Recurrence, transience and anti-concentration of Rademacher random walks

Satyaki Bhattacharya* Edward Crane† Tom Johnston†

October 29, 2025

Abstract

The Rademacher random walk associated with a deterministic sequence $(a_n)_{n\geq 1}$ is the walk which starts at zero and, at step i, independently steps either up or down by a_i with equal probability. We continue the study begun by Bhattacharya and Volkov in 2023 into the transience or recurrence of one-dimensional Rademacher random walks. In particular, we show that if the sequence of step sizes is bounded, the walk is weakly recurrent, meaning that it returns infinitely often to a random finite interval, while if the step sizes tend to infinity arbitrarily slowly the walk may be transient. On the other hand, we show that the step sizes may grow arbitrarily fast and still give a weakly recurrent random walk, and this is still true even if we restrict to non-decreasing step sizes. However, if $a_n = n^{\alpha+o(1)}$ for some $\alpha > 1/2$, we show that the walk is transient. We also show that the bound on α is tight by giving an example where $a_n = \Theta(n^{1/2})$ and the walk is weakly recurrent.

Keywords: inhomogeneous random walk, anti-concentration, modular Erdős–Littlewood–Offord inequality

1 Introduction

Throughout this paper, let $\epsilon_1, \epsilon_2, \dots$ be a sequence of i.i.d. Rademacher(1/2) random variables, taking the values ± 1 each with probability 1/2. We will refer to these simply as Rademacher random variables (dropping the 1/2). Let $(a_n)_{n\geq 1}$ be a deterministic sequence of non-negative real numbers, and define the associated *Rademacher random walk* $(X_n)_{n\geq 0}$ by

$$X_n = \sum_{k=1}^{n} \epsilon_k a_k.$$

*Centre for Mathematical Sciences, Lund University, Box 118 SE-22100, Lund, Sweden.

†School of Mathematics, University of Bristol, Bristol, BS8 1UG, UK and Heilbronn Institute for Mathematical Research, Bristol, UK.

Email addresses: satyaki.bhattacharya@matstat.lu.se, edward.crane@bristol.ac.uk, tom.johnston@bristol.ac.uk.

When all of the a_k take the value 1, the associated Rademacher random walk is exactly the simple symmetric random walk in one dimension, and the recurrence of this random walk is a fundamental result in any introductory probability course. However, by simply changing the sequence of step sizes (a_n), the problem becomes much less elementary and the behaviour of the Rademacher random walk is still unknown in many cases, including the simple-looking case where $a_n = n^\alpha$.

We mention that recurrence and transience of two-dimensional Rademacher random walks has been studied recently by Bhattacharya and Volkov in [4]. In this paper we will focus on the one-dimensional case, continuing the study begun by Bhattacharya and Volkov in [3].

If a particular sample of a Rademacher random walk satisfies $\{X_n \leq C\}$ infinitely often (i.o.), we call it *C-recurrent*. When X is almost surely 0-recurrent we say X is *recurrent*. Note that C-recurrence is not necessarily a tail event. Indeed, if we take the sequence $(a_n)_{n\geq 1}$ where $a_1 = 1$, $a_2 = 1$ and $a_i = 3$ for all $i \geq 2$, then whether the walk returns to zero i.o. depends on the value after two steps. However, the event $\{\exists C : |X_n| \leq C \ i.o.\}$ is a tail event, and we shall call a Rademacher random walk *weakly recurrent* if the probability of this event is 1. Otherwise, we will call the walk *transient*. Finally, we say that a random walk is *topologically recurrent* if its range is almost surely dense in $\mathbb{R}$.

Note that weak recurrence of the Rademacher random walk is unaffected by making any finite number of changes, deletions, or insertions to $(a_n)_{n\geq 1}$.

One interesting class of Rademacher random walks is those whose step sizes $(a_n)_{n\geq 1}$ satisfy $\sum_n a_n^2 < \infty$ but $\sum_n a_n = \infty$. In this case, X_n almost surely converges to a random limit by the $p = 2$ case of Doob's L^p-martingale convergence theorem. However, the distribution of the limit has unbounded support, so X is weakly recurrent but there is no $C < \infty$ such that X is almost surely C-recurrent.

Our main motivation is the following problem.

Problem 1. *What conditions on the growth rate of (a_n) guarantee that the associated Rademacher random walk is transient? What growth conditions guarantee that the Rademacher random walk is weakly recurrent?*

In Bhattacharya and Volkov [3] it is shown that the Rademacher random walk is C-recurrent for every $C > 0$, and hence weakly recurrent, in the following cases:

- $a_n = \log n$, and
- $a_n = \lfloor (\log_\gamma n)^\beta \rfloor$ for constants $\gamma > 1$ and $0 < \beta \leq 1$.

Two transient cases with integer step sizes are also given:

- the steps $(a_n)_{n\geq 1}$ are all distinct integers, and
- $a_n = \lfloor n^\beta \rfloor$ for any constant $0 < \beta < 1$.

One particularly natural case of Problem 1 is to assume that $a_n \approx n^\alpha$ for some $\alpha > 0$. With this constraint, we can ask whether there are values of α for which we can guarantee that the walk is transient or weakly recurrent. Combining the two transience results from [3], we see that the Rademacher walk with step sizes $a_n = \lfloor n^\alpha \rfloor$ is transient for all $\alpha > 0$. Hence, there is no $\alpha > 0$ for which the condition $a_n \approx n^\alpha$ implies the walk is weakly recurrent. In contrast, our main result shows that there are values of α for which this growth rate guarantees that the walk is transient.

Theorem 1. *Let $(a_n)_{n\geq 1}$ be a sequence and suppose that $a_n = n^{\alpha+o(1)}$ for some $\alpha > 1/2$. Then the associated Rademacher random walk is transient.*

The condition on α in Theorem 1 is best possible, even under the stronger assumption that $a_n = \Theta(n^\alpha)$ for some $\alpha > 0$ as $n \to \infty$: there is a sequence of step sizes $(a_n)_{n\geq 1}$ where $a_n = \Theta(n^{1/2})$ and the associated Rademacher random walk is weakly recurrent.

Theorem 2. *There exists a sequence $(a_n)_{n\geq 1}$ of integers such that $a_n = \Theta(n^{1/2})$ as $n \to \infty$ and the associated Rademacher random walk is weakly recurrent.*

We remark that from our proof of Theorem 1 we can already relax the condition that $a_n = n^{\alpha+o(1)}$ slightly to allow $n^{\alpha-\delta} \leq a_n \leq n^{\alpha+\delta}$ for some sufficiently small $\delta = \delta(\alpha)$. However, we still need both upper and lower bounds on a_n. Perhaps surprisingly, the upper bound is necessary: the sequence (a_n) can grow arbitrarily fast and still yield a recurrent Rademacher walk.

Theorem 3. *Let $f : \mathbb{N} \to \mathbb{R}$ be any non-decreasing function. There is an integer sequence $(a_n)_{n\geq 1}$ such that $a_n \geq f(n)$ for all n and the associated Rademacher random walk is recurrent.*

Our proof of Theorem 3 constructs the sequence in blocks of increasing length such that within each block the terms alternate between two consecutive integers. Allowing non-integer step sizes, we can get topological recurrence from a strictly increasing sequence that grows as fast as we like:

Theorem 4. *Let $f : \mathbb{N} \to \mathbb{R}$ be any non-decreasing function. There is a strictly increasing real sequence $(a_n)_{n\geq 1}$ such that $a_n \geq f(n)$ for all n and the associated Rademacher random walk is topologically recurrent.*

Recall that Bhattacharya and Volkov showed that if (a_n) consists of distinct integers then the Rademacher random walk is transient. In particular, the walk associated to any strictly increasing integer sequence (a_n) is transient, and there cannot be a single sequence that demonstrates both Theorem 3 and Theorem 4.

These results leave open the possibility that there could be a powerful sufficient condition for transience of the Rademacher random walk, in terms of the asymptotic behaviour of the sequence (a_n), when we restrict to non-decreasing integer sequences. The asymptotically fastest-growing non-decreasing integer sequences that we know to yield weakly recurrent Rademacher random walks are the examples $a_n = \lfloor c \log n \rfloor$ from [3]. On the other hand, increasing the growth rate slightly, but still using all the non-negative integers, we can obtain a transient Rademacher random walk.

Theorem 5. *Let $(a_n)_{n\geq 1}$ be a non-decreasing sequence of integers, and let L_i be the number of times i appears in the sequence. Suppose that, for all large enough n,*

$$\sum_{\substack{i\leq n\\ \gcd(i,n)=1}} L_i \geq 2n^2 \tag{1}$$

and

$$\sum_{i=1}^{n-1} i^2 L_i \geq 4n^2 \log^3(n) \cdot L_n. \tag{2}$$

Then the Rademacher random walk X associated with $(a_n)_{n\geq 1}$ is transient.

Corollary 6. *For any $\alpha > 1$, the Rademacher random walk with step sizes $a_n = \lfloor \log^\alpha(n) \rfloor$ is transient*

The (Lévy) *concentration function* Q of a real valued random variable A is defined by[1]

$$Q_r(A) = \sup_{x\in\mathbb{R}} \mathbb{P}(x < A \leq x + r).$$

Any upper bound for a value of the concentration function is called an *anti-concentration* bound, while a lower bound is called a *concentration* bound. Some known anti-concentration bounds for sums of independent random variables are discussed in Section 2.

To prove Theorem 1 we show, using the theorem below, that the position of the Rademacher random walk at each time n is sufficiently anti-concentrated, and then apply the Borel–Cantelli lemma. We remark that, as alluded to earlier, the following theorem allows us to slightly relax the condition that $a_n = n^{\alpha+o(1)}$.

Theorem 7. *Let $(a_n)_{n\geq 0}$ be a sequence and suppose that there are constants $c, C, \alpha > 0$ and $\delta \geq 0$ such that, for all large enough n,*

$$cn^\alpha \leq a_n \leq Cn^{\alpha+\delta}.$$

Then, for any $\gamma > 0$, we have the anti-concentration bound

$$Q_1\left(\sum_{i=1}^n \epsilon_i a_i\right) = O\Big(n^{-(\frac{1}{2}+\alpha f(\alpha,\delta)-\gamma)}\Big),$$

where

$$f(\alpha,\delta) = \begin{cases} \frac{\alpha^2}{(\alpha+\delta)(\alpha+2\delta+2\sqrt{\delta^2+\alpha\delta})} & \text{if } \delta \leq \frac{\sqrt{\alpha^2+1}-\alpha}{2}, \\ \frac{\alpha^2}{(\alpha+\delta)(1+2\delta)(\alpha+1/2+\delta)} & \text{if } \delta \geq \frac{\sqrt{\alpha^2+1}-\alpha}{2}. \end{cases}$$

From this it is easy to deduce Theorem 1.

[1]We warn the reader that authors disagree about the strictness of the inequalities in the definition of the concentration function, so care is needed in interpreting anti-concentration inequalities in the literature. We have made the same choice as in [10], because it gives $Q_1(A)$ the meaning that we want in the case of an integer-valued random variable A.

Proof of Theorem 1 from Theorem 7. Fix $\lambda > 0$ which is small enough that $1/2 + \alpha - \lambda > 1$. For any $\delta > 0$, we have

$$n^{\alpha-\delta} \le a_n \le n^{\alpha+\delta}$$

for all sufficiently large n. As $f(\alpha, \delta) \to 1$ as $\delta \to 0$, there is some $\delta > 0$ such that this condition on a_n is enough to get the anti-concentration bound $Q_1(X_n) = O(n^{-(1/2+\alpha-\lambda)})$.

For any fixed C, the probability that $|X_n| \le C$ is $O(n^{-(1/2+\alpha-\lambda)})$ and this is summable by our choice of λ. Hence, by the Borel–Cantelli lemma, the probability that $|X_n| \le C$ infinitely often is 0, and the walk is transient. $\square$

We remark that by substituting $\delta = 0$ into Theorem 7, we get the bound $O(n^{-(\alpha+1/2-\gamma)})$, which is easily seen to be tight up to the γ term by considering the sequence $a_n = n^\alpha$. Also of interest is what happens as $\delta \to \infty$. In this case, the anti-concentration gets close to $O(n^{-1/2})$, and it is also straightforward to show that there must be points where this is the correct behaviour. More generally, we have the following easy lower bounds which complement Theorem 7.

Proposition 8. *Fix $\alpha > 0$. Then*

$$Q_1\left(\sum_{i=1}^{n} \epsilon_i n^\alpha\right) = \Omega(n^{-(1/2+\alpha)}).$$

Moreover, for any $\delta \ge 1/2$, there exists a sequence $(a_n)_{n\ge1}$ of step sizes and a sequence of times $(n_i)_{i\ge1}$ such that

$$Q_1\left(\sum_{i=1}^{n_i} \epsilon_i a_i\right) \ge n_i^{-(1/2+\frac{\alpha}{2\delta}+o(1))}.$$

Let us now turn to the second part of Problem 1: are there conditions on the growth rate of (a_n) that guarantee that the associated Rademacher random walk is weakly recurrent?

There is no unbounded non-decreasing function f for which the condition $a_n \le f(n)$ suffices to imply weak recurrence, as the following example shows. Let $f : \mathbb{N} \to [1, \infty)$ be unbounded and non-decreasing. Take $a_n = 2^{b_n}$ where $b_n = \lfloor \log_2 f(n) \rfloor$ for all n, so that $a_n \le f(n)$. Once $b_n \ge k$, the congruence class of X_n modulo 2^k becomes constant and it is not difficult to deduce that the random walk is transient. In light of this example, one might ask for the walk to be "irreducible". We say an integer-valued random walk is *irreducible* if, for any $a, b \in \mathbb{Z}$ and any $n \in \mathbb{N}$ such that $\mathbb{P}(X_n = a) > 0$, there is some $m > n$ such that $\mathbb{P}(X_m = b \mid X_n = a) > 0$. Clearly, the Rademacher random walk that we just constructed is not irreducible. However, imposing irreducibility does not change the situation, as the following lemma shows.

Lemma 9. *Let $f : \mathbb{N} \to [1, \infty)$ be any function such that $f(n) \to \infty$ as $n \to \infty$. Then there exists a non-decreasing sequence $(a_n)_{n\ge1}$ of integer step sizes for which $a_n \le f(n)$ for all n and the associated Rademacher random walk is both irreducible and transient.*

To address the case where $(a_n)_{n\geq 1}$ is bounded, we prove the following lemma, which extends a theorem of Bhattacharya and Volkov [3, Theorem 2].

Lemma 10. *Let $(a_n)_{n\geq 1}$ be a sequence of real numbers such that $\sum_{n=1}^{\infty} a_n^2 = \infty$, and let $X = (X_n)_{n\geq 0}$ be the associated Rademacher random walk. Then, almost surely, X is unbounded below and unbounded above, and in particular, X changes sign infinitely often.*

In [3] it was shown that X changes sign infinitely often almost surely under the stronger assumption that (a_n) is a non-decreasing sequence of positive reals. Note that the condition $\sum_{n=1}^{\infty} a_n^2 = \infty$ is necessary, since otherwise X_n almost surely converges to a random limit. When X converges, it changes sign only finitely many times.

