 \leq c \leq 16$, take S attaining $M = M_c(p)$; as $t + 1 \in \{4, 5\}$ for $(k, t) \in S$, Lemmas 11.3 and 11.2 give

$$\varphi_c(p) \leq 1 - \frac{M}{c-1}(1 - e^{-z}) \leq 1 - z \frac{M}{c-1} \left(1 - \frac{z}{2}\right) \leq e^{-zR_c^{\text{old}}(p)}.$$

- (3) By definition $R_c(p) \geq -z^{-1} \log \max(\varphi_c(p), 2^{-16})$. □

Definition 11.8 (White factor and penalty). For $p \in \mathcal{P}$ the *white factor* is $w(p) = e^{-z}$ if p is white, and $w(p) = 1$ otherwise. For a word of raw letters $\mathbf{c} = (\mathbf{c}_1, \dots, \mathbf{c}_j) \in \mathbb{Z}^j$, the *penalty* is

$$\text{Pen}(\mathbf{c}) = \prod_{i=1}^j e^{-zR_{\mathbf{c}_i}(i, \mathbf{c}_1 + \dots + \mathbf{c}_i)} \in (0, 1].$$

11.5. Splitting the penalty into holds. Splitting the raw letters into holds turns the penalty into a product along the hold walk: the white factors w at the hold points carry the reward $\mathbf{1}[\text{white}]$ of the closing letters, and the internal factor carries the residual rewards of all letters of a hold (Lemma 11.10).

Definition 11.9 (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* is

$$l(x; b) = \prod_{i=1}^{|b|} e^{-zR_{\mathbf{c}_i}^{\text{res}}(x + (i, \mathbf{c}_1 + \dots + \mathbf{c}_i))} \in (0, 1].$$

Lemma 11.10 (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. For a hold $b = (\mathbf{c}_1, \dots, \mathbf{c}_r)$ and $x \in \mathbb{N} \times \mathbb{Z}$, the letters $\mathbf{c}_i$, $i < r$, are not closing, so $R_{\mathbf{c}_i}^{\text{res}} = R_{\mathbf{c}_i}$, while the closing letter sits at $y = x + \text{pt}(b)$, where $w(y)e^{-zR_{\mathbf{c}_r}^{\text{res}}(y)} = e^{-zR_{\mathbf{c}_r}(y)}$. Hence

$$w(x + \text{pt}(b)) l(x; b) = \prod_{i=1}^r e^{-zR_{\mathbf{c}_i}(x + (i, \mathbf{c}_1 + \dots + \mathbf{c}_i))}.$$

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, with

$x = \text{pt}_{k-1}(\mathbf{b})$, the right side of the claim is $\prod_i e^{-zR_{\mathbf{c}'_i}(i, \mathbf{c}'_1 + \dots + \mathbf{c}'_i)}$ over the entire raw word $\mathbf{c}'$, whose length is at least $K \geq J$. The raw letters beyond position J have points of raw length greater than J , where the rewards vanish (Lemma 11.7(1)), so their factors are 1. $\square$

11.6. Reduction to the expected penalty. The character sum is bounded by the average penalty of the hold walk (Proposition 11.13), the analogue of [9, (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 11.11 (The expected penalty). For $K \in \mathbb{N}$ and $p \in \mathcal{P}$ let

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

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

Lemma 11.12 (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) $\mathbf{Q}(p) = \mathbf{w}(p) \sum_{b \in \mathfrak{B}} \mathbf{q}(b) \mathbf{l}(p; b) \mathbf{Q}(p + \text{pt}(b))$.
- (3) $\mathbf{Q}(p) \leq \mathbf{w}(p) \sum_{y \in \mathcal{P}} \eta(y) \mathbf{Q}(p + y)$.

Proof. Part (2) is the analogue of [9, (7.35)], with internal factors and truncated sums. All factors in Definition 11.11 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

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

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 $\mathbf{w}(p + \text{pt}_{K+1}(\mathbf{b}))$ and $\mathbf{l}(p + \text{pt}_K(\mathbf{b}); b^{K+1})$ are evaluated at points of raw length at least $j(p) + K + 1 > J$, which are not white and where the rewards vanish, so they equal 1, and the last hold then sums to mass 1. By induction, $\mathbf{U}_K(p)$ does not depend on K once $j(p) + K \geq J$; as $K = J$ qualifies, the common value is $\mathbf{Q}(p)$. If $j(p) > J$, the same reasoning makes every factor equal to 1, and $\mathbf{Q}(p) = 1$. This proves (1). For (2), apply (11.3) with $K = J$, noting $\mathbf{U}_{J+1}(p) = \mathbf{Q}(p)$ by (1). For (3), drop $\mathbf{l} \leq 1$ in (2) and group the holds by their point using (10.1). $\square$

Proposition 11.13 (Character sum and expected penalty).

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

Proof. If $J = 0$, then $\mathbf{Q}(y) = 1$ for every $y \in \mathcal{P}$ (Lemma 11.12(1)), so the right side is 1, while $|\hat{\mu}_n(\xi)| \leq \sum_x \mu_n(x) = 1$.

Let $J \geq 1$. By Lemma 11.7(2), $\varphi_c(p) \leq e^{-zR_c(p)}$ for every integer c and every $p \in \mathcal{P}$, so (11.2) gives

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

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 11.10, 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 $q(b^1) l((0, 0); b^1) \cup_{J-1}(y)$, by Definition 11.11. Since $j(y) \geq 1$, Lemma 11.12(1) gives $\cup_{J-1}(y) = Q(y)$. Finally drop $l \leq 1$ and group the holds b^1 by their point, using (10.1). $\square$

11.7. Weighted maxima near the horizon. To show that $Q(p)$ decays like a power of the distance to the horizon $j = J$, we follow [9, (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 11.14 (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 $j(p) + m = J$. For a real exponent $A \geq 0$, suppressed from the notation, and 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 11.15 (Basic properties of the weighted maxima). *Let $A \geq 0$ and 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 $d_J(p)^A Q(p) \leq Q_m$.*

Proof. As in [9, (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 11.12(1)), and the set in Definition 11.14 contains the point $(J + 1, 0)$. $\square$

Lemma 11.16 (Boundary reduction). *Let $A \geq 0$ and 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 [9, §7.4] (the bound opening Case 1, and the reduction of (7.40) to (7.41)), by Lemma 11.15(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 $d_J(p) = m$. $\square$

For the index ι take $A = A_\iota$. By Lemma 11.16(2), the descent step $Q_m \leq Q_{m-1}$ (Proposition 14.4) reduces to the bound $Q(p) \leq m^{-A_\iota} Q_{m-1}$ at every boundary point p for m , which is white or black as $j(p) < J$. We prove it in three cases, with the Case-2 split $\mathbf{s}_\iota$ of Table 1: *Case 1*, p is white; *Case 2*, p is black with $\text{gap}(p) \leq \mathbf{s}_\iota m$; and *Case 3*, p is black with $\text{gap}(p) > \mathbf{s}_\iota m$. Case 1 needs one hold (Section 14); Cases 2 and 3 follow the hold walk through the trace of Section 12, and are completed in Sections 14 and 13.

12. THE TRACE

This section proves the explicit form of [9, Lemma 7.9] used in Case 3 (Section 13): from any starting point, the average of the discounted reward $e^{-z \text{Rw}}$ at the $(k + 1)$ st trace time is at most $(\frac{77}{100})^k$ (Lemma 12.16). It rests on the bound $\frac{77}{100}$ for one step of the entry kernel (Proposition 12.15). At most gaps this bound comes from the reward collected at the first raw crossing of the top edge of a triangle, whose angle is an odd dyadic fraction (Lemma 9.14); at six gaps it comes from finite trees of raw prefixes. The same crossing bounds give the bound $1 - d_{\text{pass}}$ for one passage at every gap, used in Case 2 (Lemma 14.3). The discount rate of the trace is the cancellation z at a white point (Lemma 11.2). 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. Its reward adds the paid raw rewards R_c of Definition 11.6, which can be positive at black as well as white points; each white closing point contributes at least 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 $|\mathbf{b}| = L$, $\mathbf{b}_{[i,i']} = (b^i, \dots, b^{i'})$, and $\mathbf{w}\mathbf{w}'$ for the concatenation of lists. 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

$$\text{rw}(x, b) = \sum_{i=1}^{|b|} R_{\mathbf{c}_i}(x + (i, \mathbf{c}_1 + \dots + \mathbf{c}_i)),$$

and the *accumulated reward* is $\text{Rw}(\bar{o}, \mathbf{b}; u) = \sum_{i=1}^{\min(u,L)} \text{rw}(y_{i-1}(\bar{o}, \mathbf{b}), b^i)$. The *raw word* of $\mathbf{b}$ is the concatenation of its holds, of length $M = j(\text{pt}_L(\mathbf{b}))$; $\sigma_i(\mathbf{b})$, $0 \leq i \leq M$, denotes the sum of its first i raw letters.

The raw letters of a hold other than its closing letter lie outside $\{4, 5\}$, so Definitions 11.8 and 11.9 give

$$\mathbf{w}(x + \text{pt}(b)) \mathbf{l}(x; b) = e^{-z \text{rw}(x, b)}. \quad (12.1)$$

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 $\mathfrak{B}^{L_1} \times \mathfrak{B}^{L_2}$,*

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

(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 $l(y_{u'}) - l(y_u) \geq 4k$ if $\mathbf{q}(\mathbf{b}) \neq 0$.*

(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.*

(6) *With $\mathbf{c}$ the raw word of $\mathbf{b}$ and M its length, $\text{Rw}(\bar{o}, \mathbf{b}; L) = \sum_{i=1}^M R_{\mathbf{c}_i}(\bar{o} + (i, \sigma_i(\mathbf{b})))$.*

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) follows from $\text{pt}_{L_1+u}(\mathbf{w}\mathbf{w}') = \text{pt}_{L_1}(\mathbf{w}) + \text{pt}_u(\mathbf{w}')$. (3) Every hold point (j, l) has $j \geq 1$, and $l \geq 4$ if the hold has positive mass (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 are disjoint subsets of $\mathfrak{B}^{L'}$, which has mass 1 (Lemma 10.4(5)). (5) Rewards are nonnegative, and the closing letter of the i th hold sits at y_i and has reward at least 1 there if y_i is white (Definition 11.6). (6) The i th raw letter of the k th hold has index $q = j(\text{pt}_{k-1}(\mathbf{b})) + i$ in the raw word, and the point attached to it in Definition 12.1 is $\bar{o} + (q, \sigma_q(\mathbf{b}))$. $\square$

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. A black point p lies in $\Delta(p)$ and $l_{\Delta(p)} = l_*(p)$ (Lemma 9.12 and Definition 9.9), so its gap $\text{gap}(p) = l_*(p) - l(p) \geq 0$ (Definition 9.8) is the distance from p up to the top edge of $\Delta(p)$.

Call the path $y_u = y_u(\bar{o}, \mathbf{b})$ of $\mathbf{b} \in \mathfrak{B}^L$ *complete* if $j(\bar{o}) + L > J$. If y_u is black on a complete path, then $u < L$ and $j(\bar{o}) + u \leq J$, since 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; 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. Since hits are less than L , the sequence is finite.

For $p \in \mathcal{P}$ 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 . For $\bar{o} \in \mathcal{P}$ let $\mathcal{U}(\bar{o})$, the *bridge lists*, be the set of lists $\mathbf{w}$ of positive mass 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 $\mathbf{w}$ of positive mass 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. Both sets are prefix-disjoint, and their lists $\mathbf{w}$ satisfy $j(\bar{o}) + |\mathbf{w}| \leq j(\bar{o}_{\mathbf{w}}) \leq J$. 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 a list $\mathbf{b} \in \mathfrak{B}^L$ of positive mass 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 $y'_u = y_u(y_{\tilde{u}}, \mathbf{b}_{[\tilde{u}+1, L]})$. Then $j(y_{\tilde{u}}) + (L - \tilde{u}) > J$, and for every $k \geq 1$, u_{i-1+k} exists if and only if the k th trace time u'_k of y' 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. For $\omega \in \mathbb{N}$ let Π_ω be the set of nonempty lists $\mathbf{f}$ of positive mass 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 $\sum_{\mathbf{f} \in \Pi_\omega} \mathbf{q}(\mathbf{f}) = 1$. The reward of a passage is bounded below at the place where its raw word first exceeds ω .

Lemma 12.5 (First raw crossings). *Let $\omega \in \mathbb{N}$. For integers $c \geq 2$, $1 \leq o \leq c$ and $r \geq 1$ let $\mathcal{X}_\omega(c, o, r)$ be the set of $\mathbf{f} \in \Pi_\omega$ whose raw word has length at least r , r th raw letter c , and $\sigma_r(\mathbf{f}) = \omega + o$.*

- (1) *The sets $\mathcal{X}_\omega(c, o, r)$ partition Π_ω ; for $\mathbf{f} \in \mathcal{X}_\omega(c, o, r)$, r is the least index with $\sigma_r(\mathbf{f}) > \omega$, and the position r lies in the last hold of $\mathbf{f}$.*
- (2) *For $\Psi \geq 0$ on finite words,*

$$\sum_{\mathbf{f} \in \mathcal{X}_\omega(c, o, r)} \mathbf{q}(\mathbf{f}) \Psi(\mathbf{c}(\mathbf{f})) = \mathbf{q}(c) \sum_{\mathbf{a}} \sum_{\mathbf{s}} \mathbf{q}(\mathbf{a}) \mathbf{q}(\mathbf{s}) \Psi(\mathbf{a} c \mathbf{s}),$$

where $\mathbf{c}(\mathbf{f})$ is the raw word, $\mathbf{a}$ runs over the words of $r-1$ raw letters with sum $\omega + o - c$, and $\mathbf{s}$ is empty if $c \in \{4, 5\}$ and runs over the hold words if $c \notin \{4, 5\}$. In particular the mass of $\mathcal{X}_\omega(c, o, r)$ is $\mathbf{q}(c) \mathbf{R}(r-1, \omega + o - c)$, and if $c \notin \{4, 5\}$ and $H \in \mathbb{N}$, the lists of $\mathcal{X}_\omega(c, o, r)$ whose raw word has length greater than $r + H$ have mass $\mathbf{q}(c) \mathbf{R}(r-1, \omega + o - c) (\frac{11}{16})^H$.

- (3) *For $v \in \mathcal{P}$ and $\mathbf{f} \in \mathcal{X}_\omega(c, o, r)$, $\text{Rw}(v, \mathbf{f}; |\mathbf{f}|) \geq R_c(v + (r, \omega + o))$.*

Proof. Raw letters of lists of positive mass are at least 2, so $i \mapsto \sigma_i(\mathbf{f})$ increases strictly, from 0 to $l(\text{pt}_{|\mathbf{f}|}(\mathbf{f})) > \omega$. If $\mathbf{f} \in \mathcal{X}_\omega(c, o, r)$, then $\sigma_{r-1}(\mathbf{f}) = \omega + o - c \leq \omega < \sigma_r(\mathbf{f})$, so r is the least index with $\sigma_r > \omega$; conversely that index r , its letter c and $o = \sigma_r - \omega \in [1, c]$ define a class containing $\mathbf{f}$. Every hold but the last ends at height at most ω , and the hold containing position r ends above ω , so it is the last one. This proves (1). For (2), a word of raw letters whose last letter is closing is the raw word of exactly one list of positive mass (cut after each closing letter), with mass the product of the letter masses. By (1), the raw words of $\mathcal{X}_\omega(c, o, r)$ are exactly the words $\mathbf{a} c \mathbf{s}$ described: if

c is closing the word ends at r , and otherwise the letters after r up to the next closing letter form a hold word. Summing over $\mathbf{a}$ gives $R(r-1, \omega + o - c)$ (Definition 10.1), and the hold words have total mass 1 and those with more than H raw letters, i.e. at least H nonclosing ones, have mass $\sum_{N \geq H} (\frac{11}{16})^N \frac{5}{16} = (\frac{11}{16})^H$. (3) follows from Lemma 12.2(6), since rewards are nonnegative. $\square$

The angle at a crossing point within the exit width $f_\omega = \lfloor (5\omega + 16)/16 \rfloor$ is an odd dyadic fraction, perturbed by at most $9^r \varepsilon 2^{-\omega}$ (Lemma 9.14), and the raw-letter factor there is bounded by disjoint packets of the sum (11.1), as in the proof of Lemma 11.5. For a raw letter $c \geq 3$ write the factor as Φ_{c-1} , so that $\varphi_c(p) = \Phi_{c-1}(\vartheta(p))$, and put

$$\Phi_c^{\text{cap}}(\theta) = \max\{\Phi_{c-1}(\theta), 2^{-16}\} \quad (3 \leq c \leq 16), \quad \Phi_c^{\text{cap}} = 1 \text{ otherwise.}$$

By Definition 11.6, $e^{-zR_c(p)} \leq \Phi_c^{\text{cap}}(\vartheta(p))$ for every p with $j(p) \leq J$. The function Φ_c^{cap} is even and has period 1, as the frequencies $2^a - 1$ are integers.

Definition 12.6 (Packet gains). An *envelope table* is a finite set T of pairs (s, t) of integers $s \geq 1$, $t \geq 2$ with numbers $V_{s,t} \in [0, 1]$. For $3 \leq c \leq 16$ and $1 \leq o \leq c$, a *packet family* for (c, o) is a set of pairs (i, t) with $i \geq 0$, $i + t \leq c - 1$ and $(o - i, t) \in T$, whose index sets $\{i, \dots, i + t - 1\}$ are pairwise disjoint, and

$$a_{c,o}[V] = \max_{(i,t)} \sum t(1 - V_{o-i,t})$$

over the packet families (computed by the downward recursion $F(c-1) = 0$, $F(i) = \max(\{F(i+1)\} \cup \{t(1 - V_{o-i,t}) + F(i+t)\})$, $a_{c,o}[V] = F(0)$). The *packet gains* are $\delta_{c,o} = a_{c,o}[U]/(c-1)$, with the table U of Lemma 11.5.

Lemma 12.7 (Packet families). *Let V be an envelope table with $\Phi_t((m+\beta)/2^s) \leq V_{s,t}$ for every (s, t) in it, every odd m and every real $|\beta| \leq \bar{\beta}$. Let $3 \leq c \leq 16$, $1 \leq o \leq c$, m odd, $|\beta| \leq \bar{\beta}$ and $\theta \equiv (m+\beta)/2^o \pmod{1}$. Then $\Phi_{c-1}(\theta) \leq 1 - a_{c,o}[V]/(c-1)$. Moreover $0 \leq \delta_{c,o} \leq \frac{1}{2}$; in particular $\delta_{3,1} = \frac{1}{2}$ and $\delta_{4,1} = \frac{41}{100}$.*

Proof. Apply Lemma 11.3 to an optimal packet family, with $\Phi_t(2^i\theta) \leq V_{o-i,t}$ as $2^i\theta \equiv (m+\beta)/2^{o-i}$. The entries of U lie in $[\frac{1}{2}, 1]$, so each packet contributes at most $t/2$ and $\delta_{c,o} \leq \frac{1}{2}$; the two values are exact evaluations. $\square$

Lemma 12.8 (Reward at an exit point). *Let v be black, $\omega = \text{gap}(v)$, $1 \leq r \leq f_\omega$, $3 \leq c \leq 16$ and $1 \leq o \leq c$, and put $y = v + (r, \omega + o)$. Then $j(y) \leq J$, $\vartheta(y) \equiv (m+\beta)/2^o$ with m odd and $|\beta| \leq 9^r \varepsilon 2^{-\omega} < \frac{1}{3}$, and*

$$e^{-zR_c(y)} \leq \Phi_c^{\text{cap}}(\vartheta(y)) \leq 1 - \delta_{c,o}.$$

Proof. Here $y = (j(v) + r, l_{\Delta(v)} + o)$, and Lemma 9.14(a),(b) give $j(y) \leq J$ and the localization. Lemma 12.7 with $V = U$ and $\bar{\beta} = \frac{1}{3}$ (Lemma 11.5), and $2^{-16} < \frac{1}{2} \leq 1 - \delta_{c,o}$, finish the proof. $\square$

At a few gaps we need sharper bounds for single factors, obtained on finitely many cells.

Lemma 12.9 (Capped factors on cells).

- (1) With $a_j = 2^j - 1$ and $\ell_c = \frac{44}{7(c-1)} \sum_{j=0}^{c-2} |a_j - a_{\lfloor (c-2)/2 \rfloor}|$ for $3 \leq c \leq 16$ ($\ell_c = 0$ otherwise), $|\Phi_c^{\text{cap}}(\theta) - \Phi_c^{\text{cap}}(\theta')| \leq \ell_c |\theta - \theta'|$.
- (2) For rational θ let $\bar{U}_c(\theta)$ be the least $k/10^8$, $k \in \mathbb{N}$, with $(k/10^8)^2 \geq (c-1)^{-2}(c-1 + 2 \sum_{i < j < c-1} \cos^+((a_j - a_i)\theta))$, clipped to $[\frac{1526}{10^8}, 1]$, where $\cos^+(x)$ is 1 if $x \in \mathbb{Z}$ and otherwise the degree-24 Taylor polynomial of $\cos$ at $2\pi_0\tau$, $\tau \in [-\frac{1}{2}, \frac{1}{2})$ the reduction of x modulo 1 and $\pi_0 = 3141592653589793/10^{15}$, plus $(22/7)^{26}/26! + 10^{-15}$ (possibly rounded up). Then $\Phi_c^{\text{cap}}(\theta) \leq \bar{U}_c(\theta)$.
- (3) $\Phi_2((m+\beta)/2) \leq \frac{11}{7} |\beta|$ for odd m and real β .