Lemma 10 implies that if the sequence (a_n) is bounded, say $a_n \leq C$ for all n, and $\sum_{n=1}^{\infty} a_n^2 = \infty$, then $\mathbb{P}(|X_n| < C \text{ i.o.}) = 1$, so the Rademacher random walk is weakly recurrent. Recall that if $\sum_{n=1}^{\infty} a_n^2$ converges, then the Rademacher random walk converges to a random limit and the walk is also weakly recurrent. Putting these two cases together gives the following corollary.

Corollary 11. *If the sequence $(a_n)_{n\geq 1}$ is bounded, the associated Rademacher random walk is weakly recurrent.*

We also have another immediate corollary of Lemma 10:

Corollary 12. *If the sequence $(a_n)_{n\geq 1}$ satisfies $\sum_{n=1}^{\infty} a_n^2 = \infty$ and $a_n \to 0$ as $n \to \infty$, then the associated Rademacher random walk is topologically recurrent.*

This is a special case of the well-known fact (which it easily implies by conditioning on the step sizes) that any symmetric random walk with bounded but not necessarily identically distributed steps is weakly recurrent.

Although we know that there is a weakly recurrent Rademacher walk with $a_n = \Theta(n^{1/2})$, we have not determined the behaviour in the natural case where $a_n = n^{\alpha}$ for any α in the range $0 < \alpha \leq 1/2$. We can, however, say something about the possible behaviours in this range by applying the following theorem.

Theorem 13. *Suppose the sequence (a_n) is unbounded and $a_n - a_{n-1} \to 0$. Then the associated Rademacher random walk is either transient or topologically recurrent.*

Theorem 1 shows that the transient case may occur, and the recurrence result of [3] concerning the sequence $a_n = \log n$ shows that the topologically recurrent case may occur.

The rest of the paper is organised as follows. In §2 we discuss a number of known anti-concentration results and prove a simple but useful lemma about combining anti-concentration at different scales. In §3 we prove some anti-concentration estimates for Rademacher sums, including a (mod m) analogue of the Erdős–Littlewood–Offord inequality (Theorem 15). In §4 we prove the anti-concentration bound Theorem 7, from which we have already deduced Theorem 1, and the concentration bound Proposition 8. In §5 we prove Theorem 2 by exhibiting an explicit sequence $(a_n)_{n\geq 1}$ such that $a_n = \Theta(n^{1/2})$ as $n \to \infty$ and the associated Rademacher walk is weakly recurrent. In §6 we prove Lemma 9 and Lemma 10. In §7 we prove Theorem 5 and deduce Corollary 6. In §8 we prove Theorem 13. Finally, in §9 we prove Theorem 3 and Theorem 4.

2 Anti-concentration inequalities

In this section we recall several well-known anti-concentration inequalities that will be used later on, and we give a simple lemma which allows us to combine the anti-concentration of two random variables at different scales.

Recall that the *concentration function* Q of a real-valued random variable A is given by

$$Q_r(A) = \sup_{x\in\mathbb{R}} \mathbb{P}(x < A \leq x + r).$$

Note that for any integer $m \geq 1$, the union bound gives

$$Q_{mr}(A) \leq mQ_r(A).$$

We start with the general case of a one-dimensional Rademacher random walk X whose step sizes are general real numbers, not necessarily integers and not necessarily separated.

Lemma 14 (Erdős–Littlewood–Offord inequality)**.** *If the step sizes $(a_n)_{n\geq 1}$ of a Rademacher walk $X = (X_n)_{n\geq 1}$ are all greater than or equal to a constant $c > 0$, then*

$$Q_{2c}(X_n) \leq \binom{n}{\lfloor n/2 \rfloor} 2^{-n} \sim \sqrt{\frac{2}{\pi n}} \text{ as } n \to \infty. \tag{3}$$

Erdős' simple proof of this in [7] was to note that for any $x \in \mathbb{R}$, the set of assignments of $(\epsilon_1, \dots, \epsilon_n)$ for which $\sum_{i=1}^n \epsilon_i a_i \in (x, x + c]$ form an anti-chain in the hypercube $\{-1, 1\}^n$ with the coordinatewise partial order, and then to apply Sperner's theorem [20] about the size of the largest anti-chain. In §3 we will prove the following (mod m) analogue of Lemma 14 for the case of integer step sizes, which we were unable to find in the literature.

Theorem 15. *Let m be a positive integer and let $b_1, \dots, b_n$ be positive integers coprime to m. Let $\epsilon_1, \dots, \epsilon_n$ be independent Rademacher random variables, and let $X_n = \sum_{i=1}^n \epsilon_i b_i$. Then*

$$\max_{r\in\mathbb{Z}/m\mathbb{Z}} \mathbb{P}(X_n \equiv r \pmod m) \leq \begin{cases} \frac{1}{m} + \sqrt{\frac{2}{\pi n}} & \text{if } m \text{ is odd,} \\ \frac{2}{m} + \sqrt{\frac{2}{\pi n}} & \text{if } m \text{ is even.} \end{cases}$$

The Erdős–Littlewood–Offord inequality is a special case of a more general result, known as the Kolmogorov–Rogozin inequality, which gives anti-concentration bounds on a sum of independent random variables using bounds on the anti-concentration of the summands. The first results of this form were shown by Doeblin and Levy, before being improved by Kolmogorov, then Rogozin and then Kesten. The sharpest possible bounds were obtained recently by Juškevičius in [10], which summarizes the history of the problem.

For many applications, the following form of the Kolmogorov–Rogozin inequality suffices.

Theorem 16 (Rogozin, [17])**.** *There is a $C > 0$ such that for any independent random variables $X_1, \ldots, X_n$ and real numbers $0 < \lambda_1, \ldots, \lambda_n \le 2r$,*

$$Q_r(X_1 + \cdots + X_n) \le C \cdot r \cdot \left(\sum_{i=1}^{n} \lambda_i^2 (1 - Q_{\lambda_i}(X_i)) \right)^{-1/2}.$$

Although we will not use it directly, we mention for completeness a key tool in proving many anti-concentration inequalities:

Lemma 17 (Esséen's inequality [8])**.** *If X is a real-valued random variable with characteristic function ψ, then $Q_r(X) \le t \int_{-2\pi/r}^{2\pi/r} |\psi(\lambda)| \, \mathrm{d}\lambda$.*

Let us now consider the case where the step sizes are separated. Let $Y = (Y_n)_{n \ge 0}$ be the Rademacher random walk with step sizes 1, 2, 3, etc. It is known that for any Rademacher random walk $(X_n)_{n \ge 0}$ with *distinct positive integer* step-sizes, we have the anti-concentration bound

$$Q_1(X_n) = \sup_{x \in \mathbb{Z}} \mathbb{P}(X_n = x) \le \mathbb{P}(Y_n = 0) \sim \frac{\sqrt{6/\pi}}{n^{3/2}} \text{ as } n \to \infty. \tag{4}$$

The inequality in (4) was proven by Richard Stanley [21], using enumerative algebraic geometry and answering a question of Erdős and Moser. Another wonderful proof using Lie algebras was given shortly afterwards by Proctor [16]. The asymptotic in (4) is due to Sullivan [23]. Before Stanley's result, Sárközi and Szemerédi [18] had shown that $\sup_{x \in \mathbb{Z}} \mathbb{P}(X_n = x) = O(n^{-3/2})$. From their bound it already follows by a simple Borel–Cantelli argument that any Rademacher random walk whose step sizes are distinct positive integers must be transient (see [3, Theorem 4]).

Halász extended the result of Sárközi and Szemerédi as follows.

Lemma 18 (Halász [9, Theorem 2])**.** *Consider n vectors $a_1, \ldots, a_n \in \mathbb{R}^d$ such that for any unit vector e, we have $|\langle e, a_i \rangle| \ge 1$ for at least δn values of i, and also $\|a_i - a_j\| \ge 1$ whenever $i \ne j$. Then*

$$\mathbb{P}\left(\sum_{i=1}^{n} \epsilon_i a_i \in B(x, 1) \right) \le c(\delta, d) n^{-1-d/2}, \tag{5}$$

where $c(\delta, d)$ is a constant that does not depend on n.

In the case $d = 1$, the inner product condition involving δ may be dropped since it is implied by the separation condition for large enough n. We remark that (5) already gives sufficient anti-concentration to show that the Rademacher random walk with step sizes $a_n = n^\alpha$ is transient for any $\alpha > 2/3$. Indeed, if

$\alpha \geq 1$, then $|a_i - a_j| \geq 1$ for all $i \neq j$, and we can immediately apply Lemma 18 in the case $d = 1$ to get an anti-concentration bound of $O(n^{-3/2})$ (and then we can finish by applying the Borel–Cantelli lemma). On the other hand, if $\alpha < 1$, then $a_{n+1} - a_n < 1$ and there is some a_i in each interval $[k, k+1)$. Hence, we can find a subsequence $a_{i_1}, a_{i_2}, \dots$ of length $\Theta(n^\alpha)$ such that $a_{i_j} \in [2j, 2j+1)$, and the theorem gives an anti-concentration bound of $O(n^{-3\alpha/2})$, which suffices to finish using Borel–Cantelli when $\alpha > 2/3$.

All of the bounds above deal with step sizes which are of a somewhat similar scale. If the step sizes grow exponentially with base at least 2, then $Q_1(X_n) = 1/2^n$ and there is much better anti-concentration than given by any of the results above. While we will not be dealing with scales quite as different as this, combining anti-concentration at different scales and (nearly) multiplying the anti-concentration bounds of each is the key to our proof. For this we use the following lemma.

Lemma 19 (Combining anti-concentration bounds at different scales)**.** *Let $0 < r < s$ and let A and B be independent real-valued random variables. Then*

$$Q_r(A + B) \leq \mathbb{P}(|A| \geq s) + 3Q_r(A)Q_s(B).$$

Hence,

$$Q_r(A + B) \leq (1 - Q_{2s}(A)) + 3Q_r(A)Q_s(B).$$

Proof. For any $x \in \mathbb{R}$ we have

$$\begin{aligned}
\mathbb{P}(x < A + B \leq x + r) &\leq \mathbb{P}(|A| \geq s) + \mathbb{P}(|A| \leq s \text{ and } x < A + B \leq x + r) \\
&\leq \mathbb{P}(|A| \geq s) + \mathbb{P}(x - s < B \leq x + r + s) \cdot \sup_{b \in \mathbb{R}} \mathbb{P}(x - b < A \leq x - b + r) \\
&\leq \mathbb{P}(|A| \geq s) + 3Q_s(B)Q_r(A).
\end{aligned}$$

Since we may replace A by $A + c$ for any constant c without changing $Q_r(A)$ or $Q_r(A + B)$, the final statement follows. □

3 Modular anti-concentration of Rademacher sums

In the course of several of our proofs, we will need to control the probability that a sum of Rademacher random variables takes a particular value, or lies in a particular congruence class modulo some positive integer. In this section we establish some useful results of these kinds.

For a real random variable X, the *characteristic function* of X is the function $\psi_X : \mathbb{R} \to \mathbb{C}$ given by $\psi_X(t) = \mathbb{E}(e^{itX})$. We say that a real random variable X is *monotone* if $|\psi_X(t)|$ is decreasing on $[0, \pi]$.

Theorem 20 ([1, Theorem 1.1])**.** *Let $X_1, \dots, X_k$ be independent integer-valued random variables with $\mathbb{E}(X_i) = \mu_i$ and $\mathrm{Var}(X_i) = \sigma_i^2 < \infty$. Suppose that their sum $X = X_1 + \cdots X_k$ is a monotone random variable with mean μ and variance σ^2. Then, for every t for which $\mu + t\sigma$ is an integer, we have*

$$\left| \mathbb{P}(X = \mu + t\sigma) - \frac{1}{\sqrt{2\pi\sigma^2}} e^{-t^2/2} \right| \leq c \left(\frac{\sum_{i=1}^{k} \mathbb{E}(|X_i - \mu_i|^3)}{\sigma^3} \right)^2,$$

where c is a universal constant.

Note that the sum of independent monotone random variables is monotone and that Bernoulli(1/2) random variables are monotone. Hence, by applying the above theorem to a suitably shifted and rescaled sum, we obtain the following local limit theorem for sums of Rademacher random variables.

Corollary 21. *Let X be the sum of n Rademacher random variables. Then, for any $x \equiv n \mod 2$, we have*

$$\left| \mathbb{P}(X = x) - \frac{1}{\sqrt{\pi n/2}} e^{-\frac{x^2}{2n}} \right| \leq \frac{c}{n},$$

where c is an absolute constant.

In the remainder of this section we establish Theorem 15, which we restate here.

Theorem 15. *Let m be a positive integer and let $b_1, \ldots, b_n$ be positive integers coprime to m. Let $\epsilon_1, \ldots, \epsilon_n$ be independent Rademacher random variables, and let $X_n = \sum_{i=1}^n \epsilon_i b_i$. Then*

$$\max_{r \in \mathbb{Z}/m\mathbb{Z}} \mathbb{P}(X_n \equiv r \pmod m) \leq \begin{cases} \frac{1}{m} + \sqrt{\frac{2}{\pi n}} & \text{if } m \text{ is odd,} \\ \frac{2}{m} + \sqrt{\frac{2}{\pi n}} & \text{if } m \text{ is even.} \end{cases}$$

Before giving the proof, let us make a few observations about this result.

Fixing arbitrary positive integers $b_1, \ldots, b_n$ and letting $m \to \infty$ along the primes gives an alternative proof of the integer case of the Erdős–Littlewood–Offord inequality, Lemma 14.

On the other hand, taking each $b_i = 1$ in gives a simpler modular anti-concentration result that is still useful.

Corollary 22. *Let $m \geq 2$ be an integer and let $X_n = \sum_{i=1}^n \epsilon_i$ be the sum of n independent Rademacher random variables. For every $\varepsilon > 0$, if $n \geq \varepsilon m^2$, we have*

$$\max_{0 \leq r < m} \mathbb{P}(X \equiv r \pmod m) < c(\varepsilon)/m,$$

where $c(\varepsilon) = 1 + \sqrt{\frac{2}{\pi\varepsilon}}$.

Remark 23. *For the case where $n \leq m^2$ and the obvious parity condition is satisfied, a complementary lower bound for $\mathbb{P}(X \equiv r \pmod m)$, also of order $1/m$, may be found in [3, Corollary 3.2].*

If p is an odd prime, then a p-adic formal analogue of the Erdős–Littlewood–Offord inequality follows from Theorem 15, as follows. If $b_1, \ldots, b_n \in \mathbb{Q}_p$ all satisfy $\|b_i\|_p \geq c$, then for any $z \in \mathbb{Q}_p$ we have

$$\mathbb{P}(X_n \in B(z, c)) \leq \frac{1}{p} + \sqrt{\frac{2}{\pi n}},$$

where $B(z, c)$ here means the open ball in $\mathbb{Q}_p$ of p-adic radius c about z.

In the proof of Theorem 15, we will need the following easy consequence of the rearrangement inequality.