- (4) For every odd m and real $|\beta| \leq \bar{\beta}$, $\Phi_c^{\text{cap}}((m + \beta)/2^o) < u$ for the five quadruples $(c, o, \bar{\beta}, u)$:
 $(4, 1, \frac{1}{24}, \frac{9}{25})$, $(5, 2, \frac{1}{24}, \frac{23}{50})$, $(6, 3, \frac{1}{24}, \frac{14}{25})$, $(7, 4, \frac{1}{24}, \frac{4}{5})$, $(4, 1, \frac{1}{12}, \frac{21}{50})$.

Proof. (1) Multiplying the sum by $e^{-2\pi i a[(c-2)/2]\theta}$ does not change its modulus; the reverse triangle inequality, $|e^{ix} - e^{iy}| \leq |x - y|$ and $2\pi < \frac{44}{7}$ give the bound for Φ_{c-1} , and a maximum with a constant preserves it. (2) $(c-1)^2 \Phi_{c-1}(\theta)^2 = c-1 + 2 \sum_{i < j} \cos(2\pi(a_j - a_i)\theta)$. The standard value $\pi = 3.14159265358979323\dots$ gives $\pi_0 < \pi < \pi_0 + 10^{-15} < \frac{22}{7}$; cosine is 1-Lipschitz, and the Lagrange remainder of the Taylor polynomial is at most $(22/7)^{26}/26!$. Since $2^{-16} < \frac{1526}{10^8}$ and $\Phi \leq 1$, the clipping is harmless. (3) $\Phi_2(\theta) = |\cos \pi \theta|$, and $|\cos(\pi(m + \beta)/2)| = |\sin(\pi\beta/2)| \leq \frac{\pi}{2}|\beta|$. (4) By periodicity it suffices to take $1 \leq m < 2^o$ odd. Cut $[-\bar{\beta}, \bar{\beta}]$ into 64 equal closed intervals; on each, (1) and (2) at the midpoint bound the factor by $\bar{U}_c(\text{midpoint}) + \ell_c 2^{-o} \bar{\beta}/64$. Exact evaluation at the 1024 cells gives the maxima $\frac{73783307}{210000000}$, $\frac{2527137353}{5600000000}$, $\frac{222172771}{400000000}$, $\frac{8752089581}{14400000000}$ and $\frac{51834043}{131250000}$, in the order listed, each below u . $\square$

At gaps 12 and 19 every crossing index is so small that the perturbation is below $\frac{1}{12}$ and $\frac{1}{24}$, and at every gap the raw letter 3 crossing with overshoot 1 is nearly a zero of Φ_2 . This sharpens the packet coefficients as follows. For $\omega \in \mathbb{N}$ put $\varsigma_\omega = \min\{1, \max\{2^{-16}, \frac{11}{7}\varepsilon 9^{f_\omega} 2^{-\omega}\}\}$, let $\delta_{\omega,3,1}^{\text{rep}} = \max\{\frac{1}{2}, 1 - \varsigma_\omega\}$, $\delta_{12,4,1}^{\text{rep}} = \frac{29}{50}$, $\delta_{19,4,1}^{\text{rep}} = \frac{16}{25}$, and $\delta_{\omega,c,o}^{\text{rep}} = \delta_{c,o}$ otherwise, and define the *crossing gain*

$$D(\omega) = \sum_{c=3}^{16} q(c) \sum_{o=1}^c \delta_{\omega,c,o}^{\text{rep}} \sum_{r=1}^{f_\omega} R(r-1, \omega + o - c). \quad (12.3)$$

Lemma 12.10 (Reward over one passage). *Let v be black and $\omega = \text{gap}(v)$, and put $S(v) = \sum_{\mathbf{f} \in \Pi_\omega} q(\mathbf{f}) e^{-z \text{Rw}(v, \mathbf{f}; |\mathbf{f}|)}$. Then $S(v) \leq 1 - D(\omega)$; if $\omega = 3$, then $S(v) \leq 1 - \frac{37}{160}$; and for every ω ,*

$$S(v) \leq 1 - d_{\text{pass}}, \quad d_{\text{pass}} = \frac{3661359}{10^8}.$$

Proof. For a set I of triples (c, o, r) and numbers $u_{c,o,r} \in [0, 1]$ bounding $e^{-z \text{Rw}}$ on $\mathcal{X}_\omega(c, o, r)$, Lemma 12.5(1),(2) give $S(v) \leq 1 - \sum_I (1 - u_{c,o,r}) q(c) R(r-1, \omega + o - c)$, as the classes are disjoint and $e^{-z \text{Rw}} \leq 1$. Take the triples with $3 \leq c \leq 16$, $1 \leq o \leq c$, $1 \leq r \leq f_\omega$ and $u_{c,o,r} = 1 - \delta_{\omega,c,o}^{\text{rep}}$. By Lemma 12.5(3) and Lemma 12.8, $e^{-z \text{Rw}} \leq \Phi_c^{\text{cap}}(\vartheta(y)) \leq 1 - \delta_{c,o}$ at $y = v + (r, \omega + o)$. For $(c, o) = (3, 1)$, Lemma 12.9(3) also gives $\Phi_3^{\text{cap}}(\vartheta(y)) \leq \varsigma_\omega$. For $(c, o) = (4, 1)$ and $\omega = 12, 19$, $f_{12} = 4$, $f_{19} = 6$, and $9^4 \varepsilon 2^{-12} < \frac{1}{12}$, $9^6 \varepsilon 2^{-19} < \frac{1}{24}$, so Lemma 12.9(4) gives the factors $\frac{21}{50}$ and $\frac{9}{25}$. This proves $S(v) \leq 1 - D(\omega)$. At $\omega = 3$ take instead the four triples $(c, c-3, 1)$, $4 \leq c \leq 7$: here $f_3 = 1$, $|\beta| \leq 9\varepsilon/8 < \frac{1}{24}$, $R(0, 0) = 1$, and Lemma 12.9(4) gives $S(v) \leq 1 - (\frac{3}{16} \cdot \frac{16}{25} + \frac{1}{8} \cdot \frac{27}{50} + \frac{5}{64} \cdot \frac{11}{25} + \frac{3}{64} \cdot \frac{1}{5}) = 1 - \frac{37}{160}$.

For the last bound we show $D(\omega) \geq d_{\text{pass}}$ for $\omega \neq 3$, the case $\omega = 3$ being $\frac{37}{160} > d_{\text{pass}}$. Each $D(\omega)$ is an explicit rational, as $R(k, l) = \binom{l-1}{2k-1} 2^{-l}$ for $k, l \geq 1$ (Lemma 10.2). Exact evaluation gives $D(0) = \frac{59987721}{16384000000} \geq d_{\text{pass}}$, $D(1) = \frac{832326661}{65536000000}$, $D(6) = \frac{108692439667}{524288000000}$, $D(9) = \frac{911532121133}{4194304000000}$, $D(13) = \frac{7713203895523}{33554432000000}$, $D(16) = \frac{195010696398427}{858993459200000}$, and, over the other 505 integers $2 \leq \omega \leq 511$, the minimum $D(15) = \frac{61817834604121}{268435456000000}$. For $\omega \geq 512$: first, for $2 \leq l \leq \omega$, $\sum_{r=1}^{f_\omega} R(r-1, l) \geq \frac{1}{4} - \frac{25}{2\omega} \geq \frac{231}{1024}$. Indeed the full row sums to $\frac{1}{4}$ (Lemma 10.2), and the omitted part is at most $\frac{1}{2} \mathbb{P}(X \geq 2f_\omega - 1)$ for X binomial with parameters $l-1, \frac{1}{2}$, whose threshold exceeds the mean by at least $\frac{\omega-3}{8} \geq \frac{\omega}{10}$, while $\text{Var } X \leq \frac{\omega}{4}$; apply Chebyshev's inequality. Second, writing $\omega = 16k + s$, $0 \leq s < 16$, $k \geq 32$, we have $f_\omega = 5k + 1 + \lfloor 5s/16 \rfloor$ and $9^{\lfloor 5s/16 \rfloor} \leq 2^s$, so $9^{f_\omega} 2^{-\omega} \leq 9(9^5/2^{16})^{32}$ and $\varsigma_\omega < \frac{1}{50}$ by exact comparison. Hence

$$D(\omega) \geq \frac{231}{1024} \left(\sum_{c=3}^{16} q(c) \sum_{o=1}^c \delta_{c,o} + \frac{1}{4} \left(\frac{49}{50} - \frac{1}{2} \right) \right) = \frac{231}{1024} \left(\frac{15036545331}{16384000000} + \frac{3}{25} \right) = \frac{3927606451461}{16777216000000},$$

which exceeds d_{pass} . $\square$

Lemma 12.11 (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 $\mathbf{b}$, $l(y_u(v, \mathbf{b})) > l_*(v)$ 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 $l(\text{pt}_k(\mathbf{b})) > \omega$. Then $\mathbf{f} = \mathbf{b}_{[1,k]}$ is the only prefix of $\mathbf{b}$ in Π_ω , every $y_u(v, \mathbf{b})$ with $u \geq k$ lies above $l_*(v)$ (Lemma 12.2(3)), and the conditions defining $\mathcal{E}(v)$ say that $\mathbf{b}_{[k+1,|\mathbf{b}|]} \in \mathcal{U}(y_k(v, \mathbf{b}))$. The same facts show that each $\mathbf{fw}$ lies in $\mathcal{E}(v)$. $\square$

12.4. The entry kernel. For $p \in \mathcal{P}$ and a list $\mathbf{w}$ write $\mathbf{q}^p(\mathbf{w}) = \mathbf{q}(\mathbf{w})e^{-z\text{Rw}(p, \mathbf{w}; |\mathbf{w}|)}$. 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 and homogeneous. The *bridge mass* of $p \in \mathcal{P}$ is $\text{Br}(p) = \sum_{\mathbf{w} \in \mathcal{U}(p)} \mathbf{q}^p(\mathbf{w}) \leq 1$, by Lemma 12.2(4). By Lemma 12.11 and Lemma 12.2(1), (2),

$$\mathbf{K}\mathbf{1}(v) = \sum_{\mathbf{f} \in \Pi_{\text{gap}(v)}} \mathbf{q}^v(\mathbf{f}) \text{Br}(v_{\mathbf{f}}) \leq S(v), \quad (12.4)$$

with $S(v)$ as in Lemma 12.10. Hence $\mathbf{K}\mathbf{1}(v) \leq 1 - D(\text{gap}(v)) < \frac{77}{100}$ unless $\text{gap}(v) \in \{0, 1, 6, 9, 13, 16\}$ (for $\text{gap}(v) = 3$ use $\frac{37}{160} > \frac{23}{100}$). At these six gaps a single crossing is not enough, and we follow the raw word further: as long as a raw letter cannot end the entry, the path is continued, and each factor read off along the way is paid.

Definition 12.12 (Entry trees). A *policy* is a set W of triples (r, h, c) of integers with $c \geq 2$ and $h + c$ bounded; the triple (r, h, c) means that the raw letter c at depth r and height h is *continued*. A *tree datum of height H* is a pair (cut, W) of a function $\text{cut} : \mathbb{N} \rightarrow \mathbb{N}$ and a policy such that $c \leq \text{cut}(h)$ for $(r, h, c) \in W$, $\text{cut}(0) \leq \min(16, H)$, and $\text{cut}(h + c) \leq 16$, $h + c + \text{cut}(h + c) \leq H$ for $(r, h, c) \in W$. For real X put $\psi_{r,h,c}(X) = \Phi_c^{\text{cap}}(9^r X / 2^{h+c})$ if $2 \leq c \leq \text{cut}(h)$ and $\psi_{r,h,c} = 1$ otherwise, and define the *tree cost* by downward recursion on h :

$$F_{r,h}(X) = 1 - \sum_{c=2}^{\text{cut}(h)} \mathbf{q}(c) [1 - \psi_{r,h,c}(X) G_{r,h,c}(X)],$$

with $G_{r,h,c} = F_{r+1,h+c}$ if $(r, h, c) \in W$ and $G_{r,h,c} = 1$ otherwise, and its *Lipschitz constant* $L_{r,h} = 9^r 2^{-h} \sum_{c=3}^{\text{cut}(h)} \mathbf{q}(c) \ell_c 2^{-c} + \sum_{(r,h,c) \in W} \mathbf{q}(c) L_{r+1,h+c}$. For $\omega \in \mathbb{N}$, integers C, H', m and reals $a \leq b$, the *white-test policy* $W_{m,a,b}^{\omega,C,H'}$ consists of the triples with $2 \leq c \leq C$, $h + c \leq H'$ and either $c \notin \{4, 5\}$, or $h + c \leq \omega$, or

$$\text{dist}\left(\frac{9^r(2^\omega m + \frac{a+b}{2})}{2^{h+c}}, \mathbb{Z}\right) - \frac{9^r(b-a)}{2^{h+c+1}} > \varepsilon.$$

A policy is *valid* at a black point v if, for every $(r, h, c) \in W$ with $c \in \{4, 5\}$, the point $v + (r+1, h+c)$ is white or $l(v) + h + c \leq l_*(v)$. The *root correction* is $T_{\omega,k} = \sum_{c=\max(k,\omega,2)+1}^{16} \mathbf{q}(c) \delta_{c,c-\omega}$.

Lemma 12.13 (Tree costs). *Let v be black, $\omega = \text{gap}(v)$ and $B \geq 1$.*

- (1) *There are an odd m with $0 < m < 2^B$ and a real α with $|\alpha| \leq 9\varepsilon$ such that $\vartheta(v + (r, h)) \equiv 9^{r-1}(2^\omega m + \alpha)/2^h \pmod{1}$ for all $r \geq 1$ and $0 \leq h \leq \omega + B$.*
- (2) *If m, α are as in (1), $H' \leq \omega + B$ and $a \leq \alpha \leq b$, then $W_{m,a,b}^{\omega,C,H'}$ is valid at v .*
- (3) *If m, α are as in (1) with $\omega + B = H$, and (cut, W) is a tree datum of height H valid at v , then $\mathbf{K}\mathbf{1}(v) \leq F_{0,0}(2^\omega m + \alpha) - T_{\omega, \text{cut}(0)}$.*
- (4) *$0 \leq F_{r,h} \leq 1$ and $|F_{r,h}(X) - F_{r,h}(X')| \leq L_{r,h}|X - X'|$; $L_{r,h}$ increases with W .*
- (5) *For the cutoff $\text{cut}^0(h) = \min\{12, 17 - h\}$ ($h \leq 17$), $W = W_{m,a,b}^{0,12,14}$ and every real X , $F_{0,0}(X)$ is unchanged when (m, a, b, X) is replaced by $(2^{17} - m, -b, -a, 2^{17} - X)$.*

Proof. (1) Let $\hat{v} = (j(v), l_*(v))$ and $N = 3^n \vartheta(\hat{v})$; by Lemma 9.14(c) with $\Delta = \Delta(v)$ and $p = v$, N is odd and $|N|/3^n \leq \varepsilon 2^{-\omega}$. Choose m with $m3^n \equiv -9N \pmod{2^B}$; it is odd. Put $\alpha = 9 \cdot 2^\omega N/3^n$. For $y = v + (r, h)$, residue transport (Lemma 9.4) along the row and the column gives $2^h \text{res}(y) = 9^r 2^\omega \text{res}(\hat{v})$, and the integer $K = 9^{r-1} 2^\omega (m3^n + 9N)/2^h$ satisfies $2^h K \equiv 9^r 2^\omega N \pmod{3^n}$; since 2^h is invertible modulo 3^n , K represents $\text{res}(y)$, and $K/3^n = 9^{r-1} (2^\omega m + \alpha)/2^h$. (2) If $h + c \leq \omega$ the point lies at height at most $l_*(v)$. Otherwise, by (1) its angle is within $9^r(b-a)/2^{h+c+1}$ of $9^r(2^\omega m + \frac{\alpha+b}{2})/2^{h+c}$ modulo 1, so the test gives $|\vartheta| > \varepsilon$ and the point is white. (3) Let $\mathbf{w} \in \mathcal{E}(v)$ with raw word $\mathbf{c}$. Follow $\mathbf{c}$ from the root: the edge $(i-1, \mathbf{c}_1 + \dots + \mathbf{c}_{i-1}, \mathbf{c}_i)$ is either in W or not. It cannot be in W at the last position of $\mathbf{c}$: a nonclosing letter does not end a hold, and by validity a continued closing letter ends at a white point or at height at most $l_*(v)$, so not at the end of an entry. Hence $\mathbf{c}$ begins with exactly one *leaf* $\mathbf{l}$, a word whose edges are continued except the last. Along $\mathbf{l}$, $\text{Rw}(v, \mathbf{w}; |\mathbf{w}|)$ is at least the sum of the rewards at the points $v + (i, \mathbf{l}_1 + \dots + \mathbf{l}_i)$ (Lemma 12.2(6)); these points have $j \leq J$ since the entry ends at a black point further right, and their heights are at most H , so by (1) and Definition 11.6 their factors are at most $\psi_{i-1, \mathbf{l}_1 + \dots + \mathbf{l}_{i-1}, \mathbf{l}_i}(2^\omega m + \alpha)$. The entry lists beginning with $\mathbf{l}$ have mass at most $\prod_i q(\mathbf{l}_i)$ (extend them to raw length J and use (10.3)), and, by induction on h , $F_{0,0}$ is the sum over leaves of $\prod_i q(\mathbf{l}_i) \psi_i$. Finally, for $c > \max(\text{cut}(0), \omega, 2)$ the one-letter leaf (c) has factor 1 in $F_{0,0}$, while Lemma 12.8 with $r = 1$, $o = c - \omega$ bounds its factor by $1 - \delta_{c, c-\omega}$; this gives the correction $T_{\omega, \text{cut}(0)}$. (4) By downward induction on h , using $\psi \in [0, 1]$, $\sum_c q(c) = 1$, $|pg - p'g'| \leq |p - p'| + |g - g'|$ for numbers in $[0, 1]$, and (Lemma 12.9(1)) the constant $\ell_c 9^r 2^{-h-c}$ of $\psi_{r, h, c}$. (5) All retained heights satisfy $h + c \leq 17$, so $9^r(2^{17} - X)/2^{h+c} \equiv -9^r X/2^{h+c}$; the capped factors are even and 1-periodic, and $\text{dist}(\cdot, \mathbb{Z})$ is even, so the policies and the costs coincide. $\square$

Lemma 12.14 (The six special gaps). *If v is black and $\text{gap}(v) \in \{0, 1, 6, 9, 13, 16\}$, then $\mathbf{K1}(v) < \frac{77}{100}$.*

Proof. For a tree datum and a cell $[a, b]$ we bound $F_{0,0}$ by *node values*: if $f_c \geq \psi_{r, h, c}$ and $g_c \geq G_{r, h, c}$ (with $g_c \geq 1$ unless the edge is continued), then $F_{r, h} \leq 1 - \sum_c q(c)(1 - f_c g_c)$, because $\psi, G \in [0, 1]$ (Lemma 12.13(4)).

Gap 0. Lemma 12.13(1) with $B = 17$ gives an odd $m < 2^{17}$ and $|\alpha| \leq 9\varepsilon < \frac{1}{3}$. Take $\text{cut}^0(h) = \min\{12, 17 - h\}$ for $h \leq 17$ and $W = W_{m, a, b}^{0, 12, 14}$; every triple of W has $c \leq 12$ and $h + c \leq 14$, so (cut^0, W) is a tree datum of height 17, valid at v by Lemma 12.13(2). By Lemma 12.13(3), it suffices to find $a \leq \alpha \leq b$ in $[-\frac{1}{3}, \frac{1}{3}]$ with $F_{0,0}(m + \alpha) < \frac{77}{100} + T_{0,12}$, where $T_{0,12} = \sum_{c=13}^{16} q(c) \delta_{c,c} = \frac{1197}{13107200}$. By Lemma 12.13(5) we may assume $m < 2^{16}$ (replace (m, α) by $(2^{17} - m, -\alpha)$). This is a finite computation. The odd integers are split into 84 residue classes $m_0 + 2^k \mathbb{Z}$, $k \leq 17$, obtained from $1 + 2\mathbb{Z}$ by repeated halving; four of them contain no $m < 2^{16}$. The cells are the dyadic subintervals $[\frac{p-1}{3 \cdot 2^d}, \frac{p+1}{3 \cdot 2^d}]$ of $[-\frac{1}{3}, \frac{1}{3}]$. For a class, a cell and every triple with $h + c \leq k$, the white test of Definition 12.12 and the factor $\psi_{r, h, c}(m + \beta)$, up to the Lipschitz term $\ell_c 9^r 2^{-h-c}$ times the half-width (Lemma 12.9(1)), depend only on the class and the cell. Bound each factor by its value at the midpoint, computed from rigorous enclosures of the units $e^{2\pi i 2^j q}$ (a degree-17 Taylor polynomial at a rational multiple of π , followed by repeated squaring, all with error bounds on the scale 2^{-60}), plus the Lipschitz term; treat edges with $h + c > k$ or a failed white test as not continued; and compute the node values from depth 8 down to the root, rounding upward. A cell is accepted if the root value is below $\frac{77}{100} + T_{0,12}$, and otherwise replaced by its two halves. The computation, in exact integer arithmetic, accepts 211 cells of depth at most 6 for the 80 classes, covering $[-\frac{1}{3}, \frac{1}{3}]$ for each class, with least margin greater than $2.26 \cdot 10^{-4}$.

Gaps 1, 6, 9, 13, 16. Let $p = 4$ for $\omega \in \{1, 6, 9\}$ and $p = 1$ for $\omega \in \{13, 16\}$. Lemma 12.13(1) with $B = p + 8$ gives $m < 2^{p+8}$ and α . Take the constant cutoff 8 and $W = W_{m, a, b}^{\omega, 8, \omega+p}$, a tree datum of height $\omega + p + 8$, valid at v . The root corrections are $T_{1,8} = \frac{4595457}{1310720000}$, $T_{6,8} = \frac{13575489}{1638400000}$, $T_{9,8} = \frac{800801}{204800000}$, $T_{13,8} = \frac{697807}{3276800000}$ and $T_{16,8} = 0$. For each odd $m < 2^{p+8}$, start from $[-\frac{1}{3}, \frac{1}{3}]$

(from its halves if $\omega \in \{6, 9\}$, from its quarters if $\omega \in \{13, 16\}$); on a cell with half-width δ and midpoint u , bound $F_{0,0}(2^\omega m + u)$ by node values with the factors computed from the cosine formula of Lemma 12.9(2) (a degree-14 Taylor polynomial at $2\pi \cdot \frac{3141592654}{10^9}$, error at most 10^{-9}), and add $\Lambda\delta$, where Λ bounds $L_{0,0}$ for the policy of all triples with $c \leq 8$, $h + c \leq \omega + p$, which contains W (Lemma 12.13(4)). Accept the cell if the result is below $\frac{77}{100} + T_{\omega,8}$, and otherwise halve it. The computation accepts 2204, 4096, 4112, 1024 and 1054 pairs of a value of m and a cell for $\omega = 1, 6, 9, 13, 16$, with least margin greater than $1.39 \cdot 10^{-5}$ (at $\omega = 16$). $\square$

Proposition 12.15 (The entry kernel). *For every black point v and every $k \in \mathbb{N}$, $K\mathbf{1}(v) \leq \frac{77}{100}$ and $K^k\mathbf{1}(v) \leq (\frac{77}{100})^k$.*

Proof. The first bound is Lemma 12.14 at the six special gaps, and (12.4) with Lemma 12.10 elsewhere: $\frac{37}{160}$, $\frac{61817834604121}{268435456000000}$ and $\frac{3927606451461}{16777216000000}$ all exceed $\frac{23}{100}$. The powers follow by induction, as K is positive and homogeneous. $\square$

12.5. 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 $\mathcal{Z}_{k+1}(\bar{o}, \mathbf{b}) = e^{-z \text{Rw}(\bar{o}, \mathbf{b}; u_{k+1})}$ if u_{k+1} exists, and $\mathcal{Z}_{k+1}(\bar{o}, \mathbf{b}) = 0$ otherwise.

Lemma 12.16 (Explicit form of Tao's Lemma 7.9). *Let $\bar{o} \in \mathcal{P}$ and $L \in \mathbb{N}$ with $j(\bar{o}) + L > J$. Then*

- (1) *if $\bar{o} = v$ is black, then $\sum_{\mathbf{b} \in \mathfrak{B}^L} \mathbf{q}(\mathbf{b}) \mathcal{Z}_{k+1}(v, \mathbf{b}) = K^k\mathbf{1}(v)$ for every $k \in \mathbb{N}$;*
- (2) *for every $k \in \mathbb{N}$, $\sum_{\mathbf{b} \in \mathfrak{B}^L} \mathbf{q}(\mathbf{b}) \mathcal{Z}_{k+1}(\bar{o}, \mathbf{b}) \leq (\frac{77}{100})^k$.*

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

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

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

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

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

and (1), Proposition 12.15 and $\text{Br}(\bar{o}) \leq 1$ conclude. $\square$

13. BOUNDARY POINTS FAR BELOW THE TOP EDGE

This section proves Case 3 of the descent step (Section 11) for each index $\iota \in \{o, b\}$: the boundary point p for m is black and lies at gap $\text{gap}(p) > s_\iota m$ below the top edge of its triangle (Lemma 13.26). Following Tao [9, Case 3 of Proposition 7.8 and Lemma 7.10], we continue the hold walk after its first passage above the top edge, keeping the reward paid during that passage. Outside rare events, a low reward forces many trace times on a fixed schedule (Lemma 13.3), which Lemma 12.16 makes unlikely; the main rare event, a visit to a large triangle, is bounded by a packing estimate (Lemma 13.19). The paid first passage has mass less than $\mathbf{q}_{\text{mass}} = \frac{38177}{125000}$ (Lemma 13.13), and for $\iota = b$ the last trace factors are replaced by stopped stages (Lemma 13.22).

Throughout, ι is fixed and $\mathbf{Q}, \mathbf{Q}_m$ are formed with the exponent A_ι ; we put $\lambda_0 = \lambda - \frac{1}{4}$; then $\frac{5}{16} \leq \lambda < \frac{8}{25}$ and $\frac{1}{\lambda} < \frac{158497}{50000} < \frac{127}{40}$ (exact comparisons $9^5 \leq 2^{16}$, $2^{25} < 9^8$, $9^{50000} < 2^{158497}$). The schedule constants T_ι , r_ι , $p_0(\iota)$, α_ι , H_ι are those of Table 1, the size factors $K_i(\iota)$ are those of Definition A.2, and p_i , S_i , P_i , V_i are those of Definition A.3; we drop the index ι from K_i , p_i and S_i . For $u \in \mathbb{N}$ let $i(u)$ be the least $i \leq r_\iota$ with $u \leq p_i$ ($i(u) = r_\iota$ if $u > p_{r_\iota}$), and put

$$\beta_\iota(u) = S_{i(u)}, \quad h_\iota(u) = \lceil \alpha_\iota \beta_\iota(u) \rceil + H_\iota, \quad \mathbf{g}_\iota(u) = u + h_\iota(u) + T_\iota + 1.$$

Since $K_{i+1} \geq 268$ and $268\alpha_i \geq 1$, $S_i \leq S_{i+1}$; so β_i , h_i and $\mathbf{g}_i$ are nondecreasing, $\beta_i(u) = S_i$ for $p_{i-1} < u \leq p_i$ ($p_{-1} = -1$), and $\mathbf{g}_i(p_i) = p_{i+1}$. Also $T_i + 1 = p_0$, and $V_i \geq 2^{18}$ and $\Omega_i \geq \max(2^{34}, 512V_i)$ by Lemma A.4.

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 $\omega \in \mathbb{N}$; the law $\mathbb{F}_\omega$ dominates the paid first passage of Section 13.4 and serves for tail and packing bounds.

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

By Lemmas 10.6(1) and 10.4(5), $\mathbb{P}_{p,\omega}$ is a probability, and every fresh path is complete. Splitting $\mathbf{b}$ after $k \leq J$ holds and grouping by hold points (Lemma 12.2(1), (10.1)) gives, for $\Psi \geq 0$,

$$\sum_a \mathbb{P}_{p,\omega}(a) \Psi(x, (\text{pt}(b^1), \dots, \text{pt}(b^k))) = \sum_x \mathbb{F}_\omega(x) \sum_{\mathbf{h} \in \mathcal{P}^k} \eta^{\otimes k}(\mathbf{h}) \Psi(x, \mathbf{h}). \quad (13.1)$$

Definition 13.2 (Case 3 events). For $p \in \mathcal{P}$ and $m \in \mathbb{N}$ let

$$\begin{aligned} \text{Low}^p &= \{a : \text{Rw}_{P_i-1}^p(a) \leq T_i\}, & \text{Bad}_m^p &= \{a : 400(j(y_{P_i}^p(a)) - j(p)) \geq 317m\}, \\ \text{Clr} &= \left\{a : \sum_{i=u+1}^{u+h_i(u)} l(\text{pt}(b^i)) \leq \frac{13}{9}\beta_i(u) \text{ for some } u \text{ with } u + h_i(u) \leq P_i - 1\right\}. \end{aligned}$$

An atom has an *oversize hit* at time $u \leq P_i - 1$ if $y_u^p(a) \in \Delta$ for some $\Delta \in \mathfrak{T}$ with $S_\Delta \geq \beta_i(u)$.

Off Bad_m^p the advance $j(y_{P_i}^p) - j(p)$ stays below $\frac{317}{400}m$, so the factor $\max(m - \cdot, 1)^{-A_i}$ of the boundary reduction stays comparable to m^{-A_i} ; without oversize hits and slow clearing the path climbs out of every triangle it enters on schedule.

13.2. The trace times come on schedule.

Lemma 13.3 (The trace times come on schedule). *Let $p \in \mathcal{P}$, $m \in \mathbb{N}$ with $j(p) + m = J$, $P_i \leq J$, and let $a = (x, \mathbf{b}) \in \text{Low}^p \setminus \text{Bad}_m^p$ with $\mathbf{q}(\mathbf{b}) \neq 0$; write $y_u = y_u^p(a)$.*

- (1) $j(y_u) \leq J$ for $u \leq P_i$, and every set of $T_i + 1$ consecutive times in $[1, P_i - 1]$ contains a hit. In particular $u_1 \leq p_0$.
- (2) Let $q \leq p_{r_i}$ be a hit with $\mathbf{g}_i(q) \leq P_i - 1$ and no oversize hit at q , and suppose $a \notin \text{Clr}$. Then there is a hit c with $q < c \leq \mathbf{g}_i(q)$ and $l(y_c) > l_*(y_q)$.
- (3) If $a \notin \text{Clr}$, $0 \leq k \leq r_i$ and a has no oversize hit at a time $u \leq p_{k-1}$, then the fresh path has at least $k + 1$ trace times and $u_i \leq p_{i-1}$ for $1 \leq i \leq k + 1$.

Proof. (1) As $a \notin \text{Bad}_m^p$, $j(y_{P_i}) < j(p) + m = J$, and $j(y_u)$ increases with u . By Lemma 12.2(5) the number of white points among $y_1, \dots, y_{P_i-1}$ is at most $\text{Rw}_{P_i-1}^p(a) \leq T_i$, so a run of $T_i + 1$ times contains a point which is not white, hence black. As $T_i + 1 = p_0 \leq P_i - 1$, $u_1 \leq p_0$.

(2) Let $\Delta = \Delta(y_q)$; then $y_q \in \Delta$, $l_\Delta = l_*(y_q)$ and $S_\Delta < \beta_i(q)$. Membership gives $\text{gap}(y_q) \log 2 \leq S_\Delta$, and $\log 2 > \frac{9}{13}$, so $\text{gap}(y_q) < \frac{13}{9}\beta_i(q)$. As $a \notin \text{Clr}$ and $q + h_i(q) \leq P_i - 1 \leq J$, $l(y_{q+h_i(q)}) - l(y_q) > \frac{13}{9}\beta_i(q)$, so $l(y_u) > l_*(y_q)$ for $u \geq q + h_i(q)$ (Lemma 12.2(3)). By (1) the run $[q + h_i(q) + 1, \mathbf{g}_i(q)]$ contains a hit.

(3) By induction on i : (1) gives $u_1 \leq p_0$, and if $u_i \leq p_{i-1}$ with $i \leq k$, then $\mathbf{g}_i(u_i) \leq \mathbf{g}_i(p_{i-1}) = p_i \leq P_i - 1$ and there is no oversize hit at u_i , so (2) gives $u_{i+1} \leq p_i$. $\square$

Lemma 13.4 (Discounted mass at the trace times). *Let $p \in \mathcal{P}$, $k \in \mathbb{N}$, $\phi \geq 0$ on $\mathcal{P}$, $D \subseteq \mathcal{P}$, and let $\mathcal{S}$ be a set of atoms $(x, \mathbf{b})$ with $x \in D$ whose fresh path has at least $k + 1$ trace times, with*

$u_{k+1} \leq P_\ell - 1$. Then

$$\sum_{(x, \mathbf{b}) \in \mathcal{S}} \phi(x) \mathbf{q}(\mathbf{b}) e^{-z \text{Rw}_{P_\ell-1}^p(x, \mathbf{b})} \leq \left(\frac{77}{100}\right)^k \sum_{x \in D} \phi(x).$$

Proof. As the accumulated reward is nondecreasing, $e^{-z \text{Rw}_{P_\ell-1}^p} \leq \mathcal{Z}_{k+1}(p+x, \mathbf{b})$. Enlarge $\mathcal{S}$ to all atoms with $x \in D$ and apply Lemma 12.16(2) with $L = J$. $\square$

Lemma 13.5 (Slow clearing is rare). *Let $p \in \mathcal{P}$, $\omega \in \mathbb{N}$ and $P_\ell \leq J$. Then $\mathbb{P}_{p, \omega}(\text{Clr}) \leq E_\ell^{\text{cl}}$, with E_ℓ^{cl} as in Definition 13.23.*

Proof. Put $y = y_\ell^{\text{cl}}$, $v = v_\ell^{\text{cl}}$ (Definition 13.14). Fix u with $u+h \leq P_\ell - 1$, $h = h_\ell(u)$, $Y = \frac{13}{9}\beta_\ell(u)$. By (13.1) the atoms with $\sum_{i=u+1}^{u+h} l(\text{pt}(b^i)) \leq Y$ have mass at most $\sum_{\mathbf{h} \in \mathcal{P}^h} \eta^{\otimes h}(\mathbf{h}) \mathbf{1}[\sum l(\mathbf{h}_i) \leq Y] \leq y^{-Y} \mathbf{V}(y)^h \leq y^{vh-Y}$, by Lemma 13.15(3). If $p_{i-1} < u \leq p_i$, then $\beta_\ell(u) = S_i$ and $h_\ell(u) = h_\ell(p_i)$, so $vh - Y = t_i/(1-y)$ with $t_i = t_i^{\text{cl}}(\ell) > 0$, and $\log y \leq -(1-y)$ gives $y^{vh-Y} \leq e^{-t_i} \leq 1/P_{80}(t_i)$ (Lemma A.1(4)). The admissible u in the block $(p_{i-1}, p_i]$ ($[0, p_0]$ for $i = 0$) are those with $u \leq P_\ell - 1 - h_\ell(p_i)$, and their number is $n_i^{\text{cl}}(\ell)$. Sum over the blocks. $\square$

13.3. Row geometry of triangles. For a triangle Δ and $l \in \mathbb{Z}$ put $\text{tip}_\Delta = l_\Delta - S_\Delta/\log 2$ and $\text{row}_\Delta(l) = j_\Delta + \lambda(l - \text{tip}_\Delta)$, the height at which the hypotenuse meets the left edge and 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. Making explicit the geometric steps of the proof of [9, Lemma 7.10], we show that large triangles met there have tips at or above l_Δ and corners in well-separated columns.

Lemma 13.6 (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 $j_\Delta \leq j \leq \text{row}_\Delta(l)$. Every $(j, l) \in \Delta$ 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 [9, (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 13.7 (Tips of nearby triangles). *Let $\Delta \neq \Delta'$ be members of $\mathfrak{T}$, $p \in \Delta$ 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 [9, Lemma 7.10], we find a point in both triangles, contradicting Proposition 9.13. Put $\bar{j} = \text{row}_{\Delta'}(l_\Delta) > j_{\Delta'}$. By Lemma 13.6, $j \leq \text{row}_{\Delta'}(l) \leq \bar{j} + \lambda W$. Let $j' = \min(j, \bar{j})$; then $j_{\Delta'} \leq j' \leq \bar{j}$, $0 \leq j - j' \leq \lambda W + 1$, and $(j', l_\Delta) \in \Delta'$ as $l_\Delta \leq l \leq l_{\Delta'}$. On the other hand $\lambda_0 \leq \frac{1}{4}$, so $Y + \lambda W + 1 \leq \omega/4$; hence $j' \geq j(p) \geq j_\Delta$ and $j' \leq j(p) + (\frac{1}{4} + \lambda_0)\omega = j(p) + \lambda\omega$. Multiplying by $\log 9$ and using the membership of p , $(j' - j_\Delta) \log 9 \leq (j(p) - j_\Delta) \log 9 + \omega \log 2 \leq S_\Delta$, so $(j', l_\Delta) \in \Delta$ 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 $j_{\Delta'}$ of the $\Delta' \in \mathfrak{T}$ with $S_{\Delta'} \geq S$ and $l_\Delta \leq \text{tip}_{\Delta'} \leq l_\Delta + W$.

Lemma 13.8 (Anchors are separated). *Let $\Delta \in \mathfrak{T}$, and let $W \geq 0$ and $S > 0$ be reals with $A = b_{\text{an}}S - W - 1 > 0$. Distinct elements $\zeta < \zeta'$ of $\text{Anc}(\Delta, S, W)$ satisfy $\zeta' - \zeta > \lambda A$.*

Proof. Let Δ', Δ'' define ζ and ζ' ; they are distinct. As in the proof of [9, Lemma 7.10], we compare their rows at the height $\hat{l} = l_\Delta + \lfloor S/\log 2 \rfloor$. Since $b_{\text{an}} < 1/\log 2$ (Lemma 13.15(4)), $\lfloor S/\log 2 \rfloor > S/\log 2 - 1 > W$, so $\text{tip}_{\Delta'}, \text{tip}_{\Delta''} \leq l_\Delta + W < \hat{l}$ and $l_{\Delta'} = \text{tip}_{\Delta'} + S_{\Delta'}/\log 2 \geq \hat{l}$, and likewise for Δ'' . By Lemma 13.6, $(\zeta', \hat{l}) \in \Delta''$; if $\zeta' \leq \text{row}_{\Delta'}(\hat{l})$, then $(\zeta', \hat{l}) \in \Delta'$ as well, contradicting Proposition 9.13. Hence $\zeta' - \zeta > \lambda(\hat{l} - \text{tip}_{\Delta'}) \geq \lambda(S/\log 2 - 1 - W) > \lambda A$. $\square$

13.4. The paid first passage. In Tao's argument the reward of the first passage above the top edge is discarded. We keep it: the first passage then has total mass less than $\mathbf{q}_{\text{mass}}$ at large gaps (Lemma 13.13), which pays for most of the budget. For $v \in \mathcal{P}$ put $H(v) = Q(v)/w(v)$, and for a black point e with $\omega = \text{gap}(e)$ define the *paid first passage*

$$\tilde{F}_e(x) = \sum_{\mathbf{f} \in \Pi_\omega: \text{pt}_{|\mathbf{f}|}(\mathbf{f})=x} \mathbf{q}(\mathbf{f}) e^{-z \text{Rw}(e, \mathbf{f}; |\mathbf{f}|)}, \quad \tilde{\mathbb{P}}_e(x, \mathbf{b}) = \tilde{F}_e(x) \mathbf{q}(\mathbf{b}).$$

As rewards are nonnegative, $0 \leq \tilde{F}_e \leq F_\omega$ by (12.2), so $\tilde{\mathbb{P}}_e \leq \mathbb{P}_{e, \omega}$.

Lemma 13.9 (Paid continuation). *Let e be black and $1 \leq L \leq J$. Then*

$$Q(e) \leq \sum_a \tilde{\mathbb{P}}_e(a) e^{-z \text{Rw}_{L-1}^e(a)} Q(y_L^e(a)).$$

Proof. By Lemma 11.12(2) and (12.1), $H(v) = \sum_b \mathbf{q}(b) e^{-z \text{rw}(v, b)} H(v + \text{pt}(b))$ for every $v \in \mathcal{P}$, and iterating, $H(v) = \sum_{\mathbf{b} \in \mathfrak{B}^K} \mathbf{q}(\mathbf{b}) e^{-z \text{Rw}(v, \mathbf{b}; K)} H(v_{\mathbf{b}})$ (Lemma 12.2(1), (2)). Apply the one-step identity at e and expand only the lists whose holds all end at height at most $l(e) + \omega$; after $\lfloor \omega/4 \rfloor + 1$ steps none remain (Lemma 12.2(3)), and the expanded terms are exactly the lists of Π_ω . As e is black, $w(e) = 1$, so $Q(e) = \sum_x \tilde{F}_e(x) H(e + x)$. For $o = e + x$, the iterated identity with $K = L$ and (12.1) for the last hold give $H(o) \leq \sum_{\mathbf{b} \in \mathfrak{B}^L} \mathbf{q}(\mathbf{b}) e^{-z \text{Rw}(o, \mathbf{b}; L-1)} Q(o_{\mathbf{b}})$, as $l \leq 1$. Extending $\mathbf{b}$ to J holds changes neither factor (Lemma 12.2(1), (2)). $\square$

13.5. The paid mass of a long first passage. A first passage over a long gap ω first crosses height ω at a raw letter c , raw index r and overshoot o (Lemma 12.5). For $r \leq \lfloor 9\omega/32 \rfloor$ the angles at the crossing point and at the following raw positions of the same hold, up to height 17 above the top edge, are determined by seventeen bits of the column-top residue up to a perturbation below 10^{-6} ; the paid factors there are bounded by cost tables which are periodic in r , and averaging them against the rows of raw-prefix masses gives the bound $\mathbf{q}_{\text{mass}}$.

Lemma 13.10 (Narrow envelopes). *Let U' be the envelope table with $U'_{1,2} = 2^{-16}$ and the other entries*

$s \backslash t$	2	3	4	5	6	7
1		.333340	.500010			
2	.707108	.333337	.353560			
3	.923880	.577352	.461943			
4	.980786	.883638	.604555			
5	.995185	.970239	.867542			
6	.998796	.992518	.965743	.864601		
7					.868067	
8						.874299

Then $\Phi_t((b + \beta)/2^s) \leq U'_{s,t}$ for every (s, t) in U' , every odd b and every real $|\beta| \leq 10^{-6}$. Put $\delta'_{c,o} = \min\{a_{c,o}[U']/(c-1), 1 - 2^{-16}\}$; then $\Phi_c^{\text{cap}}((b + \beta)/2^o) \leq 1 - \delta'_{c,o}$ for $3 \leq c \leq 16$, $1 \leq o \leq c$.