Lemma 24. *Let $n, m \geq 1$ and let $y_1, \ldots, y_m$ be given positive numbers. Consider the set $\mathcal{A}$ of n by m matrices in which the elements of each row are some permutation of $(y_1, \ldots, y_m)$. Then*

$$\max_{A \in \mathcal{A}} \sum_{j=1}^{m} \prod_{i=1}^{m} A_{ij} = \sum_{j=1}^{m} y_j^n.$$

Proof. The set $\mathcal{A}$ is finite, so there exists a matrix $A \in \mathcal{A}$ which maximizes $\sum_{j=1}^m \prod_{i=1}^m A_{ij}$. For this A, the terms in each row must be sorted in the same as order as the products over the columns after that row is deleted, by the rearrangement inequality. Hence, they are sorted in the same way as the overall column products, which means that they are all sorted in the same way. $\square$

We now turn to the proof of Theorem 15.

Proof of Theorem 15. For each $x \in \mathbb{Z}/m\mathbb{Z}$, let $\delta_x : \mathbb{Z}/m\mathbb{Z} \to \mathbb{C}$ be the indicator function of x, and for each $\lambda \in \mathbb{Z}/m\mathbb{Z}$, let $e_\lambda : \mathbb{Z}/m\mathbb{Z} \to \mathbb{C}$ be the function $e_\lambda(x) = \exp(2\pi i \lambda x/m)/\sqrt{m}$. Note that the functions e_λ form an orthonormal basis of $\ell^2(\mathbb{Z}/m\mathbb{Z})$. We have

$$\delta_0 = \sum_{\lambda \in \mathbb{Z}/m\mathbb{Z}} \frac{1}{\sqrt{m}} e_\lambda.$$

A calculation shows that the effect on e_λ of convolution with $\frac{1}{2}(\delta_b + \delta_{-b})$ is to multiply it by $\cos(2\pi b\lambda/m)$. Hence, the probability mass function of X_n is given by

$$\mathbb{P}(X_n \equiv r \pmod{m}) = \sum_{\lambda \in \mathbb{Z}/m\mathbb{Z}} \frac{1}{\sqrt{m}} \prod_{i=1}^{n} \cos(2\pi b_i \lambda/m) e_\lambda(r)$$

Since $|e_\lambda(r)| = \frac{1}{\sqrt{m}}$, the triangle inequality gives that

$$\mathbb{P}(X_n \equiv r \pmod{m}) \leq \frac{1}{m} \sum_{\lambda \in \mathbb{Z}/m\mathbb{Z}} \prod_{i=1}^{n} |\cos(2\pi b_i \lambda/m)|. \tag{6}$$

The multi-set of values $(\cos(2\pi b\lambda/m) : \lambda \in \mathbb{Z}/m\mathbb{Z})$ is the same for every integer b coprime to m. Therefore, by Lemma 24, the right-hand side of (6) is maximized when all b_i are congruent (mod m) to any constant value b that is coprime to m, for example $b = 1$. In this case it takes the value $\frac{1}{m} \sum_{\lambda \in \mathbb{Z}/m\mathbb{Z}} |\cos(2\pi\lambda/m)|^n$. In the case where m is even,

$$\frac{1}{m} \sum_{\lambda \in \mathbb{Z}/m\mathbb{Z}} |\cos(2\pi\lambda/m)|^n = \frac{1}{m} \left(2 + 4 \sum_{\ell=1}^{\lfloor m/4 \rfloor} \cos(2\pi\ell/m)^n \right).$$

In the case where m is odd,

$$\frac{1}{m}\sum_{\lambda\in\mathbb{Z}/m\mathbb{Z}}|\cos(2\pi\lambda/m)|^n = \frac{1}{m}\left(1+2\sum_{\ell=1}^{\lfloor m/2\rfloor}\cos(\pi\ell/m)^n\right).$$

Using the inequality $\cos x \le e^{-x^2/2}$, which holds for $x \in [-\pi/2, \pi/2]$, for even m we obtain

$$\begin{aligned}\frac{1}{m}\sum_{\lambda\in\mathbb{Z}/m\mathbb{Z}}|\cos(2\pi\lambda/m)|^n &\le \frac{1}{m}\left(2+4\sum_{\ell=1}^{\infty}e^{-2\pi^2\ell^2n/m^2}\right)\\ &\le \frac{1}{m}\left(2+4\int_0^\infty e^{-(4\pi^2n/m^2)\cdot x^2/2}\,\mathrm{d}x\right)\\ &= \frac{2}{m}+\sqrt{\frac{2}{\pi n}},\end{aligned}$$

and similarly for odd m we obtain

$$\begin{aligned}\frac{1}{m}\sum_{\lambda\in\mathbb{Z}/m\mathbb{Z}}|\cos(2\pi\lambda/m)|^n &\le \frac{1}{m}\left(1+2\sum_{\ell=1}^{\infty}e^{-\pi^2\ell^2n/(2m^2)}\right)\\ &\le \frac{1}{m}+\sqrt{\frac{2}{\pi n}}.\end{aligned}$$

□

4 Anti-concentration of Rademacher random walks

The main aims of this section are to prove the anti-concentration bound Theorem 7 and the concentration bound Proposition 8.

Our main tool in proving Theorem 7 is Lemma 19, for which we need both anti-concentration and concentration bounds on A. For the concentration bound, we will use the following well-known bound (see e.g. [14]), which is a straightforward application of Hoeffding's inequality.

Lemma 25.

$$\mathbb{P}\left(\sum_{i=1}^{n}a_i\epsilon_i \ge t\|a\|_2\right) \le e^{-t^2/2}.$$

The result in Theorem 7 only requires the bounds on a_n to hold for large enough n, and there is no control over the initial terms. The following result shows that modifying a prefix of length m of the sequence (a_n) can only change the concentration by a factor of at most 2^{m+1}, and in particular, this will allow us to assume that the bounds on a_n hold for all n when proving Theorem 7.

Lemma 26. *Let $(a_n)_{n\ge1}$ be a sequence and let $(a'_n)_{n\ge1}$ be a sequence which differs from (a_n) only in the first m terms. Let $(X_n)_{n\ge0}$ and $(X'_n)_{n\ge0}$ be the corresponding Rademacher random walks. Then, for every n,*

$$2^{-(m+1)}Q_r(X'_n) \le Q_r(X_n) \le 2^{m+1}Q_r(X'_n).$$

Proof. Suppose that A and B are discrete independent random variables. We claim that

$$\frac{Q_r(A)Q_r(B)}{2} \le Q_r(A+B) \le Q_r(A). \tag{7}$$

For the upper bound we have

$$\begin{aligned}
\mathbb{P}(x < A + B \le x + r) &= \sum_b \mathbb{P}(x - b < A \le x - b + r)\mathbb{P}(B = b) \\
&\le \sup_b \mathbb{P}(x - b < A \le x - b + r) \cdot \sum_b \mathbb{P}(B = b) \\
&\le \sup_x \mathbb{P}(x < A \le x + r) \\
&= Q_r(A),
\end{aligned}$$

which implies $Q_r(A + B) \le Q_r(A)$.

For the lower bound, let $\varepsilon > 0$ be given and choose x, y such that

$$\begin{aligned}
&\mathbb{P}(x < A \le x + r) \ge (1 - \varepsilon)Q_r(A), \\
&\mathbb{P}(y < B \le y + r) \ge (1 - \varepsilon)Q_r(B).
\end{aligned}$$

Then the events

$$\{A + B \in (x + y, x + y + r]\} \quad \text{and} \quad \{A + B \in (x + y + r, x + y + 2r]\}$$

are disjoint and cover the event

$$\{x < A \le x + r,\ y < B \le y + r\}.$$

The latter event has probability at least $(1 - \varepsilon)^2 Q_r(A)Q_r(B)$ and so one of the first two events must have probability at least $(1 - \varepsilon)^2 Q_r(A)Q_r(B)/2$. Letting $\varepsilon \to 0$ we obtain the lower bound in (7).

Now let $A = \sum_{i=1}^m \epsilon_i a_i$ and $B = \sum_{i=m+1}^n \epsilon_i a_i$. Define A' by $A' = \sum_{i=1}^m \epsilon_i a_i'$, so that $X_n = A + B$ and $X_n' = A' + B$. Note that $Q_1(A), Q_1(A') \ge 2^{-m}$. We have

$$Q_r(A + B) \le Q_r(B) \le 2^{m+1}\frac{Q_r(A')Q_r(B)}{2} \le 2^{m+1}Q_r(A' + B)$$

and

$$Q_r(A + B) \ge \frac{Q_r(A)Q_r(B)}{2} \ge \frac{Q_r(A)Q_r(A' + B)}{2} \ge 2^{-(m+1)}Q_r(A' + B).$$

□

We are now armed with all the tools we need to prove Theorem 7, which we restate here for convenience.

Theorem 7. *Let $(a_n)_{n\ge 0}$ be a sequence and suppose that there are constants $c, C, \alpha > 0$ and $\delta \ge 0$ such that, for all large enough n,*

$$cn^\alpha \le a_n \le Cn^{\alpha+\delta}.$$

Then, for any $\gamma > 0$, we have the anti-concentration bound

$$Q_1\left(\sum_{i=1}^{n} \epsilon_i a_i\right) = O\left(n^{-(\frac{1}{2}+\alpha f(\alpha,\delta)-\gamma)}\right),$$

where

$$f(\alpha,\delta) = \begin{cases} \frac{\alpha^2}{(\alpha+\delta)(\alpha+2\delta+2\sqrt{\delta^2+\alpha\delta})} & \text{if } \delta \leq \frac{\sqrt{\alpha^2+1}-\alpha}{2}, \\ \frac{\alpha^2}{(\alpha+\delta)(1+2\delta)(\alpha+1/2+\delta)} & \text{if } \delta \geq \frac{\sqrt{\alpha^2+1}-\alpha}{2}. \end{cases}$$

Before we give the full details, let us briefly sketch the proof in the case where $a_n = n^\alpha$. The key idea is to group some of the terms in the sum $\sum_{i=1}^{n} \epsilon_i a_i$ into subsets which are at different scales. In the interval $[2^k, 2^{k+\alpha})$ we expect around $2^{k/\alpha}$ of the a_n, and let us set A to be the Rademacher sum of any such terms. Using the Erdős–Littlewood–Offord inequality, the anti-concentration of A at the scale 2^{k+1} is around $2^{-k/(2\alpha)}$. On the other hand, if k is much larger than α, we expect the magnitude of A to be around $2^k \cdot 2^{k/(2\alpha)}$, and there should be concentration around this scale. Hence, if we set $k' = (1 + 1/(2\alpha) + \varepsilon)k$ and let A' be the Rademacher sum of the terms in the interval $[2^{k'}, 2^{k'+\alpha})$, then the sum A' is anti-concentrated at the scale $2^{k'}$ while the sum A_k is concentrated at the $2^{k'}$. This means we can use Lemma 19 to combine the anti-concentration.

This can clearly be repeated and we will take a sequence $k_1, k_2, \dots$ where $k_i = (1 + 1/(2\alpha) + \varepsilon)^i$ and try to combine the anti-concentration of the random variables $A_1, A_2, \dots$, where we think of A_i as the sum of the terms $\epsilon_i a_i$ for which $a_i \in [2^{k_i}, 2^{k_i+\alpha})$. In actuality, we will have to widen the interval in which we take the a_n and also limit the number that we take from each interval.

To get the required anti-concentration for all large enough n, the anti-concentration from sums of the form $A_1 + \cdots + A_m$ is not enough. Indeed, while we get the required anti-concentration of $n^{-(\alpha+1/2-\gamma)}$ when n is not much bigger than $2^{k_m/\alpha} + 1$, the anti-concentration drops to around $n^{-\alpha}$ by the time n is close to $2^{k_{m+1}/\alpha}$. In order to apply Lemma 19 we need $A_1 + \cdots + A_{i-1}$ to be concentrated at the scale of A_i for each i. We can achieve this by ensuring that none of the terms in A_j are too big for each $j < m$, but observe that we don't need A_m to be concentrated as it is the last summand. To get the required concentration for all n between $2^{k_m/\alpha+1}$ and $2^{k_{m+1}/\alpha+1}$, we will replace A_m by the sum of the terms $\epsilon_i a_i$ for which $a_i \in [2^{k_m}, \infty)$ and $i \leq n$, which we can do as we do not need concentration for the last summand.

We make this argument rigorous below. We also weaken the conditions on a_n. This means we must make our intervals wider to be able to guarantee that we can find enough a_n which fall in the intervals. This in turn means that the intervals must be further apart and we pay a penalty in the anti-concentration that we obtain. However, this is unavoidable; there must be some penalty to pay (for large δ) as shown by Proposition 8.

Proof. Pick a constant λ such that $\delta/\alpha < \lambda$. We will later take λ to be arbitrarily close to δ/α.

By Lemma 26, we can modify the first m terms in the sequence and only change the anti-concentration by at worst a factor of 2^{m+1}. We will therefore assume that

$$cn^\alpha \le a_n \le Cn^{\alpha+\delta}$$

holds for all $n \ge 1$ and not just n large enough. Using our assumptions on a_n, if n is such that

$$(2^k/c)^{1/\alpha} \le n \le (2^{(1+\lambda)k}/C)^{1/(\alpha+\delta)}$$

then $2^k \le a_n \le 2^{(1+\lambda)k}$. Hence, the number of n for which a_n is in the interval $[2^k, 2^{(1+\lambda)k}]$ is at least

$$\left(\frac{2^{(1+\lambda)k}}{C}\right)^{\frac{1}{\alpha+\delta}} - 1 - \left(\frac{2^k}{c}\right)^{1/\alpha} = \left(\frac{2^{\frac{\alpha\lambda-\delta}{\alpha(\alpha+\delta)}k}}{C^{1/(\alpha+\delta)}} - \frac{1}{c^{1/\alpha}}\right)2^{k/\alpha} - 1.$$

Since we have taken $\alpha\lambda > \delta$, this certainly is at least $2^{k/\alpha}$ for large enough k.

Now fix $\beta \ge \alpha$ and $\varepsilon > \lambda$, and define k_i by

$$k_i = \left(1 + \frac{1}{2\beta} + \varepsilon\right)^i.$$

We will later take ε arbitrarily close to λ and choose the value of $\beta \ge \alpha$ to optimise the anti-concentration bound. Let A_i be the random sum corresponding to the first $\lceil 2^{k_i/\beta} \rceil$ of the a_n which lie in the interval $[2^{k_i}, 2^{(1+\lambda)k_i}]$, noting that this is possible when k_i is large by the above argument and our assumption that $\beta \ge \alpha$.

We now claim that the anti-concentration of the sum of the A_i is roughly equal to the product of their individual anti-concentrations.

Claim 27. *There is $M > 0$ such that*

$$Q_1(A_1 + \cdots + A_m) \le M \cdot 3^m \cdot 2^{-\frac{(1+1/(2\beta)+\varepsilon)^{m+1}}{1+2\beta\varepsilon}}$$

for all m.

Proof. First, note that by the Erdős–Littlewood–Offord inequality (Lemma 14), we have

$$Q_{2^{k_i+1}}(A_i) \le \sqrt{\frac{2}{\pi}} \cdot 2^{-\frac{k_i}{2\beta}}$$

for large enough i. We will induct on m, taking $s = 2^{k_m+1}$ and $r = 1$ in Lemma 19 to get anti-concentration at different scales.