Proof. For $(1, 2)$, $\Phi_2((b + \beta)/2) = |\sin(\pi\beta/2)| \leq \frac{11}{7} \cdot 10^{-6} < 2^{-16}$. For the other pairs let $f = \Phi_t^2$; then $|f'| \leq \frac{4\pi}{t^2} \sum_{i < j < t} (2^j - 2^i) < L_{s,t} = \frac{13}{t^2} \sum_{i < j < t} (2^j - 2^i)$, so $f((b + \beta)/2^s) \leq f(b/2^s) + L_{s,t} 10^{-6}/2^s$. Bounding f at the finitely many centres $b/2^s$, $1 \leq b < 2^s$ odd, by the cosine bounds of Lemma 12.9(2), the 20 exact comparisons with $(U'_{s,t})^2$ hold; the smallest margin is at $(2, 4)$, where $f(\frac{1}{4}) = \frac{1}{8}$ and $\frac{1}{8} + \frac{13 \cdot 23}{16} \cdot \frac{10^{-6}}{4} < 0.353560^2$. The second claim follows from Lemma 12.7, as $\Phi_c^{\text{cap}} = 2^{-16}$ when $\Phi_{c-1} < 2^{-16}$. $\square$

Lemma 13.11 (Rows and residues). *Let $l \geq 2$ and put $R_l(r) = R(r-1, l)$ for $r \geq 1$, so $R_l(1) = 0$, $R_l(r) = \binom{l-1}{2r-3} 2^{-l}$ for $r \geq 2$ and $\sum_r R_l(r) = \frac{1}{4}$.*

- (1) If $\omega \geq 680$ and $\omega - 127 \leq l \leq \omega$, then $\sum_{r > \lfloor 9\omega/32 \rfloor} R_l(r) \leq e^{-2\omega/289}$, which is at most $37/\omega$ for $\omega \geq 1129$.
- (2) $9^{2^{14}} \equiv 1 \pmod{2^{17}}$, and every odd M is $\equiv 9^a s \pmod{2^{17}}$ for some $0 \leq a < 2^{14}$ and $s \in \{1, 3, 5, 7\}$, $s \equiv M \pmod{8}$.
- (3) For $L = 2^{14}$ and $f : \mathbb{N} \rightarrow [0, 1]$ of period L , $\sum_r R_l(r)f(r) \geq \frac{1}{4L} \sum_{a < L} f(a) - 8192e^{-(l-1)/2^{29}}$.
- (4) Let $f : \mathbb{N} \rightarrow [0, 1]$ have period $L \geq 1$, mean μ , and range $D = \max_{k \leq L} Q(k) - \min_{k \leq L} Q(k)$, $Q(k) = \sum_{j < k} (f(j) - \mu)$. Then $|\sum_{k=u}^{u+t-1} f(k) - t\mu| \leq D$ for all $u, t \in \mathbb{N}$, and for $a \in \mathbb{N}$,

$$\sum_{r \geq 1} R_l(r)f(r+a) \geq \frac{\mu}{4} - \frac{D}{\sqrt{2(3l-1)}}.$$

If moreover $N \geq 1$, $d \geq N$, $Z \subseteq \mathbb{Z}$ has pairwise distances at least d and $\mathcal{V} = \{\zeta + i : \zeta \in Z, 0 \leq i < N\}$, then

$$\sum_{r \in \mathcal{V}} R_l(r)f(r+a) \leq (\mu N + D) \left(\left(1 + \frac{N-2}{d}\right) \frac{1}{\sqrt{2(3l-1)}} + \frac{1}{4d} \right).$$

Proof. (1) The omitted part is at most $\frac{1}{2}\mathbb{P}(X \geq 2 \lfloor 9\omega/32 \rfloor - 1)$ for X binomial with parameters $l-1, \frac{1}{2}$; the threshold exceeds the mean by at least $\frac{\omega}{16} - \frac{5}{2} \geq \frac{\omega}{17}$, and Hoeffding's inequality (Lemma 10.11) gives $2e^{-2\omega/289}$. Then $e^y \geq y^2/2$ gives the second bound. (2) By induction, 2^{j+3} exactly divides $9^{2^j} - 1$; so 9 has order 2^{14} modulo 2^{17} , and its powers times s are the 2^{14} residues $\equiv s \pmod{8}$. (3) $r \equiv a \pmod{L}$ if and only if $2r - 3 \equiv 2a - 3 \pmod{2L}$, so the mass of a class is $\frac{1}{2}\mathbb{P}(X \equiv 2a - 3 \pmod{2^{15}})$, and by orthogonality of the characters of $\mathbb{Z}/2^{15}$ and $|\cos(\pi k/m)| \leq \cos(\pi/m) \leq e^{-2/m^2}$ this is within $\frac{1}{2}e^{-(l-1)/2^{29}}$ of $1/(4L)$; sum over the classes. (4) $Q(k+L) = Q(k)$, so the interval sum is $Q(u+t) - Q(u)$. The binomial coefficients are unimodal, so every level set $\{R_l \geq v\}$, $v > 0$, is an interval of integers, and $R_l = \sum_i (v_i - v_{i-1}) \mathbf{1}_{I_i}$ over the distinct values $0 = v_0 < v_1 < \dots$; here $\sum_i (v_i - v_{i-1}) |I_i| = \frac{1}{4}$ and $\sum_i (v_i - v_{i-1}) = \max R_l \leq (2(3l-1))^{-1/2}$ (Lemma 10.11). Apply the interval bound on each I_i . For the second inequality, an interval of k integers meets at most $1 + (k + N - 2)/d$ of the blocks of $\mathcal{V}$, and each intersection carries at most $\mu N + D$. $\square$

For the cost tables, the retained factors along the final hold are bounded as follows. For $h \geq 1$ and real X define, by downward recursion, the *future cost* $\text{cf}_h(X) = 1$ for $h \geq 15$ and, for $h \leq 14$,

$$\text{cf}_h(X) = 1 - \sum_{d=2}^{17-h} q(d) \left[1 - \Phi_d^{\text{cap}} \left(\frac{X}{2^{h+d}} \right) \cdot \begin{cases} \text{cf}_{h+d}(9X) & d \notin \{4, 5\}, h+d \leq 15, \\ 1 & \text{otherwise,} \end{cases} \right],$$

with Lipschitz constant $L_h = 2^{-h} \sum_{d=3}^{17-h} q(d) \ell_d 2^{-d} + 9 \sum_{d \notin \{4, 5\}, d \leq 15-h} q(d) L_{h+d}$ ($L_h = 0$ for $h \geq 15$); put $\text{xc}_{c,o}(X) = \Phi_c^{\text{cap}}(X/2^o)$ for $c \in \{4, 5\}$ and $\text{xc}_{c,o}(X) = \Phi_c^{\text{cap}}(X/2^o) \text{cf}_o(9X)$ otherwise. Let $\text{cf}_h^+(X)$ be defined by the same recursion for integers X with Φ^{cap} replaced by $\overline{U}$ (Lemma 12.9(2), cosine bounds rounded up on the mesh 10^{-18}) and each value rounded up on the mesh 10^{-8} . For $2 \leq c \leq 16$, $1 \leq o \leq c$ and odd m , the *cost table* $\text{ct}_{c,o}(m)$ is the least of: 1; if $c \geq 3$, $1 - \delta'_{c,o}$ and $\overline{U}_c(m/2^o) + \lceil 10^8 \cdot 10^{-6} \ell_c 2^{-o} \rceil 10^{-8}$; and, if $c \notin \{4, 5\}$, $\lceil 10^8 \overline{U}_c(m/2^o) \text{cf}_o^+(9m) \rceil 10^{-8} + \lceil 10^2 (\ell_c 2^{-o} + 9L_o) \rceil 10^{-8}$. All arguments of $\overline{U}$ are multiples of 2^{-17} , so $\text{ct}_{c,o}(m)$ depends only on $m \bmod 2^{17}$. For $s \in \{1, 3, 5, 7\}$ let $\mu_{c,o,s}$ and $\text{rg}_{c,o,s}$ be the mean and the range, in the sense of Lemma 13.11(4), of the sequence $k \mapsto \text{ct}_{c,o}(9^k s)$ of period 2^{14} , and put

$$S_s^h = \sum_{c=2}^{16} q(c) \sum_{o=1}^c (1 - \mu_{c,o,s}), \quad D_s^h = \sum_{c=2}^{16} q(c) \sum_{o=1}^c \text{rg}_{c,o,s}.$$

Lemma 13.12 (Cost of an early crossing).

- (1) Let e be black with $\omega = \text{gap}(e) \geq 224$, let $\hat{e} = (j(e), l_*(e))$ and $N = 3^n \vartheta(\hat{e})$, and let M_0 be odd with $2^{17} \mid 3^n M_0 + N$. For $2 \leq c \leq 16$, $1 \leq o \leq c$ and $1 \leq r \leq \lfloor 9\omega/32 \rfloor$,

$$\sum_{\mathbf{f} \in \mathcal{X}_\omega(c, o, r)} \mathbf{q}(\mathbf{f}) e^{-z \text{Rw}(e, \mathbf{f}; |\mathbf{f}|)} \leq \mathbf{q}(c) \mathbf{R}(r-1, \omega + o - c) \text{ct}_{c, o}(9^r M_0).$$

- (2) The aggregates are

$$(\mathbf{S}_s^h, \mathbf{D}_s^h) = \begin{cases} \left(\frac{149160843578876007}{53687091200000000}, \frac{90751350494755481}{21474836480000000} \right), & s \in \{1, 7\}, \\ \left(\frac{31902235342658719}{107374182400000000}, \frac{258514465939645767}{53687091200000000} \right), & s \in \{3, 5\}. \end{cases}$$

Proof. (1) By Lemma 9.14(c), N is odd and $|\vartheta(\hat{e})| \leq \varepsilon 2^{-\omega}$. Residue transport as in Lemma 12.13(1) gives $\vartheta(j(e) + s', l_*(e) + h) \equiv 9^{s'}(M_0 + \vartheta(\hat{e}))/2^h$ for $s' \geq 0$, $0 \leq h \leq 17$. Put $x = 9^r(M_0 + \vartheta(\hat{e}))$; then $|9^r \vartheta(\hat{e})| \leq 2^{16r/5} \varepsilon 2^{-\omega} < 2^{-\omega/10} < 10^{-6}$. By Lemma 12.5(2), a list of $\mathcal{X}_\omega(c, o, r)$ has raw word $\mathbf{a}cs$. The crossing point $e + (r, \omega + o)$ and the points of the next raw letters $\mathbf{s}_1, \mathbf{s}_2, \dots$ up to height 15 above $l_*(e)$ (at most seven of them, as raw letters are at least 2) lie within raw index $\lfloor 9\omega/32 \rfloor + 7 \leq f_\omega$ of e , hence have $j \leq J$ (Lemma 9.14). Their factors are therefore at most $\Phi_c^{\text{cap}}(x/2^o)$, $\Phi_{\mathbf{s}_1}^{\text{cap}}(9x/2^{o+\mathbf{s}_1})$, $\dots$, the last retained one being at height at most 17. Discarding the other rewards and summing over $\mathbf{s}$ with the hold-word masses gives, by induction on the height, at most $\mathbf{q}(c) \mathbf{R}(r-1, \omega + o - c) \text{xc}_{c, o}(x)$. It remains to show $\text{xc}_{c, o}(m + \beta) \leq \text{ct}_{c, o}(m)$ for $m = 9^r M_0$ and $|\beta| \leq 10^{-6}$. The value 1 and the packet value hold by Lemma 13.10; the direct value by Lemma 12.9(1), (2); and the joint value because $\text{xc}_{c, o}$ is a product of two $[0, 1]$ -valued factors with Lipschitz constants $\ell_c 2^{-o}$ and $9\mathbf{L}_o$ (by induction, as in Lemma 12.13(4)), and $\text{cf}_h \leq \text{cf}_h^+$ (by induction).

(2) This is a finite computation in exact rational arithmetic. The cosine bounds are needed only at the $2^{16} + 1$ turns $j/2^{17}$, $0 \leq j \leq 2^{16}$; the arrays cf_h^+ , $14 \geq h \geq 1$, on the 2^{16} odd residues modulo 2^{17} are computed in decreasing order of h ; then the 135 tables $\text{ct}_{c, o}$ on the odd residues (131050 distinct direct values and 8126464 distinct joint values), and the 540 orbit sequences $\text{ct}_{c, o}(9^k s \bmod 2^{17})$, $k < 2^{14}$, which give $\mu_{c, o, s}$, $\text{rg}_{c, o, s}$ and the four pairs displayed. $\square$

Lemma 13.13 (Paid mass of a long first passage). *Let e be black and $\omega = \text{gap}(e)$. If $\omega \geq 2^{34}$, then $\sum_x \tilde{\mathbf{F}}_e(x) < \mathbf{q}_{\text{mass}} = \frac{38177}{125000}$. If $\omega \geq 4096$, then $\sum_x \tilde{\mathbf{F}}_e(x) < b_*$ (Definition A.2).*

Proof. Choose M_0 as in Lemma 13.12(1) and s, a with $M_0 \equiv 9^a s \pmod{2^{17}}$ (Lemma 13.11(2)), and put $\bar{r} = \lfloor 9\omega/32 \rfloor$. By Lemma 12.5, the classes $\mathcal{X}_\omega(c, o, r)$ partition Π_ω with masses $\mathbf{q}(c) \mathbf{R}(r-1, \omega + o - c)$ of total 1. Bounding the discount by 1 outside $2 \leq c \leq 16$, $r \leq \bar{r}$, and by Lemma 13.12(1) there,

$$\sum_x \tilde{\mathbf{F}}_e(x) \leq 1 - \sum_{c=2}^{16} \mathbf{q}(c) \sum_{o=1}^c \sum_{r=1}^{\bar{r}} \mathbf{R}(r-1, l_{c, o}) (1 - \text{ct}_{c, o}(9^{r+a} s)), \quad l_{c, o} = \omega + o - c.$$

Each row $l = l_{c, o}$ lies in $[\omega - 15, \omega]$. If $\omega \geq 2^{34}$, Lemma 13.11(3), (1) bound each inner sum below by $\frac{1}{4}(1 - \mu_{c, o, s}) - E$ with $E = 8192(\frac{3}{8})^{31} + 37/2^{34}$ (as $e^{-(l-1)/2^{29}} \leq e^{-31} < (\frac{3}{8})^{31}$), so, with $\sum_{c=2}^{16} c \mathbf{q}(c) = \frac{65459}{16384} \leq 4$,

$$\sum_x \tilde{\mathbf{F}}_e(x) \leq 1 - \frac{\mathbf{S}_s^h}{4} + 4E < \frac{38177}{125000}$$

by Lemma 13.12(2) and exact arithmetic. If $\omega \geq 4096$, Lemma 13.11(4) (with $\mathbf{f}(k) = 1 - \text{ct}_{c, o}(9^k s)$, of mean $1 - \mu_{c, o, s}$ and range $\text{rg}_{c, o, s}$) and (1) give instead

$$\sum_x \tilde{\mathbf{F}}_e(x) \leq 1 - \frac{\mathbf{S}_s^h}{4} + \frac{\mathbf{D}_s^h}{\sqrt{2(3\omega - 46)}} + \frac{65459}{16384} e^{-2\omega/289} \leq 1 - \frac{\mathbf{S}_s^h}{4} + \frac{\mathbf{D}_s^h}{156} + \frac{65459}{16384 \cdot 2^{40}},$$

using $2(3\omega - 46) \geq 156^2$ and $e^{28} > 2^{40}$; the right side is $0.33250 \dots$ for $s \in \{1, 7\}$ and $0.28808 \dots$ for $s \in \{3, 5\}$, both below b_* . $\square$

13.6. Deviations, the initial filter and the packing estimate. Outside a small deviation event the first passage and the next holds stay close to their typical displacements, so a visit to a large triangle pins the column of the first passage near an anchor (Lemma 13.17), and the anchors are sparse. The allowances below are affine in the number v of holds; their slopes are chosen so that the generating functions of Lemma 13.15(3) control the deviations along the whole schedule at once.

Definition 13.14 (Case 3 allowances). For each index the allowances are X_ι° , $X_\iota^l(v) = X_\iota^l(0) + s_\iota^v v$, $X_\iota^d(v) = X_\iota^d(0) + s_\iota^d v/4$ and $W_\iota(v) = X_\iota^\circ + X_\iota^l(v)$, with the following constants:

ι	X_ι°	$X_\iota^l(0)$	s_ι^v	y_ι^v	$X_\iota^d(0)$	s_ι^d
o	194	1503	$\frac{6579}{500}$	$\frac{169}{168}$	131	$\frac{257}{2000}$
b	182	1174	$\frac{1653}{125}$	$\frac{144}{143}$	102	$\frac{383}{2500}$

ι	b_ι^d	H_ι^{pk}	y_ι^{cl}	v_ι^{cl}
o	$\frac{51}{50}$	35	$\frac{8191}{8192}$	$\frac{1279}{100}$
b	$\frac{43}{42}$	33	$\frac{255}{256}$	$\frac{25}{2}$

The common constants are $b_{\text{an}} = \frac{12500000000}{8664339757}$, $q_{\text{pack}} = \frac{1527079}{5000000}$ and $\Gamma_{\text{pk}} = \frac{2113}{500}$. For $0 \leq i < r_\iota$ put $A_i = b_{\text{an}} S_i - W_\iota(p_i) - 1$, $B_i = 2X_\iota^d(p_i) + \frac{8}{25}W_\iota(p_i) + 1$, $C_i = \frac{127}{40}(2X_\iota^d(p_i) + 1) + W_\iota(p_i)$ and $r_i = \frac{127}{40}(B_i + H_\iota^{\text{pk}})/A_i$.

Lemma 13.15 (Gates for the allowances).

- (1) If $v < V_\iota$ and $\omega \geq \Omega_\iota$, then $\frac{\omega}{32} + X_\iota^d(v) + \frac{1}{4}X_\iota^l(v) + \lambda W_\iota(v) + 1 \leq \omega \lambda_0$.
- (2) For $0 \leq i < r_\iota$, $A_i > 0$ and $B_i + H_\iota^{\text{pk}} \leq \frac{5}{16}A_i$.
- (3) $V(y_\iota^{\text{cl}}) < (y_\iota^{\text{cl}})^{v_\iota^{\text{cl}}}$, $G(1, y_\iota^v) < (y_\iota^v)^{s_\iota^v}$, and $G((b_\iota^d)^{4\varsigma}, (b_\iota^d)^{-\varsigma}) < (b_\iota^d)^{s_\iota^d}$ for $\varsigma = \pm 1$; and $v_\iota^{\text{cl}} \alpha_\iota > \frac{13}{9}$.
- (4) $b_{\text{an}} < 1/\log 2$.

Proof. (1) The allowances increase with v , $\lambda < \frac{8}{25}$ and $\lambda_0 \geq \frac{1}{16}$, so the left side is at most $\frac{\omega}{32} + 8V_\iota \leq \frac{3\omega}{64}$ provided $X_\iota^d(V) + \frac{1}{4}X_\iota^l(V) + \frac{8}{25}W_\iota(V) + 1 \leq 8V$ at $V = V_\iota$; this affine inequality holds as $V_\iota \geq 2^{18}$, and $\Omega_\iota \geq 512V_\iota$ (Lemma A.4); so the left side is at most $\frac{\omega}{16} \leq \omega \lambda_0$. (2) is an exact evaluation at the r_ι schedule positions. (3) By Lemma 10.8(4) these generating functions are explicit rationals: $V(\frac{8191}{8192}) = \frac{6187886789300271379243896831}{6197558195857188412641546239}$, $V(\frac{255}{256}) = \frac{178454632614069375}{187461831868810367}$, $G(1, \frac{169}{168}) = \frac{9577713826824949}{8858255458075381}$, $G(1, \frac{144}{143}) = \frac{97132784504832}{88581797925983}$, and G at the four drift-free points is $\frac{106892500}{106627249}$, $\frac{393435063}{392435263}$ ($\iota = o$), $\frac{45463572}{45304601}$, $\frac{164732957}{164140421}$ ($\iota = b$). Raising to the denominators of the exponents turns each claim into an exact integer comparison. Finally $\frac{1279}{100} \cdot \frac{113}{1000} - \frac{13}{9} = \frac{743}{900000}$ and $\frac{25}{2} \cdot \frac{3}{25} - \frac{13}{9} = \frac{1}{18}$. (4) $\sum_{j \leq 25} x^j/j! > 2$ at $x = 1/b_{\text{an}}$, so $e^{1/b_{\text{an}}} > 2$. $\square$

For $\omega \in \mathbb{N}$ and $v \leq J$, the *deviation event* $\text{Dev}_{\omega, v}$ is the set of atoms $(x, \mathbf{b})$ with $l(x) - \omega \geq X_\iota^\circ$, or $|j(x) - \omega/4| \geq \omega/32$, or $\sum_{k \leq v} l(\text{pt}(b^k)) \geq X_\iota^l(v)$, or $\max_{1 \leq u \leq v} |\sum_{k \leq u} (4j - l)(\text{pt}(b^k))| \geq 4X_\iota^d(v)$. The *fresh deviation event* Dev^{fr} is the set of atoms for which some $1 \leq u \leq \min(V_\iota - 1, J)$ has $\sum_{k \leq u} l(\text{pt}(b^k)) \geq X_\iota^l(u)$ or $|\sum_{k \leq u} (4j - l)(\text{pt}(b^k))| \geq 4X_\iota^d(u)$. For $p \in \mathcal{P}$ and real S , the *large-triangle event* $\text{Big}_{v, S}^p$ is the set of atoms with $y_u^p(a) \in \Delta$ for some $u \leq v$ and $\Delta \in \mathfrak{T}$ with $S_\Delta \geq S$.