To apply Lemma 19, we first need to show that the sum $A_1 + \cdots + A_{m-1}$ is concentrated at the scale s. Using Lemma 25, we have that

$$\begin{aligned}\mathbb{P}(|A_1 + \cdots + A_{m-1}| \ge s) &\le \sum_{i=1}^{m-1} \mathbb{P}\left(|A_i| \ge \frac{s}{(m-1)}\right) \\ &\le 2(m-1)\exp\left(-\frac{s^2}{4(m-1)^2 2^{(2+2\lambda+1/\beta)k_{m-1}}}\right).\end{aligned}$$

Substituting in $s = 2^{k_m+1} = 2 \cdot 2^{(1+\frac{1}{2\beta}+\varepsilon)k_{m-1}}$, we have

$$\mathbb{P}\big(|A_1 + \cdots + A_{m-1}| \ge 2^{k_m+1}\big) \le 2(m-1)\exp\left(-\frac{2^{2(\varepsilon-\lambda)k_{m-1}}}{(m-1)^2}\right).$$

Pick m_0 such that for all $m \geq m_0$, the quantity given above is at most $2^{-\frac{(1+1/(2\beta)+\varepsilon)^{m+1}}{1+2\beta\varepsilon}}$, and choose $M \geq 1$ such that

$$Q_1(A_1 + \cdots + A_m) \leq M \cdot 3^m \cdot 2^{-\frac{(1+1/(2\beta)+\varepsilon)^{m+1}}{1+2\beta\varepsilon}}$$

holds for all $m \leq m_0$.

Now suppose that the claimed bound holds for $m-1 \geq m_0$. Then, using the above, we have

$$\begin{aligned}
Q_1(A_1 + \cdots + A_m) &\leq 2^{-\frac{(1+1/(2\beta)+\varepsilon)^{m+1}}{1+2\beta\varepsilon}} + 3 \cdot M 3^{m-1} 2^{-\frac{(1+1/(2\beta)+\varepsilon)^{m}}{1+2\beta\varepsilon}} \cdot \sqrt{\frac{2}{\pi}} 2^{-\frac{k_m}{2\beta}} \\
&= 2^{-\frac{(1+1/(2\beta)+\varepsilon)^{m+1}}{1+2\beta\varepsilon}} + M \cdot 3^m \cdot \sqrt{\frac{2}{\pi}} \cdot 2^{-\frac{(1+1/(2\beta)+\varepsilon)^{m+1}}{1+2\beta\varepsilon}} \\
&= \left(1 + M \cdot 3^m \cdot \sqrt{\frac{2}{\pi}}\right) 2^{-\frac{(1+1/(2\beta)+\varepsilon)^{m+1}}{1+2\beta\varepsilon}} \\
&\leq M \cdot 3^m \cdot 2^{-\frac{(1+1/(2\beta)+\varepsilon)^{m+1}}{1+2\beta\varepsilon}}.
\end{aligned}$$

□

Pick N and suppose that $2(2^{(1+\lambda)k_m}/c)^{1/\alpha} \leq N < 2(2^{(1+\lambda)k_{m+1}}/c)^{1/\alpha}$. Let us assume that N is large enough that $m \geq m_0$, where m_0 is defined as in the claim above. Note that $a_n > 2^{(1+\lambda)k_m}$ for all $n \geq N$ so, in particular, the sum X_N contains all of the terms in the sum $A_1 + \cdots + A_m$. Let B be the Rademacher sum of the a_n for which $n \leq N$ and $a_n \geq 2^{k_m}$, so that the terms in this sum are a superset of the terms in A_m. We will bound the anti-concentration of $A_1 + \cdots + A_{m-1} + B$ using Lemma 19.

First, note that B contains at least

$$N - \left(\frac{2^{k_m}}{c}\right)^{1/\alpha} \geq \frac{N}{2}$$

terms, all of which are at least 2^{k_m}. By the Erdős–Littlewood–Offord inequality, Lemma 14, we have

$$Q_{2^{k_m+1}}(B) \leq \sqrt{\frac{4}{\pi N}}.$$

We again have the concentration bound

$$\mathbb{P}\big(|A_1 + \cdots + A_{m-1}| \geq 2^{k_m+1}\big) \leq 2(m-1)\exp\left(-\frac{2^{2(\varepsilon-\lambda)k_{m-1}}}{(m-1)^2}\right),$$

but this time we bound it directly. Since $m = \Theta(\log\log(N))$ and this term is $2^{-2^{2^{\Theta(m)}}}$, it decays quicker than any polynomial. The term 3^m is also at most polylogarithmic in N, so is certainly in $O(N^{\gamma/2})$ for any fixed $\gamma > 0$.

We also have that

$$\begin{aligned}
2^{-\frac{(1+1/(2\beta)+\varepsilon)^{m}}{1+2\beta\varepsilon}} &= 2^{-\frac{(1+\lambda)(1+1/(2\beta)+\varepsilon)^{m+1}}{\alpha} \cdot \frac{\alpha}{(1+2\beta\varepsilon)(1+\lambda)(1+1/(2\beta)+\varepsilon)}} \\
&\leq \left(\frac{c^{1/\alpha}N}{2}\right)^{-\frac{\alpha}{(1+2\beta\varepsilon)(1+\lambda)(1+1/(2\beta)+\varepsilon)}}.
\end{aligned}$$

Using Lemma 19, we find that $Q_1(A_1 + \cdots + A_{m-1} + B)$ is bounded by

$$M \cdot O(N^{\gamma/2}) \cdot \left(\frac{c^{1/\alpha}N}{2}\right)^{-\frac{\alpha}{(1+2\beta\varepsilon)(1+\lambda)(1+1/(2\beta)+\varepsilon)}} \cdot \sqrt{\frac{4}{\pi N}} + O\left(2^{-2^{2^{\Theta(\log\log(N))}}}\right)$$

which is

$$O\left(N^{-1/2-\frac{\alpha}{(1+2\beta\varepsilon)(1+\lambda)(1+1/(2\beta)+\varepsilon)}+\gamma/2}\right).$$

It remains to choose the values λ, ε and β, and we recall that we require $\alpha\lambda > \delta$, $\varepsilon > \lambda$ and $\beta \geq \alpha$. If $\delta < \frac{\sqrt{\alpha^2+1}-\alpha}{2}$, then we take $\beta = 1/(2\sqrt{\lambda^2+\lambda})$, else we take $\beta = \alpha$. We then take λ sufficiently close to δ/α and ε sufficiently close to λ so that $\beta \geq \alpha$ and

$$\frac{\alpha}{(1+2\beta\varepsilon)(1+\lambda)(1+1/(2\beta)+\varepsilon)} \geq \alpha f(\alpha,\delta) - \frac{\gamma}{2}.$$

□

From this it is easy to deduce the following corollary.

Corollary 28. *Let $(a_n)_{n\geq 1}$ be a sequence such that $a_n = n^{\alpha+o(1)}$ for some constant $\alpha > 0$, and let $(X_n)_{n\geq 0}$ be the associated Rademacher random walk. Then $Q_1(X_n) = O(n^{-(\alpha+1/2+o(1))})$.*

We end this section by proving Proposition 8, which we restate here for convenience.

Proposition 8. *Fix $\alpha > 0$. Then*

$$Q_1\left(\sum_{i=1}^{n} \epsilon_i n^{\alpha}\right) = \Omega(n^{-(1/2+\alpha)}).$$

Moreover, for any $\delta \geq 1/2$, there exists a sequence $(a_n)_{n\geq 1}$ of step sizes and a sequence of times $(n_i)_{i\geq 1}$ such that

$$Q_1\left(\sum_{i=1}^{n_i} \epsilon_i a_i\right) \geq n_i^{-(1/2+\frac{\alpha}{2\delta}+o(1))}.$$

Proof. First, let us consider the sequence $(a_n)_{n\geq 1}$ where $a_n = n^{\alpha}$, and let $(X_n)_{n\geq 1}$ be the associated Rademacher random walk. Chebyshev's inequality gives that

$$\mathbb{P}\left(|X_n| < 2\sqrt{\mathrm{Var}(X_n)}\right) \geq \frac{3}{4}.$$

The interval $(-2\sqrt{\mathrm{Var}(X_n)}, 2\sqrt{\mathrm{Var}(X_n)})$ may be covered using no more than $\lceil 4\sqrt{\mathrm{Var}(X_n)}\rceil$ intervals of the form $[x, x+1)$ so there exists $x \in \mathbb{R}$ such that

$$\mathbb{P}(X_n \in [x, x+1)) \geq \frac{3}{16\lceil\sqrt{\mathrm{Var}(X_n)}\rceil}.$$

The bound now follows by substituting in

$$\mathrm{Var}(X_n) = \sum_{k=1}^{n} (k^\alpha)^2 = \frac{n^{2\alpha+1}}{2\alpha+1} + O(n^{2\alpha}).$$

Now set $r = \delta/\alpha$ and consider the sequence where $a_n = 2^{(1+r)^k}$ for $2^{\frac{(1+r)^{k-1}}{\alpha}} \leq n < 2^{\frac{(1+r)^k}{\alpha}}$. Note that we have $n^\alpha \leq a_n \leq n^{\alpha+\delta}$.

Let A_k be the Rademacher sum of the a_n which are equal to $2^{(1+r)^k}$. The most likely value for A_k has probability

$$\binom{N}{\lfloor N/2 \rfloor} 2^{-N} \geq \sqrt{\frac{1}{\pi N}}$$

where $N \geq 2$ is the number of terms equal to $2^{(1+r)^k}$. We have that $N \leq 2^{\frac{(1+r)^k}{\alpha}}$, so the most likely value for A_k has probability at least $2^{-\frac{(1+r)^k}{2\alpha}}/\sqrt{\pi}$. Hence, the most likely value for $A_1 + \cdots + A_k$ has probability at least

$$\frac{2^{-\frac{((1+r)^k-1)(r+1)}{2r\alpha}}}{\pi^{k/2}} \geq \frac{2^{-\frac{(1+r)^{k+1}}{2r\alpha}}}{\pi^{k/2}}.$$

Hence, if we take $n_k = \lceil 2^{\frac{(1+r)^k}{\alpha}} \rceil - 1$ to the last step with step size $2^{(1+r)^k}$, then the most likely value has probability at least $n^{-(1/2+1/(2r))}/\pi^{k/2}$, as required. $\square$

5 A recurrent example with $a_n = \Theta(n^{1/2})$

Consider the sequence

$$(a_n)_{n \geq 1} = (3, 1, 5, 3, 5, 3, 5, 3, 5, 3, 9, 7, 9, 7, 9, 7, 9, 7, 9, 7, \dots)$$

which is made up of consecutive blocks, where the k^{th} block has length $4^k/2$ and the steps in the k^{th} block alternate between $2^k + 1$ and $2^k - 1$, starting with $2^k + 1$. Denote the index of the beginning of the $(k)^{th}$ block by n_k. That is,

$$n_1 = 1,\ n_2 = 3,\ n_3 = 11,\ \dots,\ n_k = 1 + \sum_{i=1}^{k-1} 4^i/2 = (4^k + 2)/6.$$

It is not hard to see that $\sqrt{n/2} \leq a_n \leq 3\sqrt{n}$ for all n, and so $a_n = \Theta(\sqrt{n})$. We will prove that the Rademacher random walk associated with this sequence is weakly recurrent (and hence prove Theorem 2). In fact we will show that the walk visits every even integer infinitely often.

Let $X = (X_n)_{n \geq 0}$ be the associated Rademacher random walk. Note that all the steps of X are odd integers, so to study the return times of X to 0 we may focus our attention on the random walk Y defined by $Y_n = X_{2n}$, which only visits even integers. We will use hitting probability estimates for the simple symmetric random walk on $\mathbb{Z}^2$ and the Kochen–Stone theorem to show that almost surely Y visits every even integer infinitely often. The Kochen–Stone theorem has been used before to prove recurrence of random walks; see e.g. [2, 15].

Theorem (Kochen and Stone [12, Theorem 1]). *Let $Z_1, Z_2, \dots$ be a sequence of random variables, each of which has nonzero mean and positive finite second moment. Suppose in addition that $\limsup_{n\to\infty}(\mathbb{E}(Z_n))^2/\mathbb{E}(Z_n^2) > 0$. Then*

(i) $\mathbb{P}(\liminf_{n\to\infty} Z_n/\mathbb{E}(Z_n) \le 1) > 0$,

(ii) $\mathbb{P}(\limsup_{n\to\infty} Z_n/\mathbb{E}(Z_n) \ge 1) > 0$, *and*

(iii) $\mathbb{P}(\limsup_{n\to\infty} Z_n/\mathbb{E}(Z_n) > 0) \ge \limsup_{n\to\infty}(\mathbb{E}(Z_n))^2/\mathbb{E}(Z_n^2)$.

Let E_k be the event that X visits 0 during the $(2k)^{th}$ block. We will show the following:

Lemma 29. *For E_k as defined above, we have $\mathbb{P}(E_k) = \Omega(1/k)$ as $k \to \infty$, so that $\sum_{k=1}^{\infty} \mathbb{P}(E_k) = \infty$. Moreover, there is a finite constant C such that for all $j < k$ we have*

$$\mathbb{P}(E_k \mid E_j) \le C\mathbb{P}(E_k).$$

Before proving Lemma 29, let us explain how it implies our claim about the recurrence of the walk Y. Let $Z_n = \sum_{k=1}^{n} \mathbb{1}(E_k)$. From the first statement in Lemma 29 we deduce that $\mathbb{E}(Z_n) = \Omega(\log n)$, and in particular $\mathbb{E}(Z_n) \to \infty$ as $n \to \infty$. From the second statement in Lemma 29 we obtain for all $j \neq k$ that

$$\mathbb{P}(E_j \cap E_k) \le C\,\mathbb{P}(E_j)\mathbb{P}(E_k),$$

so

$$\begin{aligned}
\mathbb{E}[Z_n^2] &= \mathbb{E}\left(\left(\sum_{k=1}^{n} \mathbb{1}_{E_k}\right)^2\right) \\
&= \sum_{k=1}^{n} \mathbb{P}(E_k) + \sum_{j=1}^{n}\sum_{k=1}^{n} \mathbb{1}_{\{j\neq k\}} \mathbb{P}(E_j \cap E_k) \\
&\le \mathbb{E}(Z_n) + C\sum_{j=1}^{n}\sum_{k=1}^{n} \mathbb{1}_{\{j\neq k\}} \mathbb{P}(E_j)\mathbb{P}(E_k) \\
&\le \mathbb{E}(Z_n) + C\big(\sum_{j=1}^{n} \mathbb{P}(E_j)\big)^2 \\
&= \mathbb{E}(Z_n) + C(\mathbb{E}(Z_n))^2
\end{aligned}$$

Since $\mathbb{E}(Z_n) \to \infty$ as $n \to \infty$ we obtain

$$\frac{(\mathbb{E}(Z_n))^2}{\mathbb{E}(Z_n^2)} \ge \frac{1}{C} - o(1).$$

This allows us to apply part (iii) of the Kochen–Stone theorem to deduce that

$$\mathbb{P}\left(\limsup_{n\to\infty} \frac{Z_n}{\mathbb{E}(Z_n)} > 0\right) \ge \frac{1}{C}.$$

Since $\mathbb{E}(Z_n) \to \infty$ as $n \to \infty$, this implies $\mathbb{P}(E_k \text{ occurs i.o.}) > 0$ and in particular $\mathbb{P}(X_{2n} = 0 \text{ i.o.}) > 0$. It then follows from Kolmogorov's zero-one law that

$$\mathbb{P}(\exists r \in \mathbb{Z} : Y \text{ visits } r \text{ i.o.}) = 1,$$

since this is a tail event which occurs with positive probability.

From this it is easy to see that the probability that Y (and therefore X) visits every even integer is 1. Indeed, for every time t we have

$$\mathbb{P}(Y_{t+1} = Y_t + 2 \mid Y_t) = 1/4 = \mathbb{P}(Y_{t+1} = Y_t - 2 \mid Y_t),$$

and hence, for all $m, n \in 2\mathbb{Z}$,

$$\mathbb{P}(Y \text{ visits } m \text{ i.o. but does not visit } n \text{ i.o.}) = 0.$$

The result now follows by summing over the choices for m.

We are now ready to give a proof of Lemma 29.

Proof of Lemma 29. For any event E such that $\mathbb{P}(E) > 0$ and any integrable random variable $T \geq 0$ on the same probability space such that $\mathbb{E}(T) > 0$ and $T = 0$ on the complement of E, we have

$$\mathbb{P}(E) = \frac{\mathbb{E}(T\mathbb{1}_E)}{\mathbb{E}(T \mid E)} = \frac{\mathbb{E}(T)}{\mathbb{E}(T \mid E)}.$$

Take $E = E_k$, the event that X visits 0 during the $(2k)^{th}$ block, i.e. during the times $(n_{2k}, \ldots, n_{2k+1} - 1)$, and let $T = T_k$, where T_k is the number of visits of X to 0 during the $(2k)^{th}$ block.

We will prove the following four estimates, where c_1 and c_2 are positive constants:

(i) $\mathbb{E}(T_k) = \Omega(1)$ as $k \to \infty$,

(ii) $\mathbb{E}(T_k \mid E_k) = O(k)$ as $k \to \infty$,

(iii) $\mathbb{E}(T_k \mid E_j) < c_1$, for all j, k such that $j < k$,

(iv) $\mathbb{E}(T_k \mid E_j \cap E_k) > c_2 k$ for all j, k such that $j < k$.

Observe that (i) and (ii) imply that

$$\mathbb{P}(E_k) = \frac{\mathbb{E}(T_k)}{\mathbb{E}(T_k \mid E_k)} = \Omega(1/k) \text{ as } k \to \infty.$$

On the other hand, (iii) and (iv) imply in the same way that

$$\mathbb{P}(E_k \mid E_j) = \frac{\mathbb{E}(T_k \mid E_j)}{\mathbb{E}(T_k \mid E_j \cap E_k)} < \frac{c_1}{c_2 k} \text{ for all } j, k \text{ with } j < k.$$

It follows that there is a finite constant C such that for all $j < k$

$$\mathbb{P}(E_k \mid E_j) \leq C\mathbb{P}(E_k).$$

We make the estimates (i)-(iv) by relating the portion of the walk Y during the $(2k)^{th}$ block to a simple symmetric random walk on $\mathbb{Z}^2$. Fix k for now. Let $m_0 = (n_{2k} - 1)/2$ and $m_1 = (n_{2k+1} - 1)/2$. Note that

$$m_1 - m_0 = \frac{n_{2k+1} - n_{2k}}{2} = 4^{2k}/2 = 2^{4k-2}.$$

We define a walk (a_m, b_m) for $m_0 \le m \le m_1$ by setting $(a_{m_0}, b_{m_0}) = (0,0)$, and then, inductively for $m = m_0+1, \ldots, m_1$,

$$(a_m, b_m) = (a_{m-1}, b_{m-1}) + \begin{cases} (1,0) & \text{if } Y_m - Y_{m-1} = 2^{2k+1}, \\ (-1,0) & \text{if } Y_m - Y_{m-1} = -2^{2k+1}, \\ (0,1) & \text{if } Y_m - Y_{m-1} = 2, \\ (0,-1) & \text{if } Y_m - Y_{m-1} = -2. \end{cases}$$

Let L be the random arithmetic progression in $\mathbb{Z}^2$ defined by

$$L = \{(a,b) \in \mathbb{Z}^2 \,:\, 2^{2k+1}a + 2b + Y_{m_0} = 0\}.$$

For $m_0 \le m \le m_1$ we have $Y_m = 0$ if and only if $(a_m, b_m) \in L$. Thus, T_k is the number of times that (a_m, b_m) hits L.

Proof of (i) To get a lower bound on $\mathbb{E}(T_k)$, we first apply Chebyshev's inequality to show that Y_{m_0} is often not too large:

$$\begin{aligned} \mathrm{Var}(Y_{m_0}) = \mathrm{Var}(X_{n_{2k-1}}) &= \sum_{n=1}^{n_{2k}-1} a_n^2 \\ &= \sum_{n=1}^{2k-1} \frac{4^k}{4}((2^n+1)^2 + (2^n-1)^2) \\ &\le 4^{2k}/12 \cdot 2(1+2^{2k})^2 \\ &\le 2^{8k}. \end{aligned}$$

Hence, using that $\mathbb{E}(Y_{m_0}) = 0$, we have that

$$\mathbb{P}(|Y_{m_0}| < 2^{1+4k}) \ge 1/4.$$

When $|Y_{m_0}| < 2^{1+4k}$, we can express Y_{m_0} as $2^{2k+1}\alpha + 2\beta$, where $|\alpha| \le 2^{2k}$ and $|\beta| \le 2^{2k}$. The expected number of visits of (a_m, b_m) to $(-\alpha, -\beta)$ is

$$\sum_{m=m_0+1}^{m_1} \mathbb{P}((a_m, b_m) = (-\alpha, -\beta)).$$

We now use the well-known observation about the simple symmetric two-dimensional random walk that $(a_t + b_t)_{t=m_0}^{m_1}$ and $(a_t - b_t)_{t=m_0}^{m_1}$ are independent simple symmetric random walks on $\mathbb{Z}$, started at 0 at $t = m_0$. By Corollary 21, whenever $m - m_0 > (m_1 - m_0)/2 = 2^{4k-3}$ and $m - m_0$ has the same parity as $\alpha + \beta$, we have

$$\begin{aligned} \mathbb{P}((a_m, b_m) = (-\alpha, -\beta)) &= \mathbb{P}(a_m + b_m = -\alpha - \beta)\mathbb{P}(a_m - b_m = -\alpha + \beta) \\ &\ge \left(\frac{1}{\sqrt{\pi(m-m_0)/2}} e^{-\frac{(\alpha+\beta)^2}{2(m-m_0)}} - \frac{c}{m-m_0} \right) \\ &\quad \cdot \left(\frac{1}{\sqrt{\pi(m-m_0)/2}} e^{-\frac{(\alpha-\beta)^2}{2(m-m_0)}} - \frac{c}{m-m_0} \right) \\ &\ge \left(\frac{e^{-16}}{\sqrt{\pi 2^{4k-3}}} - \frac{c}{2^{4k-3}} \right)^2 \\ &= \Omega(2^{-4k}). \end{aligned}$$

Since the number of values of m to which this applies is 2^{4k-4}, we obtain the asymptotic lower bound (i), i.e. $\mathbb{E}(T_k) = \Omega(1)$ as $k \to \infty$.

Proof of (ii) By conditioning on the time of the first visit to 0 during the $(2k)^{th}$ block, we find that

$$\mathbb{E}(T_k \mid E_k) \leq \mathbb{E}(T_k \mid X_{n_{2k}} = 0).$$

Hence, to prove (ii) it suffices to show that $\mathbb{E}(T_k \mid X_{n_{2k}} = 0) = O(k)$ as $k \to \infty$.

The expected number of returns to $(0,0)$ in the first $2N$ steps of simple symmetric random walk on $\mathbb{Z}^2$ started at $(0,0)$ is

$$\sum_{i=1}^{N} 2^{-2i} \binom{2i}{i}^2 = \Theta(\log(N)).$$

Hence, the expected number of visits of $(a_t, b_t)_{t=m_0}^{m_1}$ to $(0,0)$ is $\Theta(\log(\frac{m_1-m_0}{2})) = \Theta(k)$. We now show that the expected number of visits of $(a_t, b_t)_{t=m_0}^{m_1}$ to all the nonzero points in the arithmetic progression L_0 is $O(1)$, where

$$L_0 = \{(a,b) \in \mathbb{Z}^2 \;:\; 2^{2k+1}a + 2b = 0\} = \langle (1, -2^{2k}) \rangle.$$

$$\begin{aligned}
&\sum_{m=m_0+1}^{m_1} \mathbb{P}((a_m, b_m) = (n, -2^{2k}n)) \\
&\qquad = \sum_{m=m_0+1}^{m_1} \mathbb{P}(a_m + b_m = n(1-2^{2k}))\mathbb{P}(a_m - b_m = n(1+2^{2k})) \\
&\qquad \leq 2^{4k-2} \sup_{m \geq 2^{4k}n} \mathbb{P}(a_m + b_m = n(1-2^{2k}))\mathbb{P}(a_m - b_m = n(1+2^{2k})) \\
&\qquad \leq 2^{4k-2} \sup_{m \geq 2^{4k}n} \left(\frac{1}{\sqrt{\pi m/2}} e^{-\frac{(n(2^{2k}-1))^2}{2m}} + \frac{c}{m} \right)^2 \\
&\qquad \leq 2^{4k-2} \left(\sqrt{\frac{2}{\pi e (n(2^{2k}-1))^2}} + \frac{c}{2^{4k}n} \right)^2 \\
&\qquad \leq c/n^2,
\end{aligned}$$

Summing over all $n \in \mathbb{Z} \setminus \{0\}$ gives an upper bound of $c\pi^2/3$, and so we have proved estimate (ii).

Proof of (iii) For estimate (iii), we have

$$\begin{aligned}
\mathbb{E}(T_k \mid E_j) &\leq \max_{r \in 2\mathbb{Z}} \mathbb{E}(T_k \mid X_{n_{2j+1}} = r) \\
&\leq \sum_{n=n_{2k}}^{n_{2k+1}} Q_1(X_n - X_{n_{2j+1}}) \\
&\leq \frac{4^{2k}}{2} \cdot Q_1(X_{n_{2k}} - X_{n_{2k-1}}).
\end{aligned}$$

Let m_k be the number of pairs $(2i, 2i+1)$ with $n_{2k-1} \le 2i < n_{2k}$ for which $\epsilon_{2i} = -\epsilon_{2i+1}$. Then m_k is a binomial random variable with $4^{2k}/4$ trials and success probability $1/2$. A Chernoff bound immediately implies that the probability that m_k lies in the interval $[4^{2k}/12, 4^{2k}/6]$ is $1 - o(4^{-2k})$. Conditional on the value of m_k, the increment $X_{n_{2k}} - X_{n_{2k-1}}$ is expressible as a sum $2A_1 + 2^{2k+1}A_2$, where A_1 is a sum of m_k Rademacher random variables, A_2 is a sum of $4^{2k}/4 - m_k$ Rademacher random variables, and A_1 and A_2 are independent.

Now condition on m_k, and assume that $m_k \in [4^{2k}/12, 4^{2k}/6]$. Consider the digits of $A_1 + 2^k A_2$ in base 4^k. The units digit is determined by A_1 alone, and by Corollary 22 there is a constant c not depending on k or on m_k such that $\sup_r \mathbb{P}(A_1 \equiv r \pmod{4^k}) \le c/4^k$.

Conditional on m_k and A_1, the next digit is determined by A_2 and again $\sup_s \mathbb{P}(A_2 \equiv s \pmod{4^k}) \le c/4^k$. Hence, each possible value of the last two digits occurs with probability at most $c^2/4^{2k}$. It follows that

$$Q_1(X_{n_{2k}} - X_{n_{2k-1}}) \le c^2/4^{2k} + o(4^{-2k}).$$

This proves estimate (iii).

Proof of (iv) For estimate (iv), we begin by noting that the arguments used for estimates (i) and (ii) still work when we further condition on E_j. Indeed, the estimate for (ii) is unchanged (as we immediately condition on $X_{n_{2k}} = 0$) and the estimate for (i) is only improved by conditioning on E_j (as this reduces the variance of Y_{m_0}). Hence, we have

$$\mathbb{P}(E_k \mid E_j) \ge c/k,$$

for a constant $c > 0$ that does not depend on j. Let A_k be the event that X visits 0 between times n_{2k} and $n_{2k+1} - 4^k$. Let T'_k be the number of visits during this interval, so $A_k = \{T'_k \ge 1\}$ and $E_k \setminus A_k \subseteq \{T_k - T'_k \ge 1\}$. This implies that $\mathbb{P}(A_k^c \mid E_k \cap E_j) \le \mathbb{E}(T_k - T'_k \mid E_k \cap E_j)$, and hence

$$\mathbb{E}(T_k - T'_k \mid E_k \cap E_j) \le \frac{\mathbb{E}(T_k - T'_k \mid E_j)}{\mathbb{P}(E_k \mid E_j)} \le \frac{k}{c}\mathbb{E}(T_k - T'_k \mid E_j).$$

To estimate $\mathbb{E}(T_k - T'_k | E_j)$ we repeat the method that we used for estimate (iii). Recall that $Q_1(X_n - X_{n_{2j+1}}) \le Q_1(X_{n_{2k}} - X_{n_{2k-1}}) = O(4^{-2k})$, and so

$$\begin{aligned}
\mathbb{E}(T_k - T'_k \mid E_j) &\le \sum_{n=n_{2k+1}-4^k}^{n_{2k+1}} \mathbb{P}(X_n = 0 \mid E_j) \\
&\le \sum_{n=n_{2k+1}-4^k}^{n_{2k+1}} Q_1(X_n - X_{n_{2j+1}}) \\
&= O(4^{-k}).
\end{aligned}$$

Hence,

$$\mathbb{P}(A_k | E_k \cap E_j) = 1 - O(k4^{-k}) = 1 - o(1) \quad \text{as } k \to \infty,$$

uniformly in j. Now

$$\begin{aligned}
\mathbb{E}(T_k \mid E_k \cap E_j) &\geq \mathbb{P}(A_k \mid E_k \cap E_j)\, \mathbb{E}(T_k \mid A_k \cap E_j) \\
&\geq (1 - o(1)) \min_{n \in \{n_{2k}, \dots, n_{2k+1} - 4^k\}} \mathbb{E}(T_k \mid X_n = 0) \\
&= \Omega(\log(4^k)) \\
&= \Omega(k),
\end{aligned}$$

where we have again used that the expected number of returns to 0 of a two-dimensional simple symmetric random walk in its first $2N$ steps is $\Omega(\log(N))$. $\square$

6 Recurrence and transience for slowly growing step sizes

We start by showing that a slowly growing non-decreasing integer sequence $(a_n)_{n\geq 1}$ gives a transient Rademacher random walk if the set of values it takes is a little sparse.