Lemma 13.16 (Mass of the deviation events). *Let $p \in \mathcal{P}$ and $\omega \in \mathbb{N}$. If $\omega \geq 2^{32}$, the atoms with $l(x) - \omega \geq X_\iota^\circ$ or $|j(x) - \omega/4| \geq \omega/32$ have $\mathbb{P}_{p, \omega}$ -mass at most $9(\frac{25}{27})^{X_\iota^\circ} + (2\omega)^{-1/2}$. For every ω , $\mathbb{P}_{p, \omega}(\text{Dev}^{\text{fr}}) \leq (y_\iota^v)^{-X_\iota^l(0)} + 2(b_\iota^d)^{-4X_\iota^d(0)}$.*

Proof. The first-passage overshoot tail (Lemma 10.10, valid for $\omega \geq 64$) bounds the first condition by $9(\frac{25}{27})^{X_\iota^\circ}$. The horizontal tail of the first passage (Lemma 10.14, with $u = \omega/32 \geq 2$) bounds the second by $4e^{-(\omega/32)^2/(2(\omega-4))} + 2e^{-\omega/8192} \leq 6e^{-\omega/8192}$, and with $x = \omega/8192 \geq 2^{19}$, $e^{2x} \geq 2x^2 \geq 2^{20}x > 72 \cdot 8192x$, that is, $36e^{-2x} < (2\omega)^{-1}$.

For the second claim we use an *affine barrier*: for reals $X, y > 0, T > 0, \Lambda \geq G(X, y)$ and $N \in \mathbb{N}$, the $\eta^{\otimes N}$ -mass of the $\mathbf{h} \in \mathcal{P}^N$ with $X^{j(\mathbf{h}_1 + \dots + \mathbf{h}_u)} y^{l(\mathbf{h}_1 + \dots + \mathbf{h}_u)} \geq T\Lambda^u$ for some $u \leq N$ is at most $1/T$. Indeed, split by the least such u ; on that event the weight $X^j y^l / (TG^u)$ of the first u points is at least 1, and the tilted measures $\eta^{\otimes u}(\mathbf{h}) X^j y^l / G(X, y)^u$ of these disjoint events, extended to N points, sum to at most 1. By (13.1) the event depends only on the first $\min(V_l - 1, J)$ hold points. For the vertical part, $l \geq X_l^1(u)$ if and only if $(y^v)^l \geq (y^v)^{X^1(0)} ((y^v)^{s^v})^u$; apply the barrier with $X = 1, T = (y^v)^{X^1(0)}$ and Lemma 13.15(3). For the drift-free part, $\varsigma(4j - l) \geq 4X_l^d(u)$ if and only if $((b^d)^{4\varsigma})^j ((b^d)^{-\varsigma})^l \geq (b^d)^{4X^d(0)} ((b^d)^{s^d})^u$; apply the barrier for $\varsigma = \pm 1$. $\square$

For $p \in \Delta \in \mathfrak{T}$, $v \in \mathbb{N}$ and real S , the *anchor columns* are

$$\mathcal{K}_{v,S} = \{\zeta - j(p) + i : \zeta \in \text{Anc}(\Delta, S, W_\ell(v)), -X_\ell^d(v) - \frac{1}{4}X_\ell^1(v) < i < \lambda X_\ell^o + \lambda_0 X_\ell^1(v) + X_\ell^d(v)\}.$$

Lemma 13.17 (Outside the deviation event, large triangles sit at anchors). *Let $p \in \Delta \in \mathfrak{T}$, $\omega = \text{gap}(p)$, $v \leq J$ and S real, and assume $\frac{\omega}{32} + X_\ell^d(v) + \frac{1}{4}X_\ell^1(v) + \lambda W_\ell(v) + 1 \leq \omega\lambda_0$. If an atom $a = (x, \mathbf{b})$ of positive $\mathbb{P}_{p,\omega}$ -mass lies in $\text{Big}_{v,S}^p$ but not in $\text{Dev}_{\omega,v}$, then $j(x) \in \mathcal{K}_{v,S}$.*

Proof. We follow [9, (7.63)–(7.66)]. Choose $u \leq v$ and Δ' with $S_{\Delta'} \geq S$ and $(j, l) = y_u^p(a) \in \Delta'$; put $j_u = \sum_{k \leq u} j(\text{pt}(b^k))$ and $l_u = \sum_{k \leq u} l(\text{pt}(b^k))$. Outside $\text{Dev}_{\omega,v}$, $0 < l(x) - \omega < X^o$, $0 \leq l_u < X^1(v)$, $|j(x) - \omega/4| < \omega/32$ and $|j_u - l_u/4| < X^d(v)$. Hence $0 < l - l_\Delta < W(v)$ and $|j - (j(p) + \omega/4)| < Y = \frac{\omega}{32} + X^d(v) + \frac{1}{4}X^1(v)$, so $\Delta' \neq \Delta$, and Lemma 13.7 gives $\text{tip}_{\Delta'} \geq l_\Delta$; also $\text{tip}_{\Delta'} \leq l < l_\Delta + W(v)$, so $\zeta = j_{\Delta'}$ is an anchor. Lemma 13.6 gives $0 \leq j - \zeta \leq \lambda((l(x) - \omega) + l_u)$. With $i = j(x) + j(p) - \zeta = (j - \zeta) - j_u$, $i > -X^d(v) - \frac{1}{4}X^1(v)$ and $i < \lambda X^o + \lambda_0 X^1(v) + X^d(v)$, as $\lambda = \frac{1}{4} + \lambda_0$. $\square$

Packing needs a filter on the first passage: its crossing letter is at most 128, its crossing index at most $\lfloor 9\omega/32 \rfloor$, and its final hold has at most H_ℓ^{pk} raw letters after the crossing. Let $\mathcal{B}_\omega^{\text{in}}$ be the set of $\mathbf{f} \in \Pi_\omega$ violating this, and for a black e let $\mathbf{F}_e^{\text{gd}}$ be $\tilde{\mathbf{F}}_e$ with the sum restricted to $\Pi_\omega \setminus \mathcal{B}_\omega^{\text{in}}$.

Lemma 13.18 (The filtered first passage). *Let $\omega \geq 12500000$.*

- (1) *The lists of $\mathcal{B}_\omega^{\text{in}}$ have total mass at most $(\frac{11}{16})^{H_\ell^{\text{pk}}} + \frac{2161}{2^{127}} + \frac{148}{\omega}$.*
- (2) *Let e be black with $\text{gap}(e) = \omega \geq 2^{34}$, $\mathbf{n} \in \mathbb{N}$, $\mathbf{B} > \mathbf{n}$ and $\mathbf{d} \geq \mathbf{B} + H_\ell^{\text{pk}}$ real, $Z \subseteq \mathbb{Z}$ with pairwise distances at least $\mathbf{d}$, and $\mathcal{U} = \{\zeta + i : \zeta \in Z, 0 \leq i < \mathbf{n}\}$. Then, with $H = H_\ell^{\text{pk}}$,*

$$\sum_{r \in \mathcal{U}} \sum_{\ell} \mathbf{F}_e^{\text{gd}}(r, \ell) \leq \frac{\mathbf{q}_{\text{pack}}(\mathbf{B} + H) + \Gamma_{\text{pk}}/4}{\mathbf{d}} + \frac{49}{120} \cdot \frac{(4\mathbf{q}_{\text{pack}}(\mathbf{B} + H) + \Gamma_{\text{pk}})(1 + (\mathbf{B} + H)/\mathbf{d})}{\sqrt{\omega}}.$$

Proof. (1) By Lemma 12.5(2), the long final holds have mass at most $(\frac{11}{16})^H$. For $c > 128$, $\sum_{o \leq c} \sum_r \mathbf{R}(r - 1, \omega + o - c) \leq 1 + \frac{c}{4}$, and $\sum_{c > 128} \mathbf{q}(c)(1 + \frac{c}{4}) = \frac{128^2 + 7 \cdot 128 + 8}{4 \cdot 2^{128}} = \frac{2161}{2^{127}}$, by the closed forms $\sum_{c > k} \mathbf{q}(c) = (k + 1)2^{-k}$ and $\sum_{c > k} c \mathbf{q}(c) = (k^2 + 3k + 4)2^{-k}$. For $c \leq 128$ the rows lie in $[\omega - 127, \omega]$, and Lemma 13.11(1) and $\sum_c c \mathbf{q}(c) = 4$ give $148/\omega$.

(2) We may assume $\mathbf{n} \geq 1$. A list counted on the left lies in a class $\mathcal{X}_\omega(c, o, r)$ with $c \leq 128$, $r \leq \lfloor 9\omega/32 \rfloor$ and raw length in $[r, r + H]$; its raw length is its column, so r lies in $\mathcal{V} = \{\zeta' + i : \zeta' \in Z - H, 0 \leq i < N\}$, $N = \mathbf{n} + H < \mathbf{B} + H \leq \mathbf{d}$. With M_0, s, a as in the proof of Lemma 13.13, Lemma 13.12(1) for $c \leq 16$ and the trivial bound for $17 \leq c \leq 128$ reduce the left side to rows, and Lemma 13.11(4) (with $\mathbf{f} \equiv 1$ for $c \geq 17$) bounds the row of (c, o) by $\mathbf{q}(c)\varrho(\mu_{c,o,s}N + \text{rg}_{c,o,s})$, where $\varrho = \frac{49}{120\sqrt{\omega}}(1 + \frac{N-2}{\mathbf{d}}) + \frac{1}{4\mathbf{d}}$, since $(2(3l - 1))^{-1/2} \leq \frac{49}{120\sqrt{\omega}}$ for $l \geq \omega - 127$. Summing, the left side is at most $\varrho((4 - S_s^h)N + D_s^h)$. By Lemma 13.12(2), $S_s^h - 4 + 4\mathbf{q}_{\text{pack}} > 0$ and $D_s^h - (H + 1)(S_s^h - 4 + 4\mathbf{q}_{\text{pack}}) < \Gamma_{\text{pk}}$ for both values of H (exact comparisons), so for $N \geq H + 1$, $(4 - S_s^h)N + D_s^h \leq 4\mathbf{q}_{\text{pack}}N + \Gamma_{\text{pk}} \leq 4\mathbf{q}_{\text{pack}}(\mathbf{B} + H) + \Gamma_{\text{pk}}$. Expanding ϱ gives the claim. $\square$

Lemma 13.19 (Filtered mass of the anchor columns). *Let $p \in \Delta \in \mathfrak{T}$ be black with $\omega = \text{gap}(p) \geq 2^{34}$, $0 \leq i < r_i$, $v = p_i$ and $\mathbf{S} = S_i$. Then*

$$\sum_{r \in \mathcal{K}_{v,S}} \sum_{\ell} F_p^{\text{gd}}(r, \ell) \leq \frac{\mathbf{q}_{\text{pack}} \mathbf{C}_i + \frac{127}{40} (H_i^{\text{pk}} \mathbf{q}_{\text{pack}} + \Gamma_{\text{pk}}/4)}{\mathbf{A}_i} + \frac{49}{120} \cdot \frac{(4\mathbf{q}_{\text{pack}}(\mathbf{B}_i + H_i^{\text{pk}}) + \Gamma_{\text{pk}})(1 + r_i)}{\sqrt{\omega}}.$$

Proof. The range of i in the definition of $\mathcal{K}_{v,S}$ is an open interval of length $2X^{\text{d}}(v) + \lambda W(v)$, so $\mathcal{K}_{v,S}$ is a union of blocks of $n < \mathbf{B} := 2X^{\text{d}}(v) + \lambda W(v) + 1 \leq \mathbf{B}_i$ consecutive integers starting at points of pairwise distance greater than $\mathbf{d} = \lambda \mathbf{A}_i$ (Lemma 13.8, with $\mathbf{A}_i > 0$ by Lemma 13.15(2)). As $\lambda \geq \frac{5}{16}$, $\mathbf{B} + H \leq \mathbf{B}_i + H \leq \frac{5}{16} \mathbf{A}_i \leq \mathbf{d}$. Apply Lemma 13.18(2), and use $(\mathbf{B} + H)/\lambda \leq \mathbf{C}_i + \frac{127}{40} H$ and $(\mathbf{B} + H)/(\lambda \mathbf{A}_i) \leq r_i$, as $1/\lambda < \frac{127}{40}$. $\square$

13.7. Stopped stages for the bootstrap index. For $\iota = \mathbf{b}$ the trace factors of the last two schedule intervals are replaced by the mass of *stopped stages*. From a black point, a stage follows complete entries until either one departure from a point of gap at least 4096 has completed its entry, or 5 departures from smaller gaps have. An entry from a gap of at least 4096 has paid mass less than b_* (Lemma 13.13, via (12.4)), every entry has paid mass at most $\frac{77}{100}$, and $(\frac{77}{100})^5 < b_*$, so a stage has mass at most b_* .

Definition 13.20 (Stopped stages). For $k \in \mathbb{N}$ and $1 \leq t \leq 5$ let $\text{nx}(k, t; \omega) = (k - 1, 5)$ if $\omega \geq 4096$ or $t = 1$, and $\text{nx}(k, t; \omega) = (k, t - 1)$ otherwise. For black v define $\mathcal{M}_{0,t}(v) = 1$ and, for $k \geq 1$, $\mathcal{M}_{k,t}(v) = \sum_{\mathbf{w} \in \mathcal{E}(v)} \mathbf{q}^v(\mathbf{w}) \mathcal{M}_{\text{nx}(k,t;\text{gap}(v))}(v_{\mathbf{w}})$. For a complete path with trace times $u_1 < \dots < u_{\nu}$ define the *stage end* $\text{se}_{0,t}(i) = i$ and, for $k \geq 1$, $\text{se}_{k,t}(i) = \infty$ if $i > \nu$ and $\text{se}_{k,t}(i) = \text{se}_{\text{nx}(k,t;\text{gap}(y_{u_i}))}(i + 1)$ otherwise; and put $\mathcal{Z}_k^{\text{st}}(\bar{o}, \mathbf{b}) = e^{-z \text{Rw}(\bar{o}, \mathbf{b}; u_{\text{se}_{k,5}(1)})}$ if $\text{se}_{k,5}(1) \leq \nu$, and 0 otherwise.

Lemma 13.21 (Mass of stopped stages).

- (1) For black v and $k \in \mathbb{N}$, $\mathcal{M}_{k,5}(v) \leq b_*^k$.
- (2) For $\bar{o} \in \mathcal{P}$ and L with $j(\bar{o}) + L > J$, $\sum_{\mathbf{b} \in \mathfrak{B}^L} \mathbf{q}(\mathbf{b}) \mathcal{Z}_k^{\text{st}}(\bar{o}, \mathbf{b}) \leq b_*^k$.
- (3) In the setting of Lemma 13.4, if every atom of $\mathcal{S}$ satisfies $\text{se}_{k,5}(1) \leq \nu$ and $u_{\text{se}_{k,5}(1)} \leq P_t - 1$ for its fresh path, the factor $(\frac{77}{100})^k$ may be replaced by b_*^k .

Proof. (1) We show $\mathcal{M}_{k,t} \leq \max(b_*, (\frac{77}{100})^t) b_*^{k-1}$ for $k \geq 1$. If $\text{gap}(v) \geq 4096$, then $\mathbf{K1}(v) \leq S(v) = \sum_x \tilde{F}_v(x) < b_*$ by (12.4) and Lemma 13.13; otherwise $\mathbf{K1}(v) \leq \frac{77}{100}$ (Proposition 12.15). Induction on (k, t) in the lexicographic order gives the claim, and $(\frac{77}{100})^5 = \frac{2706784157}{10^{10}} < b_*$. (2) As in Lemma 12.16: the stage end indices shift with the trace times under the restart of Lemma 12.4, so from a black origin the sum equals $\mathcal{M}_{k,5}(v)$ (induction on (k, t)), and from any origin it is at most $\text{Br}(\bar{o}) \max_v \mathcal{M}_{k,5}(v)$. (3) follows from (2) as Lemma 13.4 follows from Lemma 12.16. $\square$

Lemma 13.22 (Stages are completed on schedule). *Let $\iota = \mathbf{b}$ and $h^{\text{st}} = 1024 + T_{\mathbf{b}} + 1 = 1193$, and put $\mathbf{t}_0 = p_0 = 169$ and $\mathbf{t}_{s+1} = \max(\mathbf{t}_s + 5h^{\text{st}}, \mathbf{g}_{\mathbf{b}}(\mathbf{t}_s + 4h^{\text{st}}))$; then $(\mathbf{t}_0, \mathbf{t}_1, \mathbf{t}_2) = (169, 503672, 63696221)$. Let p, m, a be as in Lemma 13.3 with $a \notin \text{Clr}$, let $k \in \{1, 2\}$, and suppose that a has no oversize hit at a time $u \leq p_k$. Then $\text{se}_{k,5}(1) \leq \nu$ and $u_{\text{se}_{k,5}(1)} \leq \mathbf{t}_k \leq p_{k+1}$.*

Proof. The values follow from $p_0 = 169$, $p_1 = 15486$, $p_2 = 514217$, $p_3 = 63701994$ (Lemma A.4), $\mathbf{g}_{\mathbf{b}}(4941) = 4941 + 498562 + 169$ and $\mathbf{g}_{\mathbf{b}}(508444) = 508444 + 63187608 + 169$. If u_i is a hit of gap less than 4096 with $u_i + h^{\text{st}} \leq P_{\mathbf{b}} - 1$, then a next trace time exists with $u_{i+1} \leq u_i + h^{\text{st}}$: after 1024 holds the path lies at least 4096 above y_{u_i} , hence above $l_*(y_{u_i})$ (Lemma 12.2(3)), and Lemma 13.3(1) gives a hit in the next $T_{\mathbf{b}} + 1$ times. If the gap is at least 4096, $u_i \leq p_k$ and $\mathbf{g}_{\mathbf{b}}(u_i) \leq P_{\mathbf{b}} - 1$, Lemma 13.3(2) gives $u_{i+1} \leq \mathbf{g}_{\mathbf{b}}(u_i)$. By induction on $s \leq k$, the s th stage ends at a trace time at most $\mathbf{t}_s$: within a stage at most four small-gap steps occur before the last one, each adding at most h^{st} , and the last step adds at most h^{st} or ends by $\mathbf{g}_{\mathbf{b}}$ of a time at most $\mathbf{t}_s + 4h^{\text{st}}$. The departures of the s th stage occur

by time $t_{s-1} + 4h^{\text{st}}$, which is at most p_k because $t_0 + 4h^{\text{st}} = 4941 \leq p_1$ and $t_1 + 4h^{\text{st}} = 508444 \leq p_2$; and $t_1 = 503672 \leq p_2$, $t_2 = 63696221 \leq p_3 = P_b - 1$. $\square$

13.8. The assembly.

Definition 13.23 (Case 3 error). Put $\mathbf{e}_i = \mathbf{e}_i(\iota)$ and $H = H_\iota^{\text{pk}}$, and

$$\begin{aligned} \Pi_\iota^{\text{ad}} &= \sum_{i < r_\iota} \mathbf{e}_i \frac{\mathbf{q}_{\text{pack}} \mathbf{C}_i + \frac{127}{40} (H \mathbf{q}_{\text{pack}} + \Gamma_{\text{pk}}/4)}{\mathbf{A}_i}, \\ E_\iota^{\text{sq}} &= \Omega_\iota^{-1/2} \left(\frac{49}{120} \sum_{i < r_\iota} \mathbf{e}_i (4 \mathbf{q}_{\text{pack}} (\mathbf{B}_i + H) + \Gamma_{\text{pk}}) (1 + r_i) + 1 \right), \\ E_\iota^{\text{in}} &= 9 \left(\frac{25}{27} \right)^{X_\iota^\circ} + \left(\frac{11}{16} \right)^H + \frac{2161}{2^{127}} + \frac{148}{\Omega_\iota}, \quad E_\iota^{\text{dv}} = (y_\iota^\vee)^{-X_\iota^{\text{l}}(0)} + 2(b_\iota^{\text{d}})^{-4X_\iota^{\text{d}}(0)}, \\ E_\iota^{\text{cl}} &= \sum_{i=0}^{r_\iota} \frac{n_i^{\text{cl}}}{P_{80}(t_i^{\text{cl}})}, \end{aligned}$$

where $P_{80}(t) = \sum_{j \leq 80} t^j / j!$, $t_i^{\text{cl}} = (1 - y_\iota^{\text{cl}})(v_\iota^{\text{cl}} h_\iota(p_i) - \frac{13}{9} S_i) > 0$ and $n_i^{\text{cl}} = \max(0, \min(p_i, P_\iota - 1 - h_\iota(p_i)) - p_{i-1})$ ($p_{-1} = -1$). The *Case 3 error* is

$$\mathcal{E}_\iota = \Pi_\iota^{\text{ad}} + \mathbf{q}_{\text{mass}} \mathbf{e}_{r_\iota} + E_\iota^{\text{sq}} + E_\iota^{\text{in}} + \mathbf{q}_{\text{mass}} (E_\iota^{\text{dv}} + E_\iota^{\text{cl}}).$$

Lemma 13.24 (Case 3 low-reward bound). *Let $p \in \Delta \in \mathfrak{T}$, $\omega = \text{gap}(p) > \mathbf{s}_\iota m$, $m \geq D_{\text{sc}}(\iota)$, $j(p) + m = J$ and $P_\iota \leq J$. Then*

$$\sum_{a \in \text{Low}^P \setminus \text{Bad}_m^P} \tilde{\mathbb{P}}_p(a) e^{-z \text{Rw}_{P_\iota-1}^p(a)} \leq \mathcal{E}_\iota.$$

Proof. By Lemma A.6(2), $\mathbf{s}_\iota m \geq \Omega_\iota$, so $\omega \geq \Omega_\iota \geq \max(2^{34}, 512V_\iota)$ (Lemma A.4). Rewards are nonnegative, so all discounts lie in $(0, 1]$. Write X for the left side.