Lemma 30. *Suppose $(a_n)_{n\geq 1}$ is a non-decreasing sequence of positive integers which takes values in a set S. Suppose that the value s appears L_s times in (a_n) and suppose further that $\sum_{s\in S} 1/s < \infty$ and, for some $\varepsilon > 0$ and all large enough s, there is some $s' < s$ for which $L_{s'} \geq \varepsilon n^2$. Then the Rademacher random walk X associated to $(a_n)_{n\geq 1}$ is transient.*

Proof. Fix any finite set F. Using the hypothesis about s', we can apply Corollary 22 to see that for all large enough $s \in S$, the probability that X_n is congruent to any element of F (mod s) at the beginning of the block of steps of size s is at most $|F|c(\varepsilon)/s$. The walk X can only visit F during the s-block if it is congruent to an element of F modulo s and, since $\sum_{s\in S} 1/s < \infty$, the first Borel–Cantelli lemma shows that almost surely this happens for only finitely many $s \in S$. Hence, X is transient. $\square$

Now we can easily prove Lemma 9, which we restate here for convenience.

Lemma 9. *Let $f : \mathbb{N} \to [1, \infty)$ be any function such that $f(n) \to \infty$ as $n \to \infty$. Then there exists a non-decreasing sequence $(a_n)_{n\geq 1}$ of integer step sizes for which $a_n \leq f(n)$ for all n and the associated Rademacher random walk is both irreducible and transient.*

Proof. Since we could choose to start with any finite number of steps with a step size of 0, we may assume without loss of generality that $f(n) \geq 9$ for all n. Let $p_i = (2i + 1)^2$ for each $i \geq 1$, and choose a sequence $\ell_1, \ell_2, \dots$ of integers such that for each $i \geq 1$ we have

1. $\ell_{2i} \equiv i + 1 \pmod 2$,
2. $\ell_{2i+1} \equiv i \pmod 2$,
3. $\ell_i \geq p_{i+1}^2$,

4. $f(n) \geq p_{i+1}$ for all $n > \sum_{k=1}^{i} \ell_k$.

This is always possible as $f(n) \to \infty$ as $n \to \infty$. Construct the sequence $(a_n)_{n\geq 1}$ by letting the first ℓ_1 terms be p_1, the next ℓ_2 terms be p_2, and so on. That is, $a_n = p_i$ whenever $c_i < n \leq c_{i+1}$, where $c_i = \sum_{j=1}^{i-1} \ell_j$. The final condition in the list above ensures that $a_n \leq f(n)$ for all n. Let $(X_n)_{n\geq 0}$ be a Rademacher walk with step sizes given by the sequence $(a_n)_{n\geq 1}$. Since $\gcd(p_{i-1}, p_i) = 1$, the construction ensures that $(a_n)_{n\geq 1}$ satisfies the hypotheses of Lemma 30, so X is transient.

It remains to show that X is irreducible. Note that for any $i \geq 1$ we have $(i+1)p_{2i} - ip_{2i+1} = 1$. By our assumptions on the parity of the ℓ_i, we have that in two consecutive blocks where the step sizes are p_{2i} in the first and p_{2i+1} in the second, it occurs with positive probability that the total increment in block $2i$ is $(i+1)p_{2i}$ and the total increment in block $(2i+1)$ is $-ip_{2i+1}$, in which case the total increment from these two blocks is 1. Likewise, it occurs with positive probability that the total increment in block i is $-(i+1)p_{2i}$ and the total increment in block $(i+1)$ is ip_{2i+1}, so that the total increment from the two blocks is -1. Hence, X is irreducible. □

We now turn to the case of bounded step sizes, and prove Lemma 10, which states that if $\sum_{n=1}^{\infty} a_n^2 = \infty$ then the Rademacher random walk X associated to $(a_n)_{n\geq 1}$ is almost surely unbounded both above and below.

Proof of Lemma 10. To show that (X_n) is almost surely unbounded both below and above, it suffices to show for any constant C that almost surely $X_n \leq C$ i.o. and $X_n \geq C$ i.o. as well. To prove this, we will show that whenever $\mathbb{P}(X_m = x) > 0$, we have

$$\mathbb{P}(\exists n > m \,:\, (X_n - C)(x - C) \leq 0 \mid X_m = x) = 1. \tag{8}$$

For any $n > m$, the increment $X_n - X_m$ is independent of X_m and has a symmetric distribution with variance $\sum_{k=m+1}^{n} a_k^2$. Its fourth moment satisfies

$$\begin{aligned}
\mathbb{E}((X_n - X_m)^4) &= \mathbb{E}\left(\left(\sum_{k=m+1}^{n} \epsilon_k a_k\right)^4\right) \\
&= \sum_{k=m+1}^{n} a_k^4 + 3 \sum_{k=m+1}^{n} \sum_{j=m+1}^{n} \mathbb{1}_{\{j\neq k\}} a_k^2 a_j^2 \\
&= 3 \sum_{k=m+1}^{n} \sum_{j=m+1}^{n} a_k^2 a_j^2 - 2 \sum_{k=m+1}^{n} a_k^4 \\
&\leq 3\left(\sum_{k=m+1}^{n} a_k^2\right)^2.
\end{aligned}$$

We now apply the Paley–Zygmund inequality to the random variable Z

defined by $Z = (X_n - X_m)^2$.

$$\mathbb{P}\left(|X_n - X_m| \geq \frac{1}{2}\left(\sum_{k=m+1}^{n} a_k^2\right)^{1/2}\right) = \mathbb{P}\left(Z > \frac{1}{4}\mathbb{E}(Z)\right)$$
$$\geq \left(\frac{3}{4}\right)^2 \frac{\mathbb{E}(Z)^2}{\mathbb{E}(Z^2)}$$
$$\geq \frac{3}{16}.$$

We remark that the above inequality complements the statement of Tomaszewski's conjecture, recently proved by Keller and Klein [11], which tells us that

$$\mathbb{P}\left(|X_n - X_m| \leq \left(\sum_{k=m+1}^{n} a_k^2\right)^{1/2}\right) \geq \frac{1}{2}.$$

We can now prove (8). Define a sequence of stopping times $\tau_0 = m < \tau_1 < \tau_2 < \dots$ inductively by

$$\tau_i = \min\left(\left\{n > \tau_{i-1} \, : \, \frac{1}{2}\left(\sum_{k=1+\tau_{i-1}}^{n} a_k^2\right)^{1/2} > \left|X_{\tau_{i-1}} - C\right|\right\}\right).$$

Since $\sum_{k=1}^{\infty} a_k^2 = \infty$, we have $\tau_i < \infty$ a.s. for every integer $i \geq 0$. Let $\mathcal{G}_i$ denote the σ-algebra generated by $X_1, \dots, X_{\tau_i}$. For each $i \geq 1$ we have

$$\mathbb{P}((X_{\tau_i} - C)(X_{\tau_{i-1}} - C) \leq 0 \mid \mathcal{G}_{i-1}) \geq \frac{3}{32}.$$

By the conditional Borel–Cantelli lemma, we find that almost surely there exists a random $i < \infty$ such that $(X_{\tau_i} - C)(X_{\tau_{i-1}} - C) \leq 0$ and, taking $n = \tau_i$ for the least such i, we have $(X_n - C)(X_m - C) \leq 0$. $\square$

7 Transience for sequences that cover all natural numbers

The main aim of this section is to prove Theorem 5, and then to deduce Corollary 6.

Theorem 5. *Let $(a_n)_{n\geq 1}$ be a non-decreasing sequence of integers, and let L_i be the number of times i appears in the sequence. Suppose that, for all large enough n,*

$$\sum_{\substack{i \leq n \\ \gcd(i,n)=1}} L_i \geq 2n^2 \tag{1}$$

and

$$\sum_{i=1}^{n-1} i^2 L_i \geq 4n^2 \log^3(n) \cdot L_n. \tag{2}$$

Then the Rademacher random walk X associated with $(a_n)_{n\geq 1}$ is transient.

Proof. Fix $C > 0$, and let E_n be the event that the walk is within C of the origin after one of the steps of size n, that is, $E_n = \{\exists i : a_i = n, |X_i| \leq C\}$. We will show that the probability of these events is summable, and therefore only finitely of the E_n occur almost surely. Since the probability that the walk is C-recurrent is zero for all C, the walk must be transient.

Clearly, the probability of E_n is 0 if $L_n = 0$, so suppose that $L_n \geq 1$ and that n is large enough for (1) and (2) to hold. Let $N = \sum_{i=1}^{n-1} L_i$ so that $a_{N+1} = \cdots = a_{N+L_n} = n$. We split the event E_n into two cases based on the size of $|X_N|$. When $|X_N|$ is large, the probability that L_n steps of size n will travel far enough to be within C of the origin is small enough to be summable. When $|X_N|$ is small, we use the fact that X_N is well-distributed over the equivalence classes modulo n, and the probability that steps of size n could possibly get within C of the origin is $O(C/n)$. By combining this with the probability that $|X_N|$ is small, we find that $\sum_{n=1}^{\infty} \mathbb{P}(E_n) < \infty$.

First, consider the case where $|X_N|$ is large, by which we mean $|X_N|^2 \geq \sum_{i=1}^{n-1} i^2 L_i / \log^2(n)$. Define $M := \sum_{i=1}^{n-1} i^2 L_i / \log^2(n)$. By the reflection principle we have

$$\mathbb{P}\left(\max_{0 \leq t \leq L_n} \sum_{i=1}^{t} \epsilon_{(N+i)} n \geq \sqrt{M}\right) = 2\mathbb{P}\left(\sum_{i=1}^{L_n} \epsilon_i n \geq \sqrt{M}\right) - \mathbb{P}\left(\sum_{i=1}^{L_n} \epsilon_i n = \sqrt{M}\right)$$
$$\leq 2\exp\left(-\frac{M}{2n^2 L_n}\right).$$

By our assumption on $\sum_{i=1}^{n-1} i^2 L_i$, we have that the probability of this event is at most $2n^{-2}$.

Now consider the case where $|X_N|^2 \leq M$. We split the sum X_N into two parts, a small part which ensures that X_N is well distributed over the equivalence classes modulo n, and another part that we know must be large. Pick a set $A \subseteq \{1, \ldots, N\}$ of size $|A| = n^2$ such that $\gcd(a_i, n) = 1$ for all $i \in A$ and $\sum_{i \in A} a_i^2$ is as small as possible. Let $A^c = [N] \setminus A$, and note that $\sum_{i \in A^c} a_i^2 \geq 2n^2 \log^3(n) L_n$. Define S_A and S_{A^c} by $S_A = \sum_{i \in A} a_i \epsilon_i$ and $S_{A^c} = \sum_{i \in A^c} a_i \epsilon_i$.

Note that the event $\{|X_N| \leq M\}$ is contained in the event

$$\{|S_A| \leq 2\log(n) n^2\} \cap \{|S_{A^c}| \leq 2n^2 \log(n) L_n\},$$

and that if E_n is to occur, there must be some $c \in [-C, C]$ for which $X_N \equiv c \mod n$. By our choice of A, for every x, the probability that $S_A \equiv x \mod n$ is at most $3/n$ by Theorem 15, and this implies that, even given the value of S_{A^c}, the probability that there is some $c \in [-C, C]$ for which $X_N \equiv c \mod n$ is at most $3/n$. To finish this case, we will show that the probability of $\{|S_{A^c}| \leq 2n^2 \log(n) L_n\}$ is at most $C' / \log^2(n)$ and that the probability of $\{|S_A| \leq 2\log(n) n^2\}$ is at least $1 - n^{-2}$. These imply that the probability of $E_n \cap \{|X_N| \leq M\}$ is at most

$$\frac{C'}{\log^2(n)}\left(\frac{1}{n^2} + \frac{3}{n}\right)$$

Since the probability of $E_n \cap \{|X_N| \geq M\}$ is at most $2n^{-2}$, the probability of E_n is $O(n^{-1}\log^{-2}(n))$. Hence, the sum $\sum_{n=1}^{\infty} \mathbb{P}(E_n) < \infty$ and almost surely only finitely many E_n occur, which means the walk is transient.

Let us bound the probability that $|S_{A^c}| \leq 2n^2 \log(n) L_n$. By the Berry–Esséen Theorem, we have that

$$\begin{aligned}
\mathbb{P}\big(|S_{A^c}| \leq 2n^2\log(n)L_n\big) &= \mathbb{P}\left(\frac{|S_{A^c}|}{\sum_{i\in A^c} a_i^2} \leq \frac{2n^2\log(n)L_n}{\sum_{i\in A^c} a_i^2}\right) \\
&\leq \mathbb{P}\left(\frac{|S_{A^c}|}{\sum_{i\in A^c} a_i^2} \leq \frac{1}{2\log^2(n)}\right) \\
&\leq \Phi\left(\frac{1}{2\log^2(n)}\right) - \Phi\left(-\frac{1}{2\log^2(n)}\right) + C'\frac{n-1}{\sum_{i\in A^c} a_i^2} \\
&\leq \frac{1}{\sqrt{2\pi}\log^2(n)} + \frac{C'}{n}.
\end{aligned}$$

Finally, we need to bound the probability that $|S_A| \leq 2\log(n)n^2$, for which we use Lemma 25. Trivially, $\sum_{i\in A} a_i^2 \leq n^4$ and,

$$\mathbb{P}(S_A \geq 2\log(n)n^2) \leq e^{-2\log(n)}. \qquad \square$$

Given the theorem above, Corollary 6 follows easily. Indeed, we only need to show that $a_n = \lfloor \log^{1+\varepsilon}(n) \rfloor$ satisfies the appropriate conditions.

Corollary 6. *For any $\alpha > 1$, the Rademacher random walk with step sizes $a_n = \lfloor \log^{\alpha}(n) \rfloor$ is transient*

Proof. Clearly, (a_n) is a non-decreasing sequence of integers, so we can apply Theorem 5 provided the L_n satisfy the necessary conditions. The ith step size a_i equals n exactly when $e^{n^{1/\alpha}} \leq i < e^{(n+1)^{1/\alpha}}$, so $L_n = e^{(n+1)^{1/\alpha}} - e^{n^{1/\alpha}} + O(1)$. By the Mean Value Theorem, there is $x \in (n, n+1)$ such that

$$e^{(n+1)^{1/\alpha}} - e^{n^{1/\alpha}} = \frac{e^{x^{1/\alpha}}}{\alpha x^{1-1/\alpha}}.$$

In particular, $L_{n-1} \geq 2n^2$ for large enough n and (1) is satisfied. We also have

$$\begin{aligned}
\sum_{i=1}^{n-1} i^2 L_i &\geq \sum_{i=2}^{n-1} i^2\left(\frac{e^{i^{1/\alpha}}}{\alpha i^{1-1/\alpha}} + O(1)\right) \\
&\geq O(n^3) + \frac{1}{\alpha}\int_1^{n-1} x^{1+1/\alpha} e^{x^{1/\alpha}}\,dx \\
&= (n-1)^2 e^{(n-1)^{1/\alpha}} + O(e^{n^{1/\alpha}}/n^2).
\end{aligned}$$

Comparing this with $4n^2\log^3(n)\left(\frac{e^{(n+1)^{1/\alpha}}}{\alpha(n+1)^{1-1/\alpha}} + O(1)\right)$, we see that (2) is also satisfied, and we can apply Theorem 5 to finish the proof. $\square$

8 Unbounded step sequences whose gaps tend to zero

In this section we prove Theorem 13, which asserts that if the step sizes of a Rademacher random walk are unbounded with gaps converging to zero, the walk is either transient or topologically recurrent.