Filter. The part of X carried by $\tilde{\mathbb{F}}_p - \mathbb{F}_p^{\text{gd}}$ is at most $(\frac{11}{16})^H + \frac{2161}{2^{127}} + \frac{148}{\Omega_\iota}$ (Lemma 13.18(1)). For the rest, $\sum_x \mathbb{F}_p^{\text{gd}}(x) < \mathbf{q}_{\text{mass}}$ (Lemma 13.13) and $\mathbb{F}_p^{\text{gd}} \leq \mathbb{F}_\omega$.

Deviations. The atoms with $l(x) - \omega \geq X^\circ$ or $|j(x) - \omega/4| \geq \omega/32$ contribute at most $9(\frac{25}{27})^{X^\circ} + (2\omega)^{-1/2}$ (Lemma 13.16). Membership in $\text{Dev}^{\text{fr}} \cup \text{Clr}$ depends only on $\mathbf{b}$, so these atoms contribute at most $\mathbf{q}_{\text{mass}}(E_\iota^{\text{dv}} + E_\iota^{\text{cl}})$ (Lemmas 13.16 and 13.5).

The remaining atoms. Let $\mathcal{L}'$ be the remaining atoms of positive mass. Split them into $\mathcal{L}'_A$, those with no oversize hit at a time $u \leq p_{r_\iota-1}$, and, for $i < r_\iota$, $\mathcal{L}'_i$, those whose first oversize hit u_0 has $p_{i-1} < u_0 \leq p_i$. On $\mathcal{L}'_A$, Lemma 13.3(3) with $k = r_\iota$ and Lemma 13.4 give the contribution at most $\mathbf{q}_{\text{mass}}(\frac{77}{100})^{r_\iota} = \mathbf{q}_{\text{mass}} \mathbf{e}_{r_\iota}$ if $\iota = \circ$; if $\iota = \mathbf{b}$, Lemmas 13.22 (with $k = 2$) and 13.21(3) give $\mathbf{q}_{\text{mass}} b_*^2 = \mathbf{q}_{\text{mass}} \mathbf{e}_3$. An atom of $\mathcal{L}'_i$ lies in Big_{p_i, S_i}^p and outside Dev_{ω, p_i} (as $p_i < V_\iota$ and the allowances increase), and Lemma 13.15(1) gives the margin hypothesis of Lemma 13.17; so $j(x) \in \mathcal{K}_{p_i, S_i}$. Its fresh path has $i+1$ trace times by time p_i (Lemma 13.3(3)), and, for $\iota = \mathbf{b}$ and $i = 2$, one stage by time p_2 (Lemma 13.22, $k = 1$). So Lemma 13.4 or Lemma 13.21(3), with $D = \mathcal{K}_{p_i, S_i} \times \mathbb{Z}$, and Lemma 13.19 bound the contribution of $\mathcal{L}'_i$ by $\mathbf{e}_i$ times the bound of Lemma 13.19. Summing, and using $(2\omega)^{-1/2} \leq \Omega_\iota^{-1/2}$ and $\omega^{-1/2} \leq \Omega_\iota^{-1/2}$, gives $X \leq \mathcal{E}_\iota$. $\square$

Lemma 13.25 (Outer horizontal tail). *Let $p \in \mathcal{P}$, $\omega \in \mathbb{N}$, and let $m \geq D_{\text{bad}}(\iota)$ be an integer with $P_\iota \leq J$ and $\omega \log 2 \leq m \log 9$. Then $m^{A_\iota} \mathbb{P}_{p, \omega}(\text{Bad}_m^p) < \frac{1}{8000}$.*

Proof. As in [9, proof of (7.54)], we split the advance between the first passage and the next P_ι holds. Bad_m^p is contained in $\{j(x) \geq \frac{158499}{200000} m\} \cup \{j(\text{pt}_{P_\iota}(\mathbf{b})) \geq \frac{m}{200000}\}$, as $\frac{317}{400} = \frac{158500}{200000}$. Since $\omega \leq m \log 9 / \log 2 < \frac{158497}{50000} m$, on the first event $j(x) - \omega/4 \geq u = m/100000 > 2$, and Lemma 10.14 bounds its mass by $4e^{-u^2/(2(\omega-4))} + 2e^{-u/256} \leq 6e^{-m/2^{36}}$ (or by $4e^{-u/256}$ if $\omega \leq 5$), because

$25600000 < 2^{36}$ and $2(\omega - 4) \cdot 10^{10} < 63398800000m \leq 2^{36}m$. On the second, Lemma 10.9(3), applied through (13.1) with threshold $10 \lceil m/200000 \rceil$, gives $e^{P_\iota/2 - m/3200000}$. Now $m \geq D_{\text{bad}}(\iota) = 2^{21}(P_\iota + 64)$, so $P_\iota/2 \leq m/2^{22} - 32$, and $m \geq 2^{43}$ by Lemma A.6(1). With $x = m/2^{43} \geq 1$, $\log m < 43 + x - 1$, so $A_\iota \log m \leq \frac{5}{2} \log m < 105 + \frac{5}{2}x \leq 128x - 13 = m/2^{36} - 13$; hence $m^{A_\iota} 6e^{-m/2^{36}} < 6e^{-13} < e^{-10}$, and $m^{A_\iota} e^{P_\iota/2 - m/3200000} < e^{m/2^{36} + m/2^{22} - 32 - m/3200000} < e^{-32}$, as $2^{-36} + 2^{-22} < \frac{1}{3200000}$. Finally $e^{10} > 2e^9 > 16000$ (Lemma A.1(4)), so the sum is less than $2e^{-10} < \frac{1}{8000}$. $\square$

Lemma 13.26 (Case 3 of the descent step). *Let m be an integer with $D_*(\iota) \leq m \leq J$, let $p \in \mathcal{P}$ with $j(p) + m = J$, and suppose $p \in \Delta \in \mathfrak{T}$ with $\text{gap}(p) > s_\iota m$. Then $\mathbf{Q}(p) \leq m^{-A_\iota} \mathbf{Q}_{m-1}$.*

Proof. We follow [9, (7.52)–(7.56)], with $\frac{317}{400}m$ in place of $0.9m$ and an explicit error budget. Write $A = A_\iota$ and $\omega = \text{gap}(p)$. By Lemma A.6(2), $P_\iota \leq m \leq J$. 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 13.9 with $L = P_\iota$ gives $\mathbf{Q}(p) \leq \sum_a \tilde{\mathbb{P}}_p(a) e^{-z \text{Rw}_{P_\iota-1}(a)} \mathbf{Q}(y_{P_\iota}^p(a))$. For an atom of positive mass, $\text{Adv}(a) = j(y_{P_\iota}^p(a)) - j(p) \geq 1$ and $j(y_{P_\iota}^p(a)) + m = J + \text{Adv}(a)$, so the boundary reduction (Lemma 11.16(1)) gives $\mathbf{Q}(y_{P_\iota}^p(a)) \leq \max(m - \text{Adv}(a), 1)^{-A} \mathbf{Q}_{m-1}$. This factor is at most $m^A \cdot m^{-A}$; outside Bad_m^p , $m - \text{Adv}(a) > \frac{83}{400}m$, so it is less than $(400/83)^A m^{-A} < c_{\text{cont}}(\iota) m^{-A}$ (Lemma A.9); outside Low^p as well, $e^{-z \text{Rw}} (400/83)^A < e^{-z T_\iota} (400/83)^A < 1/P_W(\iota)$ (Lemma A.9). Let t , b and l be the $\tilde{\mathbb{P}}_p$ -masses of all atoms, of Bad_m^p and of $\text{Low}^p \setminus \text{Bad}_m^p$, and X the left side of Lemma 13.24; so $t < \mathbf{q}_{\text{mass}}$ (Lemma 13.13) and $X \leq l$. With $c = c_{\text{cont}}(\iota)$ and $w = 1/P_W(\iota) \in (0, 1)$,

$$\mathbf{Q}(p) \leq m^{-A} \mathbf{Q}_{m-1} (m^A b + cX + w(t - b - l)) \leq m^{-A} \mathbf{Q}_{m-1} (m^A b + \mathbf{q}_{\text{mass}} w + (c - w)X).$$

Since $\tilde{\mathbb{P}}_p \leq \mathbb{P}_{p,\omega}$, Lemma 13.25 gives $m^A b < \frac{1}{8000}$, and Lemma 13.24 gives $X \leq \mathcal{E}_\iota$. So the bracket is less than

$$\frac{1}{8000} + \frac{\mathbf{q}_{\text{mass}}}{P_W(\iota)} + \left(c_{\text{cont}}(\iota) - \frac{1}{P_W(\iota)} \right) \mathcal{E}_\iota,$$

which is at most the error budget (A.2), less than 1 by Lemma A.9(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) for each index $\iota \in \{\text{o}, \text{b}\}$: the boundary point p for m is white, or black with $\text{gap}(p) \leq s_\iota m$. With Case 3 (Lemma 13.26) this gives the descent step beyond the threshold $D_*(\iota)$, from which (4.2) follows. As in Section 13, $\mathbf{Q}_m$ is formed with an exponent A stated in each lemma, and $\chi = \frac{63}{2500}$ and z are the constants of Appendix A.1.

14.1. Envelope factors. To average the bound of Lemma 11.16(1) over the hold law, we dominate $\max(m - r, 1)^{-A}$ by m^{-A} times the *envelope factor* $e^{\mathbf{t}_m r}$ (compare [9, (7.42)]), whose average is an exponential moment; here, for an integer $m \geq 1$ and $A \geq 0$,

$$\mathbf{t}_m = \frac{4A(1 + \log m)}{m}. \quad (14.1)$$

Lemma 14.1 (Envelope bound). *Let $0 \leq A \leq \frac{5}{2}$, and let $m \geq 2$ and $r \geq 1$ be integers. Then*

$$\max(m - r, 1)^{-A} \leq m^{-A} e^{\mathbf{t}_m r}, \quad (14.2)$$

and $(m - r)^{-A} \leq m^{-A} e^{2Ar/m}$ if $2r \leq m$. If $m \geq 2^{40}$, then $\mathbf{t}_m \leq \chi/32 \leq 1/1024$ and

$$\sum_{r \geq 1} \nu_{45}(r) \max(m - r, 1)^{-A} \leq \mathbf{M}_{45}(\mathbf{t}_m) m^{-A} \leq e^{\chi/2} m^{-A}. \quad (14.3)$$

Proof. If $2r \leq m$, then $(1 - r/m)(1 + 2r/m) \geq 1$ gives $(m - r)^{-A} \leq m^{-A} e^{2Ar/m}$, and $2A/m \leq \mathbf{t}_m$. If $2r > m$, then $\max(m - r, 1)^{-A} \leq 1 = m^{-A} e^{A \log m}$ and $A \log m \leq \mathbf{t}_m r$. If $m \geq 2^{40}$, then $\log m \leq \sqrt{m}$ gives $\mathbf{t}_m \leq 8A/\sqrt{m} \leq \chi/32$, because $(256A/\chi)^2 \leq (256 \cdot \frac{5}{2} \cdot \frac{2500}{63})^2 < 2^{30}$; and $\chi/32 < 1/1024$. Summing (14.2) against ν_{45} gives the first inequality of (14.3), and Lemma 10.8(2) gives $\mathbf{M}_{45}(\mathbf{t}_m) \leq e^{13\mathbf{t}_m/4} \leq e^{\chi/2}$. $\square$

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 $0 \leq A \leq \frac{5}{2}$, let $m \geq 2^{40}$ be an integer, and let $p \in \mathcal{P}$ be white with $j(p) + m = J$. Then*

$$\mathbf{Q}(p) \leq e^{-z+\chi/2} m^{-A} \mathbf{Q}_{m-1} \leq m^{-A} \mathbf{Q}_{m-1}.$$

Proof. We follow [9, (7.43)]. The second inequality holds as $\chi/2 = \frac{63}{5000} < z$. By Lemma 11.12(3) and $\mathbf{w}(p) = e^{-z}$, $\mathbf{Q}(p) \leq e^{-z} \sum_y \eta(y) \mathbf{Q}(p+y)$. Every $y \in \mathcal{P}$ has $j(y) \geq 1$ and $j(p+y) + m = J + j(y)$, so Lemma 11.16(1) gives $\mathbf{Q}(p+y) \leq \max(m - j(y), 1)^{-A} \mathbf{Q}_{m-1}$. Grouping by $r = j(y)$ with (10.2) and using (14.3) gives the claim. $\square$

14.3. The near-top case. Tao's Case 2 halts the hold walk at its first passage over the top edge and uses the reward of the exit points. We instead continue $\mathbf{Q}$ through the paid first passage and one further hold. The first passage then has mass less than $\mathbf{q}_{\text{mass}}$ at large gaps and at most $1 - d_{\text{pass}}$ at every gap, while its horizontal advance is about $\omega/4 \leq \mathbf{s}_l m/4$; so the loss $(m/(m - \text{advance}))^A$ of the boundary reduction is paid by the mass.

Lemma 14.3 (Black boundary points near the top). *Let $\iota \in \{\mathbf{o}, \mathbf{b}\}$, let $0 \leq A \leq A_l$, let $m \geq 2^{42}$ be an integer, and let p be black with $j(p) + m = J$ and $\omega = \text{gap}(p) \leq \mathbf{s}_l m$. Then*

$$\mathbf{Q}(p) \leq \frac{1999}{2000} m^{-A} \mathbf{Q}_{m-1}.$$

Proof. Lemma 13.9 with $L = 1$ gives $\mathbf{Q}(p) \leq \sum_a \tilde{\mathbb{P}}_p(a) \mathbf{Q}(y_1^p(a))$. For an atom $a = (x, \mathbf{b})$ of positive mass put $R = j(x) \geq 1$ and $U = j(\text{pt}(b^1)) \geq 1$; then $j(y_1^p(a)) + m = J + R + U$, and Lemma 11.16(1) gives $\mathbf{Q}(p) \leq m^{-A} \mathbf{Q}_{m-1} Z$ with

$$Z = \sum_a \tilde{\mathbb{P}}_p(a) \theta(a)^A, \quad \theta(a) = \frac{m}{\max(m - R - U, 1)} \in [1, m].$$

Put $u = m/4096$ and let $\mathcal{G}$ be the set of atoms with $R \leq \omega/4 + u$ and $U \leq u$.

The exceptional atoms. By (13.1), the horizontal tail of the first passage (Lemma 10.14, with $u \geq 2$) bounds the $\mathbb{P}_{p,\omega}$ -mass of $\{R > \omega/4 + u\}$ by $4e^{-u/256}$ if $\omega \leq 5$ and by $4e^{-u^2/(2(\omega-4))} + 2e^{-u/256}$ otherwise; as $2(\omega - 4) \leq 4m$, both are at most $6e^{-m/2^{26}}$. The mass of $\{U > u\}$ is $(\frac{11}{16})^{\lfloor u \rfloor} \leq \frac{16}{11} e^{-5u/16} < 2e^{-m/2^{26}}$. As $\tilde{\mathbb{P}}_p \leq \mathbb{P}_{p,\omega}$ and $\theta^A \leq m^3$ ($A_l < 3$), the contribution of the complement of $\mathcal{G}$ to Z is less than $8m^3 e^{-m/2^{26}}$. With $x = m/2^{42} \geq 1$, $e^{m/2^{26}} = e^{2^{15}x} e^{2^{15}x} \geq 2^{32768} \cdot \frac{(2^{15}x)^3}{6}$, so this is less than $6 \cdot 2^{129-45-32768} < 2^{-100}$.

Large gaps. On $\mathcal{G}$, $R + U \leq \omega/4 + m/2048 \leq (\frac{\mathbf{s}_l}{4} + \frac{1}{2048})m$, so $\theta(a) \leq 2048/(2047 - 512\mathbf{s}_l)$. If $\omega \geq 2^{34}$, the total mass of $\tilde{\mathbb{P}}_p$ is $\sum_x \tilde{\mathbb{F}}_p(x) < \mathbf{q}_{\text{mass}} = \frac{38177}{125000}$ (Lemma 13.13), and the contribution of $\mathcal{G}$ is less than $\frac{999}{1000}$ by Lemma A.7, since $A \leq A_l$ and $\theta \geq 1$.

Small gaps. If $\omega < 2^{34}$, then $\omega/4 < 2^{32} \leq m/1024$, so on $\mathcal{G}$, $R + U < \frac{3m}{2048}$ and $\theta(a) < \frac{2048}{2045}$. The total mass of $\tilde{\mathbb{P}}_p$ is at most $1 - d_{\text{pass}} = \frac{96338641}{10^8}$ (Lemma 12.10), and the contribution of $\mathcal{G}$ is less than $\frac{97}{100}$ by Lemma A.7. In both cases $Z < \frac{999}{1000} + 2^{-100} < \frac{1999}{2000}$, and $\mathbf{Q}_{m-1} \geq 0$. $\square$

14.4. The descent step and the Fourier bound.

Proposition 14.4 (Descent). *Let $\iota \in \{o, b\}$ and form Q_m with $A = A_\iota$. For every integer m with $D_*(\iota) \leq m \leq J$, $Q_m \leq Q_{m-1}$.*

Proof. By Definition A.5, $m \geq D_{sc}(\iota)$ and $m \geq D_{bad}(\iota)$, which is at least 2^{43} by Lemma A.6(1). By Lemma 11.16(2) it suffices to prove $Q(p) \leq m^{-A_\iota} Q_{m-1}$ 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. $\text{gap}(p) \leq s_\iota m$. Apply Lemma 14.3 with $A = A_\iota$.

Case 3. $\text{gap}(p) > s_\iota m$. Apply Lemma 13.26. $\square$

Lemma 14.5 (An inverse moment of the first hold). *For $1 < A \leq \frac{5}{2}$ and every integer $J \geq 2^{24}$, $\sum_{r \geq 1} \nu_{45}(r) \max(J - r, 1)^{-A} \leq (J - \frac{7}{2})^{-A}$.*

Proof. Put $t = \frac{32A}{31J} \leq \frac{1}{1024}$. For $r \leq J/32$, $-\log(1 - r/J) \leq \frac{32r}{31J}$, so $\max(J - r, 1)^{-A} \leq J^{-A} e^{tr}$, and these terms sum to at most $J^{-A} M_{45}(t) \leq J^{-A} e^{104A/(31J)}$ (Lemma 10.8(2)). The terms with $r > J/32$ are at most $(\frac{11}{16})^{\lfloor J/32 \rfloor} \leq e^{-J/256}$, and $J^3 e^{-J/256} < \frac{1}{8J}$ for $J \geq 2^{24}$, as $\log J \leq \sqrt{J}$. Since $\frac{7}{2} - \frac{104}{31} = \frac{9}{62}$ and $\frac{9A}{62} > \frac{1}{8}$, $e^{104A/(31J)} + \frac{1}{8J} < e^{7A/(2J)} \leq (1 - \frac{7}{2J})^{-A}$. $\square$

Theorem 14.6 (The Fourier bound). *For each $\iota \in \{o, b\}$, every integer $n \geq 1$ and every frequency $\xi \in G_n$ prime to 3,*

$$|\hat{\mu}_n(\xi)| \leq \left(\frac{D_F(\iota)}{n} \right)^{A_\iota}.$$

This is (4.2).

Proof. Write $A = A_\iota$, $D_* = D_*(\iota)$ and $D_F = D_F(\iota) = 2D_* + 6$, so $1 < A \leq \frac{5}{2}$ and $D_* \geq 2^{43}$. If $n \leq D_F$, then $|\hat{\mu}_n(\xi)| \leq 1 \leq (D_F/n)^A$. Let $n > D_F$, so $J \geq (n-1)/2 > D_* + \frac{5}{2}$ and $J \geq D_* + 3$. Form Q and Q_m for (n, ξ) with the exponent A . Then $Q_m \leq (D_* - 1)^A$ for $0 \leq m \leq J$, by induction on m : for $m \leq D_* - 1$ by Lemma 11.15(1), and for $m \geq D_*$ by Proposition 14.4. So Lemma 11.15(2) with $m = J$ gives $Q(y) \leq (D_* - 1)^A d_J(y)^{-A}$ for every $y \in \mathcal{P}$, where $d_J(y) = \max(J - j(y), 1)$. By Proposition 11.13, grouping by $r = j(y)$ with (10.2), and Lemma 14.5,

$$|\hat{\mu}_n(\xi)| \leq \sum_y \eta(y) Q(y) \leq (D_* - 1)^A \sum_{r \geq 1} \nu_{45}(r) \max(J - r, 1)^{-A} \leq \left(\frac{D_* - 1}{J - 7/2} \right)^A.$$

Finally $n \leq 2J + 1$ gives $D_F(J - \frac{7}{2}) - n(D_* - 1) \geq (2D_* + 6)(J - \frac{7}{2}) - (2J + 1)(D_* - 1) = 8(J - D_*) - 20 \geq 4$, so $\frac{D_* - 1}{J - 7/2} < \frac{D_F}{n}$. $\square$

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 deficit of the first block family (Lemma 15.8), the deficits of the first-crossing families at every shift and their mask failure (Lemma 15.7), and the decay and the products of the deficits of the block families (Lemmas 15.9 and 15.10).

15.1. Scales and levels. The deficit bounds need linear and geometric growth of the scales, the selector needs δ_{29} , and propagation needs the levels $k_j = \lfloor \delta_j/4 \rfloor$ to reach the mixing gate $n_0 = 2^{80}$ of Proposition 4.1.

Lemma 15.1 (Scales and levels).

- (1) *For all $i, j \in \mathbb{N}$, $g^j \delta_i \leq \delta_{i+j}$, $\delta_j \geq 16544 + 166j$ and $16544g^j \leq \delta_j < 16644g^j$. Moreover $\delta_{29} = 22095$.*

(2) For every $j \in \mathbb{N}$, $k_j \geq 2^{12}g^j$; for $j \geq 9853$, $k_j \geq \max(n_0, 2^{12}g^j)$.

Proof. For (1), $e_\delta = \lceil \delta/100 \rceil$ satisfies $\delta/100 \leq e_\delta \leq (\delta + 99)/100$, and $e_\delta \geq 166$ for $\delta \geq 16544$. Hence $\delta_{m+1} \geq g\delta_m$ and $\delta_{m+1} \geq \delta_m + 166$ for every m , and induction gives the first two bounds and, with $\delta_0 = 16544$, the lower bound $16544g^j$. For the upper bound, $\delta_j + 100 \leq 16644g^j$ by induction: it is an equality at $j = 0$, and $\delta_{j+1} + 100 \leq \delta_j + (\delta_j + 99)/100 + 100 < g(\delta_j + 100)$. Running the integer recurrence 29 steps from 16544 gives δ_{29} .

For (2), $k_j > \delta_j/4 - 1 \geq 4136g^j - 1 \geq 2^{12}g^j$ as $g^j \geq 1$. The integer comparison $101^{70} > 2 \cdot 100^{70}$ gives $g^{70} > 2$, so for $j \geq 9853 \geq 70 \cdot 140$, $2^{12}g^j \geq 2^{12}g^{9800} > 2^{152} > n_0$. $\square$

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 $\delta \geq 256$, 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 $\delta \geq 256$ 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 $\delta \geq 256$, 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 [7, 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 two events of Lemma 15.3 are deviations of a partial degree from twice its length. The sub-Gaussian form of the next lemma serves all scales; its exact form, with an explicit tilt, gives the sharper bound needed for the first block (Lemma 15.8).

Lemma 15.4 (Maximal inequality). *Let $d \geq 1$ be an integer and $x > 0$ real.*

(1) *If $4x/(9d) \leq 1/32$, then*

$$\mathbb{P}\left(\max_{1 \leq i \leq d} (\deg_i(W) - 2i) \geq x\right) \leq e^{-2x^2/(9d)}, \quad \mathbb{P}(\deg_d(W) - 2d \leq -x) \leq e^{-2x^2/(9d)}.$$

(2) If $0 < \theta < 1$, then

$$\mathbb{P}\left(\max_{1 \leq i \leq d} (\deg_i(W) - 2i) \geq x\right) \leq \frac{(1 + \theta)^{-x}}{(1 - \theta^2)^d}, \quad \mathbb{P}(\deg_d(W) - 2d \leq -x) \leq \frac{(1 - \theta)^x}{(1 - \theta^2)^d}.$$

Proof. For $0 < y < 2$ the geometric series gives $\mathbb{E}y^{a_1-2} = 1/(y(2-y)) \geq 1$. If $y \geq 1$, the process $(y^{\deg_i(W)-2i})_{i \leq d}$ is therefore a nonnegative submartingale, and Doob's maximal inequality [2, Theorem 4.4.2] gives $y^x \mathbb{P}(\max_{i \leq d} (\deg_i(W) - 2i) \geq x) \leq \mathbb{E}y^{\deg_d(W)-2d} = (y(2-y))^{-d}$. If $y \leq 1$, Markov's inequality gives $\mathbb{P}(\deg_d(W) - 2d \leq -x) \leq y^x (y(2-y))^{-d}$. Part (2) is the case $y = 1 \pm \theta$. For (1) take $y = e^{\pm t}$ with $t = 4x/(9d)$: then $(y(2-y))^{-1} = M_{\text{geo}}(\pm t) \leq e^{9t^2/8}$ by Lemma 4.2, and $-tx + 9dt^2/8 = -2x^2/(9d)$. $\square$

15.4. Deficits of first-crossing families. Lemma 15.3 is applied to the block families through central length ranges (Lemmas 15.5, 15.6 and 15.8) and to every shift $0 \leq u \leq 2r_\delta$, as the terminal extension needs (Lemma 15.7).

Lemma 15.5 (Central length ranges). *Let $\delta \geq 256$, $R \geq 0$ 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 $d_\delta^{\text{lo}} \leq i < \delta - y$ 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(1) with $d = \delta + y$ bounds the two events of Lemma 15.3 by $e^{-2x^2/(9(\delta+y))}$ each. $\square$

Lemma 15.6 (Deficit of the block families). *Let $\delta \geq 16544$ and $R \geq 0$ be integers. Then*

$$1 - \text{mass}(\text{FC}(\delta, R)) \leq 2 \exp\left(-\frac{529}{10^6}(\delta - 14)\right) + 2\delta^{-3/2} + 2^{-(R+1)}.$$

Proof. Put $y_{\text{sq}} = \lceil \sqrt{32\delta \lceil \log_2 \delta \rceil} \rceil$, $y_{\text{lin}} = \lfloor \delta/8 \rfloor$ and $y = \min(y_{\text{sq}}, y_{\text{lin}})$. As $\lfloor 3\delta/5 \rfloor - 1 > \delta/8 \geq y_{\text{lin}}$, $y \leq \ell_\delta^c$, so $\text{FC}(\delta, R) = \mathcal{V}_{\ell_\delta^c} \supseteq \mathcal{V}_y$ in the notation of Lemma 15.5. Also $y_{\text{lin}} \geq 2068$ and $y_{\text{sq}}^2 \geq 32 \cdot 16544 \cdot 15 > 1000^2$, so $1000 \leq y \leq \delta/8$, and Lemma 15.5 gives $1 - \text{mass}(\text{FC}(\delta, R)) \leq 2e^{-x} + 2^{-(R+1)}$ with $x = c_w y^2/(\delta + y)$ and $c_w = 2(207/500)^2/9 = 4761/125000$.

Case $y_{\text{lin}} \leq y_{\text{sq}}$. Then $y = y_{\text{lin}} \geq (\delta - 7)/8$. As x increases with $y > 0$,

$$x \geq c_w \frac{(\delta - 7)^2}{8(9\delta - 7)} \geq \frac{c_w}{72}(\delta - 14) = \frac{529}{10^6}(\delta - 14),$$

since $9(\delta - 7)^2 - (\delta - 14)(9\delta - 7) = 7\delta + 343 > 0$.

Case $y_{\text{sq}} < y_{\text{lin}}$. Then $\delta + y \leq 9\delta/8$ and $y^2 \geq 32\delta \log_2 \delta$, so $x \geq \frac{256c_w}{9} \log_2 \delta \geq \frac{3}{2} \log \delta$, because $\frac{256c_w}{9} = \frac{16928}{15625} > \frac{3}{2} \cdot \frac{347}{500} > \frac{3}{2} \log 2$ (Lemma A.1(3)). Hence $2e^{-x} \leq 2\delta^{-3/2}$. $\square$

Lemma 15.7 (First-crossing families and the mask). *Let $\delta \geq 256$, $R \geq 0$ and $0 \leq u \leq 2r_\delta$ be integers. Then*

$$1 - \text{mass}(\text{FC}_u(\delta, R)) \leq 2^{116}\delta^{-6} + 2^{-(R+1)},$$

$$\text{mass}(\{w \in \text{FC}_u(\delta, R) : 2|w| > H_{\delta, u}(|w|) + \lfloor \delta/8 \rfloor\}) \leq 2^{113}\delta^{-6};$$

the words of the second set are those failing the mask (5.9) at scale δ and shift u .

Proof. We follow [7, 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 $0 \leq u \leq 2r_\delta$,

$$H_{\delta,u}(d_\delta^{\text{lo}}) \geq 2d_\delta^{\text{lo}} + 2e_\delta, \quad H_{\delta,u}(d_\delta^{\text{hi}}) \leq 2d_\delta^{\text{hi}} - 2e_\delta.$$

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 (4.6) 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. 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 (4.6), and $e^{-\delta/16384} \leq 7! (\delta/16384)^{-7}$ gives $\text{mass}(\mathcal{V}) \leq 4 \cdot 7! \cdot 16384^7 \delta^{-6} < 2^{113} \delta^{-6}$. $\square$

The average at generation j_0 (Section 6) needs the first block family far more precisely than Lemma 15.6 gives; the exact tilts of Lemma 15.4(2) provide it.

Lemma 15.8 (The first block family). $1 - \text{mass}(\text{FC}(\delta_0, R_0)) < 1/20000000$.

Proof. Here $\delta = \delta_0 = 16544$ and $R_0 = 29$, so $\ell_\delta = 9926$, $d_\delta^{\text{lo}} = 6618$ and $\lceil \log_2 \delta \rceil = 15$; as $2817^2 < 32 \cdot 16544 \cdot 15 \leq 2818^2$, $\ell_\delta^c = 2818$, and $\text{FC}(\delta, 29)$ consists of the words of $\text{FC}_{r_\delta}(\delta, 29)$ with length in $[13726, 19362]$. Apply Lemma 15.3 with $d' = 13726$ and $d'' = 19362$; the shift cancels as in the proof of Lemma 15.5.

If $6618 \leq i < 13726$ and $\deg_i(W) \geq H_{\delta,r_\delta}(i)$, then $m = \delta - i \geq 2819$ and $\deg_i(W) - 2i \geq 2m - B_0(m) > \frac{83}{200}m - 1 > 1168$, so the integer $\deg_i(W) - 2i$ is at least 1169. Lemma 15.4(2) with $d = 13725$ and $\theta = \frac{1}{24}$, and $1 + s \leq e^s$, bound the first event of Lemma 15.3 by

$$\left(\frac{24}{25}\right)^{1169} \left(\frac{576}{575}\right)^{13725} \leq \exp\left(-\frac{1169}{25} + \frac{13725}{575}\right) = e^{-13162/575} < e^{-22}.$$

If $\deg_{19362}(W) < H_{\delta,r_\delta}(19362) = 2\delta + B_0(2818)$, then $\deg_{19362}(W) - 2 \cdot 19362 \leq B_0(2818) - 5637$, and $B_0(2818) - 5636 < 1 - \frac{83}{200} \cdot 2818 < -1168$, so $\deg_{19362}(W) - 2 \cdot 19362 \leq -1170$. Lemma 15.4(2) with $d = 19362$ and $\theta = \frac{1}{32}$ bounds the second event by

$$\left(\frac{31}{32}\right)^{1170} \left(\frac{1024}{1023}\right)^{19362} \leq \exp\left(-\frac{1170}{32} + \frac{19362}{1023}\right) = e^{-96221/5456} < e^{-35/2}.$$

By Lemma A.1(4) with $k = 60$ and $k = 50$, $e^{22} > 3 \cdot 10^9$ and $e^{35/2} > 39 \cdot 10^6$, and $\frac{1}{3 \cdot 10^9} + \frac{1}{39 \cdot 10^6} + 2^{-30} = \frac{2200586851}{817889280000000000} < \frac{1}{20000000}$. $\square$

15.5. Deficits along the scale sequence. At the scales δ_j , Lemma 15.6 gives the decay of the block deficits used for propagation and the lower bound on products of block masses used in Lemma 6.6. The cap terms satisfy

$$2^{-(R_j+1)} \leq 2^{-29} 2^{-j/56} \quad (j \in \mathbb{N}), \quad (15.1)$$

since $R_j + 1 = 30 + \lfloor j/56 \rfloor \geq 29 + (j+1)/56$.

Lemma 15.9 (Decay of the block deficits). *For every $j \in \mathbb{N}$, $1 - \text{mass}(\text{FC}(\delta_j, R_j)) \leq \frac{1}{2} g^{-3j/2} + 2^{-29} 2^{-j/56}$.*

Proof. Put $c = 529/10^6$. By Lemma 15.1, $\delta_j - 14 \geq 16530 + 166j$, and $166c > \frac{3}{200} \geq \frac{3}{2} \log g$ as $\log g < \frac{1}{100}$; with $e^{16530c} \geq 1 + 16530c > 8$ this gives $2e^{-c(\delta_j-14)} \leq \frac{1}{4} g^{-3j/2}$. Also $\delta_j \geq 16544g^j$ gives $2\delta_j^{-3/2} \leq 2 \cdot 16544^{-3/2} g^{-3j/2} \leq \frac{1}{4} g^{-3j/2}$. Lemma 15.6 and (15.1) conclude. $\square$

Lemma 15.10 (Products of block masses). *For all integers $0 \leq j_1 \leq j_2$,*

$$\prod_{j_1 \leq j < j_2} \text{mass}(\text{FC}(\delta_j, R_j)) > \frac{199}{200}.$$

Proof. All factors lie in $[0, 1]$ (Lemma 15.2), so the product is at least $1 - \sum_{j \geq 0} (1 - \text{mass}(\text{FC}(\delta_j, R_j)))$. We bound the three sums arising from Lemma 15.6. With $c = 529/10^6$, $\delta_j - 14 \geq 16530 + 166j$ gives $\sum_j 2e^{-c(\delta_j - 14)} \leq 2e^{-16530c}/(1 - e^{-166c})$, where $1 - e^{-166c} \geq 166c - (166c)^2/2 > \frac{83}{1000}$ and $e^{16530c} > 6275$ (Lemma A.1(4) with $k = 40$). Next $\delta_j \geq 16544g^j$, $16544^3 \geq 2127000^2$ and $1 - g^{-3/2} \geq 1 - g^{-1} = \frac{1}{101}$ give $\sum_j 2\delta_j^{-3/2} \leq \frac{202}{2127000}$. Each value of $\lfloor j/56 \rfloor$ occurs 56 times, so $\sum_j 2^{-(R_j+1)} \leq 56 \cdot 2^{-29}$. Finally

$$\frac{2}{6275} \cdot \frac{1000}{83} + \frac{202}{2127000} + \frac{56}{2^{29}} < \frac{1}{200}. \quad \square$$

16. SOURCES AND THE HISTOGRAM ESTIMATE

This section proves the spread of the sources (Lemma 16.1), used in Lemmas 6.10 and 6.8, and the histogram estimate (5.8) (Proposition 16.5) used in Section 6. The arguments are those of [7, proof of Proposition 3.1], made explicit.

16.1. Sources.

Lemma 16.1 (Spread of the sources). *Let $t \geq 16^{16544}$ 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 5.3 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 selector pins down the offset of every selected word of at least 29 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 $j \geq 29$ and $w = v_0 \cdots v_{j-1} \in \bar{L}_j$. Then $0 < \text{off}(w) \leq \frac{3}{4} + 2^{-8}$.*

Proof. For $29 \leq i \leq j$ 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). By Lemma 5.2, $\text{off}(v_i) < 2^{\delta_i}$ and $\text{wt}(v_i) < 16^{-e_{\delta_i}}$. If $\mathfrak{o}_{i+1} \leq 2^{\delta_{i+1}+1}$, then, as $e_{\delta_i} \geq 1$ and $\delta_{i+1} = \delta_i + e_{\delta_i}$,

$$\mathfrak{o}_i \leq 2^{\delta_i} + 16^{-e_{\delta_i}} \cdot 2^{\delta_i + e_{\delta_i} + 1} = 2^{\delta_i} + 2^{\delta_{i+1} - 3e_{\delta_i}} \leq 2^{\delta_{i+1}}.$$

Downward induction from $i = j$ gives $0 \leq \mathfrak{o}_{29} \leq 2^{\delta_{29}+1}$. By (3.2),

$$\text{off}(w) = \text{off}(v_0) + \sum_{0 < i < 29} \text{wt}(w_{<i}) \text{off}(v_i) + \text{wt}(w_{<29}) \mathfrak{o}_{29},$$

with all terms nonnegative and the first positive. Since $\sum_{l < i} e_{\delta_l} = \delta_i - \delta_0 \geq 166i$ (Lemma 15.1), $\text{wt}(w_{<i}) < 16^{-166i}$. With (5.5), the middle sum is at most $\sum_{i \geq 1} 2^{600-664i} < 2^{-63}$, and, as $\delta_{29} - \delta_0 = 5551$, the last term is less than $16^{-5551} 2^{22096} = 2^{-108}$. Hence $\text{off}(w) < \frac{3}{4} + 2^{-62}$. $\square$

16.3. The histogram estimate. The histogram estimate (5.8) 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, $j \geq 9853$, $0 \leq q \leq 2\delta_j$, $y \in G_q$, and $d, l \in \mathbb{Z}$. Then*

$$3^q \sum_{\substack{w \in L_j, |w|=d, \deg(w)=l \\ [s_w]_q = y}} \text{wt}(w) < \frac{97}{128}.$$

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). Lemma 16.2 places $s_w = (t - \text{off}(w))/\text{wt}(w)$ in an interval of length $(\frac{3}{4} + 2^{-8})\text{wt}(w)^{-1}$. 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 $(\frac{3}{4} + 2^{-8})\text{wt}(w)^{-1}3^{-q} + 1$ of them. The left side is therefore at most $\frac{3}{4} + 2^{-8} + 3^q \text{wt}(w)$.

By Lemma 5.2, $\text{wt}(v_i) < 16^{-e_{\delta_i}}$ for every block v_i of w , and $\sum_{i < j} e_{\delta_i} = \delta_j - \delta_0$, so $\text{wt}(w) < 16^{16544 - \delta_j}$. With $3^q \leq 9^{\delta_j}$ and $16 \cdot 9^5 < 16^5$,

$$3^q \text{wt}(w) < 2^{66176} (9/16)^{\delta_j} \leq 2^{66176 - 4\lceil \delta_j/5 \rceil} \leq 2^{-1255536},$$

using $\delta_j \geq 16544 + 166 \cdot 9853 = 1652142$ (Lemma 15.1). Since $\frac{3}{4} + 2^{-8} + 2^{-8} = \frac{97}{128}$, the lemma follows. $\square$

Lemma 16.4 (Lengths and degrees of words). *Let t be a positive odd integer and $j \geq 9853$.*

- (1) *The lengths $|w|$, $w \in L_j$, take at most $2 \sum_{i < j} \ell_{\delta_i}^c + 1 < 37170 \sqrt{j} g^{j/2}$ values.*
- (2) *For each length, the degrees $\deg(w)$, $w \in L_j$, take fewer than R_Σ values.*

Proof. For (1), every block v_i of w has length in $[\delta_i - \ell_{\delta_i}^c, \delta_i + \ell_{\delta_i}^c]$ by the length restriction, so $|w|$ is one of the $2 \sum_{i < j} \ell_{\delta_i}^c + 1$ integers within $\sum_{i < j} \ell_{\delta_i}^c$ of $\sum_{i < j} \delta_i$. By Lemma 15.1, $\delta_i < 16644g^i < 2^{15}g^i$, so $\log_2 \delta_i < 15 + i/69$ (as $\log_2 g < 1/69$), and for $i < j$, $\lceil \log_2 \delta_i \rceil < 16 + j/69 \leq j/62$ because $7j \geq 68448$. With $32 \cdot 16644 \cdot 10^4 \leq 62 \cdot 9269^2$,

$$\ell_{\delta_i}^c < \sqrt{32\delta_i \lceil \log_2 \delta_i \rceil} + 1 < \frac{9269}{100} \sqrt{j} g^{i/2} + 1 \quad (i < j).$$

The integer comparison $100 \cdot 403^2 < 101 \cdot 401^2$ gives $g^{1/2} - 1 > 2/401$, hence $\sum_{i < j} g^{i/2} < \frac{401}{2} g^{j/2}$. The binomial theorem gives $g^j \geq \binom{j}{4} 100^{-4} \geq 49j$, as $(j-1)(j-2)(j-3) \geq 24 \cdot 49 \cdot 10^8$; thus $2j+1 \leq 3j \leq \frac{3}{7} \sqrt{j} g^{j/2}$, and

$$2 \sum_{i < j} \ell_{\delta_i}^c + 1 < \left(\frac{3716869}{100} + \frac{3}{7} \right) \sqrt{j} g^{j/2} < 37170 \sqrt{j} g^{j/2}.$$

For (2), the inequalities $H_{\delta, r_\delta}(|v|) \leq \deg(v) \leq H_{\delta, r_\delta}(|v|) + R$ for $v \in \text{FC}(\delta, R)$ 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 $\sum_i |v_i| = d$, places $\deg(w)$ in an open interval, depending only on d , of integer length $\sum_{i < j} (R_i + 2) < R_\Sigma$; it contains fewer than R_Σ integers. $\square$

Proposition 16.5 (Histogram estimate). *Let t be a positive odd integer, $j \geq 9853$ and $0 \leq q \leq 2\delta_j$. Then*

$$3^q \Xi_{j,q}(t; y) < 344 j^{5/2} g^{j/2} \quad (y \in G_q).$$

Consequently (5.8) 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 and (5.1), fewer than $37170\sqrt{j} g^{j/2} \cdot j^2/82$ pairs occur, and by Lemma 16.3 each contributes less than $\frac{97}{128} 3^{-q}$; and $\frac{97 \cdot 37170}{128 \cdot 82} = \frac{1802745}{5248} < 344$. For (5.8), group a subset of L_j by the residue $y = [s_w]_q$; since weights are nonnegative,

$$\sum_w \text{wt}(w) |f(s_w)| \leq \sum_{y \in G_q} \Xi_{j,q}(t; y) |f(y)| \leq \max_y (3^q \Xi_{j,q}(t; y)) \langle |f| \rangle_q. \quad \square$$

APPENDIX A. NUMERICAL CONSTANTS AND FINITE CHECKS

This appendix defines the constants of the Fourier argument for both indices $\iota \in \{o, b\}$ and records the finite computations behind the numerical claims. Apart from the standard values stated in Lemma A.1, every claim is an exact computation: finite decimals denote exact rational numbers, all integer arithmetic is exact, a ceiling $\lceil x/y \rceil$ of a quotient of positive integers is evaluated as $\lfloor (x+y-1)/y \rfloor$, and a value written with “...” is a truncation shown for orientation only.