Lemma 31. *Suppose $(a_n)_{n\geq 1}$ is a sequence such that $\limsup_{n\to\infty} a_n = \infty$ and $|a_n - a_{n-1}| \to 0$ as $n \to \infty$. Let X and X' be Rademacher random walks with step sizes given by (a_n) started at time k at locations $X_k = N$ and $X'_k = N + d$ for some $d \neq 0$. Then, for any $\varepsilon > 0$, the walks X' and X' can be coupled so that a.s. $\lim_{n\to\infty}(X_n - X'_n)$ exists and lies in $[0, \varepsilon]$.*

Proof. Think of the problem of coupling X and X' as a game as follows. Just before time i, we know $\epsilon_1, \ldots, \epsilon_{i-1}$ and $\epsilon'_1, \ldots, \epsilon'_{i-1}$ and we must choose either to couple ϵ'_i and ϵ_i so that $\epsilon'_i = \epsilon_i$, or to couple them so that $\epsilon'_i = -\epsilon_i$. Once we have made our choice, the value of ϵ_i is revealed, and ϵ'_i is either ϵ_i or $-\epsilon_i$ depending on the choice that we made. We win the game if for some $i \geq k$ we achieve $X_i - X'_i \in [0, \varepsilon]$. The proof consists of a strategy for winning this game eventually with probability 1. Once we have won the game we may couple all subsequent signs to be equal, so that $X_i - X'_i$ stabilizes.

Our strategy is organised as a sequence of episodes. In each episode we will win with probability at least 1/4, conditional on all the outcomes in previous episodes. For $i \geq 1$, episode i will begin at time $m_{i-1} + 1$ and end at time m_i, where $m_0 = k$. To describe episode i for any $i \geq 1$, assume we know m_{i-1} $X_{m_{i-1}}$ and $X'_{m_{i-1}}$, and assume we have not yet won the game at time m_{i-1}. Let $d_i = X_{m_{i-1}} - X'_{m_{i-1}}$. In particular, $d_1 = d$. Let $\delta_i = \min(\epsilon, |d_i|)$. Note that $\delta_i > 0$ since if $d_i = 0$ then we have already won the game. Choose $n_i \geq m_{i-1}$ sufficiently large that for all $n \geq n_i$ we have $|a_n - a_{n-1}| < \delta_i/2$. Let

$$x_i = \begin{cases} d_i/2 & \text{if } d_i > 0, \\ -d_i/2 + \delta_i/2, & \text{if } d_i \leq 0. \end{cases}$$

Since $\limsup_{n\to\infty} a_n = \infty$, we may find $m_i > n_i$ such that

$$a_{m_i} - a_{n_i} \in [x - \delta_i/2, x].$$

Now choose to couple the signs ϵ_i and ϵ'_i driving the movements of the walks X and X' to be equal for each i in the range $m_{i-1} \leq i \leq n_i - 1$ and for $n_i + 1 \leq i \leq m_i - 1$. Choose to couple $\epsilon_{n_i} = -\epsilon'_{n_i}$ and $\epsilon_{m_i} = -\epsilon'_{m_i}$.

There are four possible options for $X'_{m_i} - X_{m_i}$, each having probability 1/4 conditional on the outcomes preceding episode i:

$$X'_{m_i} - X_{m_i} = \begin{cases} d_i + 2(a_{m_i} + a_{n_i}) & \epsilon_{n_i} = -1, \epsilon_{m_i} = -1, \\ d_i + 2(a_{m_i} - a_{n_i}) & \epsilon_{n_i} = +1, \epsilon_{m_i} = -1, \\ d_i - 2(a_{m_i} - a_{n_i}) & \epsilon_{n_i} = -1, \epsilon_{m_i} = +1, \\ d_i - 2(a_{m_i} + a_{n_i}) & \epsilon_{n_i} = +1, \epsilon_{m_i} = +1. \end{cases}$$

If $d_i > 0$ then $d_i - 2(a_{m_i} - a_{n_i}) \in [0, \varepsilon]$. If $d_i < 0$ then $d_i + 2(a_{m_i} - a_{n_i}) \in [0, \varepsilon]$. Hence, with probability at least $1/4$, we have $X'_{m_i} - X_{m_i} \in [0, \varepsilon]$, in which case we win the game no later than time m_i.

Carrying out the procedure above repeatedly until success, we will almost surely win after finitely many episodes, since in each episode we win with probability at least $1/4$, conditional on the outcomes in previous episodes. □

Using this coupling, we show that if the walk hits the interval $[a, b]$ i.o. with positive probability, then the probability that it hits a second interval i.o. is also positive.

Corollary 32. *Suppose $(a_n)_{n\geq 1}$ is a sequence such that $\limsup_{n\to\infty} a_n = \infty$ and $|a_n - a_{n-1}| \to 0$ as $n \to \infty$. Let $X = (X_n)_{n\geq 1}$ be the Rademacher random walk with step sizes (a_n). Let $a < b$ and $e < f$ and take $m \in \mathbb{N}$ such that $m > (b - a)/(f - e)$. If $\mathbb{P}(X_n \in [a, b]$ i.o.$) \geq p$, then $\mathbb{P}(X_n \in [e, f]$ i.o.$) \geq p/m$.*

Proof. Divide the interval $[a, b]$ into m equal intervals. By a union bound, at least one such interval $[a', b']$ is visited infinitely often with probability at least p/m. We have $[a' + t, b' + t] \subset [e, f)$, where $t = e - a'$. If $t = 0$ there is nothing to do, so let us assume that $t \neq 0$ and further that $t > 0$. The case $t < 0$ is similar. Apply the coupling of Lemma 31 starting at time 0 at $X_0 = 0$ and $X'_0 = t$, taking $\varepsilon = f - (b' + t)$. Then the standard Rademacher random walk X'' defined by $X''_n = X'_n - t$ visits $[e, f]$ infinitely often if X visits $[a', b']$ infinitely often, and this occurs with probability at least p/m. □

The proof of Theorem 13 follows.

Proof. Suppose X is not transient. Then there exists $C < \infty$ such that

$$\mathbb{P}(|X_n| < C \text{ i.o.}) > 0.$$

Apply Corollary 32 taking $[a, b] = [-C, C]$ and $p = \mathbb{P}(|X_n| < C \text{ i.o.})$, to see that whenever $e < f$ we have

$$\mathbb{P}(X_n \in [e, f] \text{ i.o.}) \geq \frac{p}{\lceil 2C/(f - e)\rceil} > 0.$$

Now suppose (for a contradiction) that for some interval $[g, h]$, we have $\mathbb{P}(X_n \in [g, h] \text{ i.o.}) < 1$. Then (by a standard martingale argument) there exists a finite k' and two sequences of signs $\beta_1, \dots, \beta_{k'}$ and $\gamma_1, \dots, \gamma_{k'}$ such that

$$\mathbb{P}(X_n \in [g, h] \text{ i.o.} \mid \epsilon_1 = \beta_1, \dots, \epsilon_k = \beta_{k'}) > 2/3$$

and

$$\mathbb{P}(X_n \in [g, h] \text{ i.o.} \mid \epsilon_1 = \gamma_1, \dots, \epsilon_k = \gamma_{k'}) < 1/3.$$

Take $m = 2$ in Corollary 32, applied to the walk with step sizes $a_{k'+1}, a_{k'+2}, \dots,$ and with

$$[a, b] = [g - (\beta_1 a_1 + \cdots + \beta_{k'} a_{k'}), h - (\beta_1 a_1 + \cdots + \beta_{k'} a_{k'})]$$

and

$$[e, f] = [g - (\gamma_1 a_1 + \cdots + \gamma_{k'} a_{k'}), h - (\gamma_1 a_1 + \cdots + \gamma_{k'} a_{k'})]$$

to obtain a contradiction. □

9 Recurrent Rademacher walks where the step sizes grow arbitrarily quickly

In this section we prove Theorem 3 and Theorem 4, both of which show the existence of sequences of step sizes which grow arbitrarily quickly yet give recurrent Rademacher random walks. The constructions used in the proofs of Theorem 3 and Theorem 4 both work by considering a suitable two-dimensional random walk and understanding the range of the second coordinate at the times when the first coordinate is zero, but the proof of Theorem 3 is much simpler.

Theorem 3. *Let $f : \mathbb{N} \to \mathbb{R}$ be any non-decreasing function. There is an integer sequence $(a_n)_{n\geq 1}$ such that $a_n \geq f(n)$ for all n and the associated Rademacher random walk is recurrent.*

Proof. We define the sequence (a_n) in blocks, starting from the empty sequence. Suppose that $a_1, \dots, a_N$ have already been chosen, and let $M = \sum_{i=1}^{N} a_i$. Since the two-dimensional simple symmetric random walk is recurrent, we can choose some L such that the probability that the two-dimensional simple symmetric random walk (SSRW) has hit every point in $\{(0, y) : y \in [-M, M]\}$ by time L is at least $1/2$. Now we choose r such that $r \geq g(2L + N)$ and define the next $2L$ steps to alternate between $r + 1$ and r. This sequence clearly satisfies the necessary growth condition and it remains to prove that the sequence is weakly recurrent.

Let E_k be the event that the walk hits 0 in the kth block. We claim that uniformly for any outcome on the preceding blocks, the probability of E_k is at least $1/2$. Suppose that kth block starts at a_{N+1} and that the walk is at m immediately before the kth block, i.e. $X_N = m$. Pair up consecutive steps in the kth block and consider the walk $Y = (Y_n)$ where $Y_n = X_{N+2n}$. This starts at m and takes steps of $\pm(2r + 1), \pm 1$ each with probability $1/4$. As before, we define a two dimensional random walk (x_n, y_n) by setting $(x_0, y_0) = (0, 0)$, and then inductively defining (x_n, y_n) for $n = 1, \dots, L$ by

$$(x_n, y_n) = (x_{n-1}, y_{n-1}) + \begin{cases} (1, 0) & \text{if } Y_n - Y_{n-1} = 2r + 1, \\ (-1, 0) & \text{if } Y_n - Y_{n-1} = -(2r + 1), \\ (0, 1) & \text{if } Y_n - Y_{n-1} = 1, \\ (0, -1) & \text{if } Y_n - Y_{n-1} = -1. \end{cases}$$

Clearly, if the walk (x_n, y_n) hits the point $(0, -m)$, then the walk X has hit zero and by our choice of L this happens with probability at least $1/2$. To finish the proof that the walk is recurrent, we can apply the Kochen–Stone theorem to the random variable $Z_n = \sum_{k=1}^{n} \mathbb{1}_{E_k}$. □

We now turn to the proof of Theorem 4. We will again consider the times when the first coordinate of a two-dimensional random walk is 0, but we will have to work with a more complicated two-dimensional random walk and we will need the following lemma.

Lemma 33. *Define for each* $n \geq 1$

$$c_n = \sum_{m=1}^{n-1} \frac{m^{-3/2}}{\sqrt{1+\log m}}.$$

Consider the two-dimensional Rademacher random walk $(Y_n, Z_n)_{n\geq 0}$ *with* n^{th} *step* $\pm(1, c_n)$*, starting at* $(Y_0, Z_0) = (0,0)$*. Almost surely the set* $\{Z_n : Y_n = 0\}$ *is dense in* $\mathbb{R}$*.*

From this lemma it is relatively straightforward to prove Theorem 4.

Theorem 4. *Let* $f : \mathbb{N} \to \mathbb{R}$ *be any non-decreasing function. There is a strictly increasing real sequence* $(a_n)_{n\geq 1}$ *such that* $a_n \geq f(n)$ *for all* n *and the associated Rademacher random walk is topologically recurrent.*

Proof. Define for each $n \geq 1$

$$c_n = \sum_{m=1}^{n-1} \frac{m^{-3/2}}{\sqrt{1+\log m}}.$$

Note that $\sum_{m=1}^{\infty} \frac{m^{-3/2}}{\sqrt{1+\log m}} < \infty$, so $c_n \nearrow c_\infty$ as $n \to \infty$ where $c_\infty < \infty$. Also,

$$c_{n+1} - c_n = \frac{n^{-3/2}}{\sqrt{1+\log n}}.$$

We will define the step size sequence $(a_n)_{n\geq 1}$ as the concatenation of blocks of the form $(x_j + c_0, x_j + c_1, \ldots, x_j + c_{n_j})$, for $j \geq 1$, where for each j we choose n_j and then x_j suitably large given the previous choices. Let $M_j = \sum_{n=1}^{n_1+\cdots+n_{j-1}} a_n$ be the sum of all the terms in the blocks preceding the j^{th} block. According to Lemma 33, we may choose n_j so that with probability at least $1/2$, the walk $(Y_n, Z_n)_{n=0}^{n_j}$ visits the $(1/j)$-neighbourhood of each point in $\{0\} \times [-2M_j, 2M_j]$. Then choose x_j so large that $x_j \geq f(n_1 + \cdots + n_j)$. This ensures $a_n \geq f(n)$ for all n. Note that $n_j \to \infty$ as $j \to \infty$ and hence $M_j \to \infty$ also. For any $t \in \mathbb{R}$, once $M_j + 1/j > t$, we have that conditional on $X_{n_1+\cdots+n_{j-1}}$, the Rademacher walk X_n visits the interval $(t - 1/j, t + 1/j)$ with probability at least $1/2$. Hence, X is topologically recurrent by the conditional Borel–Cantelli lemma. $\square$

It remains to prove Lemma 33.

Proof of Lemma 33. Consider the sequence of successive times $0 = \tau_1 < \tau_2 < \tau_3 < \ldots$ at which $Y_n = 0$. This is almost surely an infinite increasing sequence tending to ∞ since the walk $Y = (Y_n)_{n\geq 0}$ on its own is a simple symmetric random walk on $\mathbb{Z}$, which is recurrent. The distribution of each increment $Z_{\tau_{i+1}} - Z_{\tau_i}$ conditional on the earlier increments has a symmetric distribution. Therefore we can condition on the increment sizes $|Z_{\tau_{i+1}} - Z_{\tau_i}|$ and obtain a *random* Rademacher walk $(Z_{\tau_i})_{i\geq 1}$ with random step sizes $b_i := |Z_{\tau_{i+1}} - Z_{\tau_i}|$ for $i \geq 1$. We claim that almost surely $b_i \to 0$ as $i \to \infty$ and $\sum_{i=1}^{\infty} b_i^2 = \infty$. Given this claim, we may apply Lemma 10 to see that almost surely the Rademacher walk $(Z_{\tau_i})_{i\geq 1}$ is topologically recurrent, which is to say that almost surely $\{Z_n : Y_n = 0\}$ is dense in $\mathbb{R}$.