Lemma A.1 (Elementary bounds).

- (1) $\log_2 3 < 317/200$.
- (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 $e^9 > 8000$.

Proof. Parts (1) and (3), and the numerical claim in (4), follow from the standard values $\log_2 3 = 1.58496\dots$, $\log 2 = 0.6931471805\dots$, $\log(4/3) = 0.2876820724\dots$ and $e^9 = 8103.08\dots$. 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 Fourier parameters and the threshold D_* . The Fourier bound (4.2) is proved twice by one argument, for the original index o and the bootstrap index b ; the two parameter sets differ in the constants of Table 1. This subsection defines them, the trace schedule built from them and the threshold $D_*(\iota)$ above which the descent step holds, and proves their numerical properties.

Definition A.2 (Fourier parameters). The constants common to both indices are

$$\varepsilon = \frac{37037}{10^6}, \quad z = \frac{133}{3125}, \quad \kappa = \frac{3187}{20000}, \quad \chi = \frac{63}{2500}, \quad b_* = \frac{333407}{10^6}, \quad \ell_{\text{cont}} = \frac{157263}{100000}.$$

The constants depending on ι are listed in Table 1, where V_ι is defined in Definition A.3. The size factors are

$$(K_0(\iota), \dots, K_{r_\iota}(\iota)) = \begin{cases} (4638, 1622, 1253, 971, 754, 587, 458, 359, 292, 2^{30}, 2^{30}), & \iota = o, \\ (718, 268, 1024, 2^{30}), & \iota = b. \end{cases}$$

Definition A.3 (Trace schedule). The *schedule positions* are $p_0 = p_0(\iota)$ and

$$p_{i+1} = p_i + \lceil \alpha_\iota S_i \rceil + H_\iota + T_\iota + 1, \quad S_i = K_i(\iota)(p_i + 1) \quad (0 \leq i < r_\iota),$$

with $S_{r_\iota} = K_{r_\iota}(\iota)(p_{r_\iota} + 1)$. Put $P_\iota = p_{r_\iota} + 1$ and $V_\iota = p_{r_\iota-1} + 1$.

TABLE 1. The constants of the two Fourier parameter sets.

quantity	$\iota = \text{o}$	$\iota = \text{b}$
exponent A_ι ; $a_\iota = A_\iota - \frac{1}{4}$	$\frac{21193}{10000}, \frac{18693}{10000}$	$\frac{17}{10}, \frac{29}{20}$
Case-2 split $\mathbf{s}_\iota$	$\frac{3}{2}$	2
reward bound T_ι ; first position $p_0(\iota) = T_\iota + 1$	192; 193	168; 169
schedule length r_ι	10	3
window height H_ι ; slope α_ι	10100; $\frac{113}{1000}$	500; $\frac{3}{25}$
event coefficients $\mathbf{e}_i(\iota)$, $0 \leq i \leq r_\iota$	$(\frac{77}{100})^i$	$1, \frac{77}{100}, b_*, b_*^2$
far-gap scale $R_\Omega(\iota)$	$\lceil 1231V_\text{o}/4 \rceil$	$45V_\text{b}$
far-gap threshold Ω_ι	$R_\Omega(\text{o})^2$	$R_\Omega(\text{b})^2$
continuation factor $c_\text{cont}(\iota)$	$\frac{14009206687631}{500000000000}$	$\frac{7245014059}{500000000}$
prefactor bound: $\log_2 B_F(\iota) <$	325.7939550288	88.1188851177

Lemma A.4 (The trace schedule). *The positions are*

i	$p_i, \iota = \text{o}$	$p_i, \iota = \text{b}$
0	193	169
1	112161	15486
2	20680179	514217
3	2948776479	63701994
4	326497388488	
5	28144727892821	
6	1895012673761891	
7	99969498591645143	
8	4155432147958923993	
9	141268071302011590393	
10	17140454471732661035095372255	

so $2^{18} \leq V_\iota < P_\iota < 2^{94}$ and $P_\iota \geq 2^{22}$. Moreover

$$R_\Omega(\text{o}) = 43475248943194066943754, \quad R_\Omega(\text{b}) = 23139810,$$

$$\Omega_\text{o} = 1890097270672696834631662319882702018199612516, \quad \Omega_\text{b} = 535450806836100;$$

both Ω_ι are divisible by 6, and $\Omega_\iota \geq \max(2^{34}, 512V_\iota)$.

Proof. The table is the recurrence evaluated in integer arithmetic; it gives V_ι , P_ι and $R_\Omega(\iota)$ (for o, $1231V_\text{o} = 4 \cdot 43475248943194066943753 + 2$), and the bounds are integer comparisons. Finally $R_\Omega(\iota) \geq V_\iota \geq 2^{18}$ gives $\Omega_\iota \geq 2^{18}R_\Omega(\iota) \geq \max(2^{34}, 512V_\iota)$. $\square$

Definition A.5 (Thresholds). For $\iota \in \{\text{o}, \text{b}\}$ put

$$D_\text{bad}(\iota) = 2^{21}(P_\iota + 64), \quad D_\text{sc}(\iota) = 2 + P_\iota + \lceil \Omega_\iota / \mathbf{s}_\iota \rceil, \quad D_*(\iota) = \max(D_\text{bad}(\iota), D_\text{sc}(\iota)),$$

the Fourier threshold $D_F(\iota) = 2D_*(\iota) + 6$, and $B_F(\iota) = (\frac{15}{2}D_F(\iota))^{A_\iota}$.

Case 3 of the descent step needs $m \geq D_\text{bad}(\iota)$ for its bad advance and $m \geq D_\text{sc}(\iota)$ for its horizon P_ι and its gaps $\omega > \mathbf{s}_\iota m \geq \Omega_\iota$ (Section 13); the near-top case needs $m \geq 2^{42}$.

Lemma A.6 (The thresholds).

- (1) $D_*(\iota) = D_\text{sc}(\iota) \geq 2^{47}$ and $D_\text{bad}(\iota) \geq 2^{43}$, with $D_\text{bad}(\text{o}) < 2^{116}$, $D_\text{bad}(\text{b}) = 133592900435968$ and

$$D_\text{sc}(\text{o}) = 1260064847115131240228229351654462380561780602, \quad D_\text{sc}(\text{b}) = 267725467120047.$$

Hence

$$D_F(\text{o}) = 2520129694230262480456458703308924761123561210, \quad D_F(\text{b}) = 535450934240100.$$

- (2) Every real $m \geq D_{\text{sc}}(\iota)$ satisfies $P_\iota \leq m$ and $\mathbf{s}_\iota m \geq \Omega_\iota$.
 (3) The prefactors satisfy

$$325.7939550287 < \log_2 B_F(\mathbf{o}) < 325.7939550288, \quad 88.1188851176 < \log_2 B_F(\mathbf{b}) < 88.1188851177.$$

In particular $B_F(\mathbf{o}) < 2^{349}$.

Proof. For (1), $\Omega_\iota/\mathbf{s}_\iota$ is an integer, so $D_{\text{sc}}(\iota) = 2 + P_\iota + 2\Omega_\iota/3$ for $\mathbf{o}$ and $2 + P_\iota + \Omega_\iota/2$ for $\mathbf{b}$; the values and the comparisons are integer computations from Lemma A.4, and $D_{\text{bad}}(\iota) \geq 2^{21}P_\iota \geq 2^{43}$. Part (2) follows from $D_{\text{sc}}(\iota) \geq P_\iota$ and $D_{\text{sc}}(\iota) \geq \Omega_\iota/\mathbf{s}_\iota$. For (3), write $\frac{15}{2}D_F(\iota) = 2^e u$ with $e \in \mathbb{N}$ and $1 \leq u < 2$ ($e = 153$ for $\mathbf{o}$, $e = 51$ for $\mathbf{b}$). Then $\log_2 B_F(\iota) = A_\iota(e \log 2 + \log u)/\log 2$, and (A.1) with $N = 80$ for $\log 2$ and $\log u$, the lower bounds in the numerator and the upper bound for $\log 2$ in the denominator (and conversely), gives rational enclosures of width below 10^{-20} that lie in the stated intervals. $\square$

The near-top case (Section 14) bounds the paid mass of a first passage over the top edge, times an inverse power of the remaining distance to the horizon; it closes by four integer comparisons.

Lemma A.7 (Numerical facts for Case 2). *The following integer comparisons hold:*

$$\begin{aligned} (38177 \cdot 1000)^{10000} 2048^{21193} &< (125000 \cdot 999)^{10000} 1279^{21193}, \\ 96338641^{10000} 2048^{21193} &< (97 \cdot 10^6)^{10000} 2045^{21193}, \\ (38177 \cdot 1000)^{10} 2048^{17} &< (125000 \cdot 999)^{10} 1023^{17}, \quad 96338641^{10} 2048^{17} < (97 \cdot 10^6)^{10} 2045^{17}. \end{aligned}$$

Equivalently, with $f_\iota = (2047 - 512\mathbf{s}_\iota)/2048$, that is $f_{\mathbf{o}} = \frac{1279}{2048}$ and $f_{\mathbf{b}} = \frac{1023}{2048}$, and $b_s = \frac{2045}{2048}$,

$$\frac{38177}{125000} f_\iota^{-A_\iota} < \frac{999}{1000}, \quad \frac{96338641}{10^8} b_s^{-A_\iota} < \frac{97}{100}.$$

Proof. Raise the second set of inequalities to the power 10000 (for $\mathbf{o}$) or 10 (for $\mathbf{b}$) and clear denominators. $\square$

A.2. The error budget. This subsection defines the error budget (A.2) of Case 3 and shows that it is less than 1 for both indices. It bounds the Case 3 error $\mathcal{E}_\iota$ of Definition 13.23 after rounding each of its pieces upward.

Definition A.8 (Error budget). Put $x_W(\iota) = zT_\iota - A_\iota \ell_{\text{cont}}$ and $P_W(\iota) = \sum_{j=0}^{20} x_W(\iota)^j/j!$. With the notation of Definition 13.23, let Π_ι^+ be Π_ι^{ad} with each of its r_ι terms rounded up to a multiple of 10^{-16} , let $E_\iota^{\text{cl},+}$ be E_ι^{cl} with each of its terms rounded up to a multiple of 10^{-24} , and let $U_\iota^{\mathbf{o}}, U_\iota^{\mathbf{v}}, U_\iota^{\mathbf{d}}$ be $9(\frac{25}{27})^{X_\iota^{\mathbf{o}}}, (y_\iota^{\mathbf{v}})^{-X_\iota^{\mathbf{v}}(0)}$ and $2(b_\iota^{\mathbf{d}})^{-4X_\iota^{\mathbf{d}}(0)}$ rounded up to multiples of 10^{-20} . The *error budget* is

$$\mathcal{B}_\iota = \frac{1}{8000} + \frac{\mathbf{q}_{\text{mass}}}{P_W(\iota)} + \left(c_{\text{cont}}(\iota) - \frac{1}{P_W(\iota)}\right) \mathcal{E}_\iota^+, \quad (\text{A.2})$$

where

$$\mathcal{E}_\iota^+ = \Pi_\iota^+ + \mathbf{q}_{\text{mass}} \mathbf{e}_{r_\iota}(\iota) + E_\iota^{\text{sq}} + U_\iota^{\mathbf{o}} + \left(\frac{11}{16}\right)^{H_\iota^{\text{pk}}} + \frac{2161}{2^{127}} + \frac{148}{\Omega_\iota} + \mathbf{q}_{\text{mass}}(U_\iota^{\mathbf{v}} + U_\iota^{\mathbf{d}} + E_\iota^{\text{cl},+}) \geq \mathcal{E}_\iota.$$

In order, the terms of $\mathcal{B}_\iota$ pay for the bad advance, the paid mass outside the low-reward event, and, through the continuation factor, the packing, the trace survival, the square-root deviations, the overshoot deviation, the common initial exception, and the vertical and drift-free deviations and the clearing of the paid mass.

Lemma A.9 (The budget closes). *For both indices:*

- (1) $(400/83)^{A_\iota} < c_{\text{cont}}(\iota)$ and $e^{-zT_\iota} (400/83)^{A_\iota} < 1/P_W(\iota) < 1$.

(2) *The rounded quantities are*

$$\begin{aligned}\Pi_o^+ &= \frac{117855141445251}{10^{16}}, & \Pi_b^+ &= \frac{172936863842033}{10^{16}}, \\ E_o^{\text{cl},+} &= \frac{25092942460067403877}{10^{24}}, & E_b^{\text{cl},+} &= \frac{12788}{10^{24}}, \\ 10^{20}(U_o^o, U_o^v, U_o^d) &= (295143990568991, 13370061336019939, 6230745211954701), \\ 10^{20}(U_b^o, U_b^v, U_b^d) &= (743222777209534, 27986092430721741, 13539748627993287).\end{aligned}$$

(3) $0.019300993707926 < 1 - \mathcal{B}_o < 0.019300993707927$ and $0.002435391461730 < 1 - \mathcal{B}_b < 0.002435391461731$. In particular $\mathcal{B}_i < 1$.

Proof. For (1), raising to the power 10000 (for o) or 10 (for b) turns the first claim into an integer comparison, for o

$$400^{21193} \cdot 5000000000000^{10000} < 83^{21193} \cdot 14009206687631^{10000}.$$

Lemma A.1(4) with $k = 20$ gives $e^{\ell_{\text{cont}}} \geq 4.819306\dots > 400/83 = 4.819277\dots$, so $(400/83)^{A_i} < e^{A_i \ell_{\text{cont}}}$; and $x_W(o) = \frac{4838645241}{10^9} > 0$, $x_W(b) = \frac{4476609}{10^6} > 0$ give $1 < P_W(i) \leq e^{x_W(i)}$, whence the second claim. Part (2) is Definition 13.23 evaluated with Lemma A.4 in rational arithmetic; for instance U_o^o is checked by $9 \cdot 25^{194} \cdot 10^{20} \leq 295143990568991 \cdot 27^{194} < 9 \cdot 25^{194} \cdot 10^{20} + 27^{194}$. In E_i^{cl} the counts $n_{r_i}^{\text{cl}}$ vanish, and all other terms except those with $i \leq 1$ (for o) and $i = 0$ (for b) round to 10^{-24} . For (3), $\mathcal{B}_i$ is an explicit rational number; in it $P_W(o) = 126.298\dots$, $E_o^{\text{sq}} = 0.000686\dots$, $P_W(b) = 87.935\dots$ and $E_b^{\text{sq}} = 0.01726\dots$, and the stated bounds are rational comparisons. $\square$

A.3. The joint-law tables. The average at generation 9853 (Section 6) bounds below the joint law of the offset of the random word and the residue modulo 3^8 of the offset of its first eight letters by exact integer tables, and combines it with Lemmas 15.8 and 15.10. This subsection records the tables and the final comparison.

The *offset grid* is $\xi_k = n_k/512$ for $0 \leq k \leq 6198$, with $n_k = k$ for $k \leq 4096$ and $n_{4096+m} = \left\lfloor 4096(65/64)^m + \frac{1}{2} \right\rfloor$ for $1 \leq m \leq 2102$; it increases, and $\xi_{384} = \frac{3}{4}$. For $y \geq 0$ let $\text{ix}(y)$ be the largest $k \leq 6198$ with $\xi_k \leq y$; put $y_{k,a} = (2^a \xi_k - 1)/3$ and $F^-(x) = \max(0, 1 - 7/(x+2))$, a lower bound for $F_{D,0}(0, x)$, every D , by Lemma 4.3. For a table $\mathsf{T} = (\mathsf{T}(k))_{k \leq 6198}$ of integers let $c_{\mathsf{T}}(k, a)$ be 0 if $y_{k,a} < 0$, $\lfloor 2^{-a} \lfloor 2^{62} F^-(y_{k,a}) \rfloor \rfloor$ if $y_{k,a} > \xi_{6198}$, and $\lfloor 2^{-a} \mathsf{T}(\text{ix}(y_{k,a})) \rfloor$ otherwise. For tables $Y_q(r, k)$, $1 \leq q \leq 8$, $r \in G_q$, $k \leq 6198$, and $1 \leq a \leq 60$ let $c_q^{\text{jt}}(r, k, a)$ be 0 unless $3 \mid 2^a r - 1$ (with $0 \leq r < 3^q$); otherwise it is $c_{Y_0}(k, a)$ if $q = 1$, and for $q \geq 2$ it is $\left\lfloor 2^{-a} Y_{q-1}(\frac{2^a r - 1}{3} \bmod 3^{q-1}, \text{ix}(y_{k,a})) \right\rfloor$ if $0 \leq y_{k,a} \leq \xi_{6198}$ and 0 otherwise.

Lemma A.10 (The joint-law tables). *There are integer tables $Y_0, \dots, Y_8$ with*

$$\begin{aligned}Y_0(0) \leq Y_0(1) \leq \dots \leq Y_0(6198) &\leq \left\lfloor 2^{62} F^-(\xi_{6198}) \right\rfloor, & Y_0(k) &\leq \max\left(\left\lfloor 2^{62} F^-(\xi_k) \right\rfloor, \sum_{a=1}^{60} c_{Y_0}(k, a)\right), \\ Y_q(r, k) &\leq \sum_{a=1}^{60} c_q^{\text{jt}}(r, k, a) \quad (1 \leq q \leq 8), & Y_8(6275, 384) &= 617014680840434.\end{aligned}$$

Moreover, with $\mathbf{m}_8$ as in Lemma 6.6, $6561\mathbf{m}_8 - \frac{13}{15} - 2^{-28} > 0.0064$.

Proof. The table Y_0 is the stationary point, reached after 724 sweeps, of the iteration

$$Y_0^{(0)}(k) = \left\lfloor 2^{62} F^-(\xi_k) \right\rfloor, \quad Y_0^{(s+1)}(k) = \max_{l \leq k} \max \left(Y_0^{(s)}(l), \sum_{a=1}^{60} c_{Y_0^{(s)}}(l, a) \right);$$

its entries include $Y_0(384) = 997765252056324269$ and $Y_0(6198) = 4611686018427359570$, which is $\left\lfloor 2^{62} F^-(\xi_{6198}) \right\rfloor$. For $q \geq 1$ the table Y_q vanishes outside the reachable pairs: $\mathcal{Y}_8 = \{(6275, 384)\}$,

Hence $\mathfrak{p}_M^{10} \leq D_{*,M} < \mathfrak{p}_M^{11}$. As $640 \cdot 6409 < 2^{22}$ and $6409 \cdot 11 < 2^{17}$, it follows that $64090 \log_2 \mathfrak{p}_M < \log_2 C_M < 2 + 6409(22 + 11 \log_2 \mathfrak{p}_M) < 2^{17} \log_2 \mathfrak{p}_M$, and (B.1) gives

$$3^{r_M} < \log_2 C_M < 2^{20} a_M 3^{r_M}, \quad \log_2 C_M > 4 \cdot 2^{80} + 1000. \quad (\text{B.2})$$

Put $b_{0,M} = 2^{80}$. The next parameters [7, (8.1)–(8.2)] are

$$F_M = 2^{467} b_{0,M} \cdot 16^{b_{0,M}} (C_M + 1), \quad N_{*,M} = 20000(\lceil \log_2 F_M \rceil + 64), \quad N_{t,M} = 20000(\lceil \log_2 (21 F_M) \rceil + 96).$$

Since $\log_2 F_M = 547 + 4b_{0,M} + \log_2(C_M + 1)$,

$$20000 \log_2 C_M \leq N_{*,M} < 40000 \log_2 C_M, \quad N_{t,M} < 40000 \log_2 C_M.$$

The recurrence $b_{j+1,M} = b_{j,M} + \lfloor b_{j,M}/100 \rfloor$ gives $b_{j,M} \leq 2^{80+j}$. Unfolding the definition [7, (8.3)–(8.4)] of $q_{*,M}$ yields

$$q_{*,M} = \left\lfloor b_{N_{*,M},M}/4 \right\rfloor + \sum_{i < N_{*,M}} (b_{i,M} + \lfloor 3b_{i,M}/5 \rfloor), \quad N_{*,M} \leq q_{*,M} \leq (2N_{*,M} + 1)b_{N_{*,M},M}.$$

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 [7, (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 [7, (8.4)–(8.6)]

$$J_{*,M} = \max(N_{*,M} + N_{t,M}, 200(\lceil \log_2 m_{*,M} \rceil - 78)_+, 20 \cdot 10^9),$$

$$\mathfrak{l}_{0,M} = 0, \quad \mathfrak{l}_{j+1,M} = \mathfrak{l}_{j,M} + 17 + \lfloor j/100 \rfloor + 2b_{j,M} - \lfloor 19b_{j,M}/12 \rfloor,$$

$$B_{\text{ht},M} = \mathfrak{l}_{J_{*,M},M} + \left\lfloor 2b_{J_{*,M},M}/3 \right\rfloor,$$

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

$$\log_2^{o5}(Y) < 2r_M + 207 < 2^{4a_M+17}.$$

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

$$\log_2^{o6}(Y) < 4a_M + 17 = 2^{172} + 17 < 2^{173}, \quad \log_2^{o7}(Y) < 173.$$

Since $Y > \mathcal{M}_M$ and $q_{*,M} > 100$, the iterated logarithms $\log_2^{oi}(Y)$ 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] 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>.
- [7] 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.
- [8] 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.
- [9] 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