The key observation is that since the sign of Y_n is constant for n in the interval $[1+\tau_i, \tau_{i+1}-1]$, and $Y_n = 0$ when $n \in \{\tau_i, \tau_{i+1}\}$, and $c_{n+1} - c_n > 0$ for all n, we have

$$\begin{aligned} |Z_{\tau_{i+1}} - Z_{\tau_i}| &= \left|\sum_{n=1+\tau_i}^{\tau_{i+1}} (Y_n - Y_{n-1})c_n\right| = \left|-\sum_{n=1+\tau_i}^{\tau_{i+1}} Y_n(c_{n+1} - c_n)\right| \\ &= \sum_{n=1+\tau_i}^{\tau_{i+1}-1} |Y_n|(c_{n+1} - c_n) = \sum_{n=1+\tau_i}^{\tau_{i+1}-1} |Y_n| \frac{n^{-3/2}}{\sqrt{1+\log n}}. \end{aligned}$$

Since $Y_n = 0$ when $n \in \{\tau_i, \tau_{i+1}\}$, the final sum above is unchanged if we replace the lower limit by $n = \tau_i$ or the upper limit by τ_{i+1}.

Let us use the Komlós–Major–Tusnády coupling to couple the simple symmetric random walk Y_n to a standard Brownian motion $(B_s)_{s\geq 0}$. This is a coupling with the property that for every $\alpha > 0$ there is a positive constant c_α such that for all n one has

$$\mathbb{P}\left(\max_{1\leq j\leq n} \frac{|B_j - Y_j|}{\log n} > c_\alpha\right) < c_\alpha n^{-\alpha}.$$

(See [13, Theorem 7.1.1].) Taking $\alpha > 1$ and using Borel–Cantelli it follows that almost surely

$$\limsup_{s\to\infty} \frac{|B_s - Y_{\lfloor s\rfloor}|}{\log s} < \infty.$$

Since $|Y_{n+1} - Y_n| = 1$ for all n, we also have

$$\limsup_{s\to\infty} \frac{|B_s - Y_{\lceil s\rceil}|}{\log s} < \infty.$$

Let E be the (almost surely finite) random variable defined by

$$E = \sup_{s\geq 1} \frac{\max(|B_s - Y_{\lfloor s\rfloor}|, |B_s - Y_{\lceil s\rceil}|)}{1+\log s}.$$

Then, for each $i \geq 2$, we have $\tau_i \geq 2$ so the function $x^{-3/2}/\sqrt{1+\log x}$ is decreasing on the interval $[\tau_i, \tau_{i+1}]$ and hence

$$\sum_{n=\tau_i}^{\tau_{i+1}-1} |Y_n| \frac{n^{-3/2}}{\sqrt{1+\log n}} > \int_{\tau_i}^{\tau_{i+1}} (|B_s| - (1+\log s)E) \frac{s^{-3/2}}{\sqrt{1+\log s}} \, \mathrm{d}s$$

and

$$\sum_{n=1+\tau_i}^{\tau_{i+1}} |Y_n| \frac{n^{-3/2}}{\sqrt{1+\log n}} < \int_{\tau_i}^{\tau_{i+1}} (|B_s| + (1+\log s)E) \frac{s^{-3/2}}{\sqrt{1+\log s}} \, \mathrm{d}s.$$

Note that $\int_1^\infty \left(\sqrt{1+\log s}\right) s^{-3/2} \, \mathrm{d}s < \infty$, so the sequence of errors

$$\left(\int_{\tau_i}^{\tau_{i+1}} \left(\sqrt{1+\log s}\right) s^{-3/2} E \, \mathrm{d}s\right)_{i\geq 2}$$

is almost surely in $\ell^1(\mathbb{N})$ and hence also in $\ell^2(\mathbb{N})$.

Therefore, to show that the sequence $(b_i)_{i\geq 1} = \left(\left|Z_{\tau_{i+1}} - Z_{\tau_i}\right|\right)_{i\geq 1}$ almost surely tends to 0 but does not lie in $\ell^2(\mathbb{Z})$, it suffices to do the same for the sequence $(I_i)_{i\geq 1}$ defined by

$$I_i := \int_{\tau_i}^{\tau_{i+1}} |B_s| \frac{s^{-3/2}}{\sqrt{1+\log s}} \, \mathrm{d}s.$$

We make the change of variable $s = e^t$, to get

$$I_i = \int_{\log(\tau_i)}^{\log(\tau_{i+1})} |B_{e^t}| \frac{e^{-t/2}}{\sqrt{t+1}} \, \mathrm{d}t = \int_{\log(\tau_i)}^{\log(\tau_{i+1})} \frac{|W_t|}{\sqrt{t+1}} \, \mathrm{d}t,$$

where the process

$$W_t := B_{(e^t)} e^{-t/2}$$

is a stationary Ornstein–Uhlenbeck process whose stationary distribution π is Gaussian with mean 0 and variance 1. For this identity in law, see [22, §8.5.1], in which W is called an *ancient* Ornstein–Uhlenbeck process. W satisfies the SDE

$$dW_t = -\frac{1}{2} W_t \, dt + dB'_t,$$

where $(B'_t)_{t\in\mathbb{R}}$ is another (two-sided) standard Brownian motion.

The rough idea now is that large increments I_i, exceeding a positive constant size, correspond to increasingly large excursions of W from 0, which only occur finitely often, almost surely, but on the other hand the large excursions of W from 0 whose integral is at least a positive constant occur sufficiently regularly to give a subsequence of (I_i) whose sum diverges. Some care is needed to make this precise.

To show that $I_i \to 0$ almost surely as $i \to \infty$, we consider the excursions of W_t above -1 and the excursions of W_t below 1. The times $\log(\tau_i)$ are times at which $|Y_{e^t}| = 0$ and hence $|B_{e^t}| \leq (1+t)E$. There exists a random time $t_0 < \infty$ such that

$$(1+t)e^{-t/2}\, E < 1/2 \text{ for } t \geq t_0.$$

For $i \geq i_0 := \lceil e^{t_0} \rceil$, we have $\tau_i \geq i$ so for all $t > \log(\tau_i)$ we have $(1+t)e^{-t/2}E < 1/2$. For each $i \geq i_0$, on the interval $\tau_i < n < \tau_{i+1}$, either all $Y_n > 0$, in which case

$$W_t > -(1+t)e^{-t/2}E > -1/2 \text{ for all } t \in [\log(\tau_i), \log(\tau_{i+1})],$$

or all $Y_n < 0$, in which case

$$W_t < (1+t)e^{-t/2}E < 1/2 \text{ for all } t \in [\log(\tau_i), \log(\tau_{i+1})].$$

It follows that for each $i \geq i_0$, the integral I_i is dominated either by an integral of the form

$$\frac{1}{\sqrt{a_i+1}} \int_{a_i}^{b_i} (W_t + 1) \mathrm{d}t,$$

where $[a_i, b_i]$ is an excursion interval of W above the level -1 which reaches the level $-1/2$, or by an integral of the form

$$\frac{1}{\sqrt{a_i+1}} \int_{a_i}^{b_i} (-W_t - 1)\mathrm{d}t$$

where $[a_i, b_i]$ is an excursion interval of W below the level 1 which reaches the level $1/2$. Since $\log(\tau_i) \to \infty$ as $i \to \infty$, and W hits each of -1 and 1 at an unbounded set of times almost surely, we have that $a_i \to \infty$ as $i \to \infty$.

The law of the iterated logarithm for B_s as $s \to \infty$ corresponds to a simpler statement about the maximal growth of the stationary Ornstein–Uhlenbeck process (see [22, eq. (8.5.2)]):

$$\limsup_{t\to\infty} \frac{|W_t|}{\sqrt{\log t}} = \sqrt{2} \text{ a.s.}$$

Hence, there is a random time t_1 such that

$$|W_t| < 2\sqrt{\log t} \text{ for all } t \geq t_1. \tag{9}$$

It is known that the hitting times in one-dimensional Ornstein–Uhlenbeck processes have exponential tails. However, we could not find a simple single reference for this fact, and we prove it here. Denote by $\tau_x(y)$ the hitting time of level y starting from level x. First, from Sato [19] it is known that for $x \neq 0$, $\tau_x(0)$ has a probability density function on $(0, \infty)$ given by

$$\frac{|x|}{\sqrt{2\pi}}(e^t - 1)^{-3/2} e^t \exp\left(-\frac{x^2}{2(e^t-1)}\right).$$

This indeed decays exponentially as $t \to \infty$, and immediately implies that the hitting time $\tau_x(y)$ has an exponential tail whenever y lies between x and 0. The hitting time of a general level y starting from x is bounded above by the hitting time of 0 starting from x plus an independent hitting time of y starting from 0, so it suffices to show that $\tau_0(y)$ has an exponential tail, and for this we may assume $y > 0$ without loss of generality since the Ornstein–Uhlenbeck process is symmetric about 0. One way to bound $\tau_0(y)$ is to run the process starting at 0 until it hits $\{-y, y\}$; if it has hit y then we are done; otherwise run from y until you first hit 0, then try again, repeating until success. We succeed after a Geometric(1/2) number of trials. A standard exercise using moment generating functions shows that the sum of a geometric number of i.i.d. random variables with exponential tails itself has an exponential tail. Hence it suffices to show that the hitting time of $\{-y, y\}$ starting from 0 has an exponential tail. This is done in Breiman [5, Thm 1] using the explicit form of the Laplace transform of this hitting time that had been obtained earlier by Darling and Siegert [6]. (Breiman claimed to bound the tail of the first hitting time of y from 0, but we believe his proof actually bounds the first hitting time of $\{-y, y\}$ from 0.)

From the exponential tails of hitting times, it follows that the successive excursions of W above -1 that reach $-1/2$ have lengths that form an i.i.d. sequence with exponential tails. Indeed, the length of each one is the sum of two independent hitting times: the hitting time of $-1/2$ starting at -1, and the hitting time of -1 starting at $-1/2$. The same applies to successive excursions below 1 that reach 1/2. Among both such types of excursion, we now consider only those whose length is at least 1, since by (9) excursions shorter than this after time t_1 can only produce increments I_i that tend to 0. For each type of excursion, the n^{th} instance of the remaining long excursions starts at a time that is at least n, since they all have length at least 1 and are disjoint (although excursions of the two different kinds can overlap). Hence, if L_t is the length of the longest excursion of either kind up to time t, then

$$\sup_{t\geq 2} L_t/\log t < \infty \text{ a.s.}$$

because of the exponential tail bound that we established above. Combining this bound on the excursion lengths with the bound (9) on their heights, we get a bound on the largest integral $\int_{a_i}^{b_i} |W_t + 1|\,\mathrm{d}t$ where $b_i - a_i \geq 1$ and $a_i \leq t$: it is no more than

$$2\sqrt{\log t}\log t \sup_{t\geq 2} L_t/\log t,$$

which is $o(\sqrt{t+1})$ as $t \to \infty$. It follows that $|Z_{\tau_{i+1}} - Z_{\tau_i}| \to 0$ as $i \to \infty$.

It remains to show that the sequence $\left(|Z_{\tau_{i+1}} - Z_{\tau_i}|\right)_{i\geq 1}$ almost surely does not lie in $\ell^2(\mathbb{N})$.

Let us call an excursion of W above -1 a *good excursion* if it contains an excursion above 1 whose integral is at least 1.

Consider a good excursion of W_t above -1 that starts at time s and finishes at time t, where $t_0 \leq s < t-1$, with a subinterval $(s', t') \subset (s, t)$ such that $W_r \geq 1$ for all $r \in (s', t')$ and $\int_{s'}^{t'} W_r\,\mathrm{d}r > 1$. Then there is an excursion of Y above 0, say from time τ_i to time τ_{i+1}, where

$$s < \log \tau_i < s' < t' < \log \tau_{i+1} < t$$

and

$$I_i = \int_{\log(\tau_i)}^{\log(\tau_{i+1})} \frac{W_t}{\sqrt{t+1}}\,\mathrm{d}t > \frac{1}{\sqrt{t'+1}} \int_{s'}^{t'} W_t\,\mathrm{d}t > \frac{1}{\sqrt{t'+1}}\,.$$

The number of disjoint good excursions that lie entirely between times 4^{k-1} and 4^k grows as $\Theta(4^k)$, almost surely as $k \to \infty$. (Again, this follows from the fact that the hitting times of 1 starting from -1 and vice versa in the Ornstein–Uhlenbeck process have exponential tails, together with the Markov property of the Ornstein–Uhlenbeck process.) Hence, the sequence $(I_i)_{i\geq 1}$ almost surely does not lie in $\ell^2(\mathbb{N})$. This completes the proof of Lemma 33. □

Acknowledgements

The research of E.C. and T.J. is supported by the Heilbronn Institute for Mathematical Research. We thank Stanislav Volkov and Zemer Kosloff for helpful conversations about aspects of this work.

References

[1] A. Aamand, N. Alon, J. B. k. T. Houen, and M. Thorup. On sums of monotone random integer variables. *Electron. Commun. Probab.*, 27:Paper No. 64, 8, 2022.

[2] I. Benjamini, R. Pemantle, and Y. Peres. Random walks in varying dimensions. *J. Theor. Probab.*, 9:231–244, 1996.

[3] S. Bhattacharya and S. Volkov. Recurrence and transience of Rademacher series. *ALEA Lat. Am. J. Probab. Math. Stat.*, 20(1):33–51, 2023.

[4] S. Bhattacharya and S. Volkov. Two-dimensional Rademacher walk, 2025. arXiv:2506.16259.

[5] L. Breiman. First exit times from a square root boundary. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume II, Part II*, pages 9–16. University of California Press, Berkeley, 2024.

[6] D. A. Darling and A. J. F. Siegert. The first passage problem for a continuous Markov process. *The Annals of Mathematical Statistics*, 24(4):624–639, 1953.

[7] P. Erdős. On a lemma of Littlewood and Offord. *Bull. Amer. Math. Soc.*, 51:898–902, 1945.

[8] C. G. Esséen. On the Kolmogorov–Rogozin inequality for the concentration function. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 5:210–216, 1966.

[9] G. Halász. Estimates for the concentration function of combinatorial number theory and probability. *Period. Math. Hungar.*, 8(3-4):197–211, 1977.

[10] T. Juškevičius. The sharp form of the Kolmogorov–Rogozin inequality and a conjecture of Leader–Radcliffe. *Bull. Lond. Math. Soc.*, 56(11):3289–3299, 2024.

[11] N. Keller and O. Klein. Proof of Tomaszewski's conjecture on randomly signed sums. *Adv. Math.*, 407:Paper No. 108558, 39, 2022.

[12] S. Kochen and C. Stone. A note on the Borel–Cantelli lemma. *Illinois J. Math.*, 8:248–251, 1964.

[13] G. F. Lawler and V. Limic. *Random Walk: A Modern Introduction*, volume 123 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 2010.

[14] S. J. Montgomery-Smith. The distribution of Rademacher sums. *Proc. Amer. Math. Soc.*, 109(2):517–522, 1990.

[15] Y. Peres. Answer to question 3957937: Recurrence of walks with growing range of step. Mathematics Stack Exchange, Dec. 22, 2020. https://math.stackexchange.com/q/3957937.

[16] R. A. Proctor. Solution of two difficult combinatorial problems with linear algebra. *Amer. Math. Monthly*, 89(10):721–734, 1982.

[17] B. A. Rogozin. An estimate of the concentration functions. *Teor. Verojatnost. i Primenen*, 6:103—-105, 1961.

[18] A. Sárközi and E. Szemerédi. Über ein problem von Erdős und Moser. *Acta Arithmetica*, 11:205–208, 1965.

[19] S. Sato. Evaluation of the first-passage time probability to a square root boundary for the Wiener process. *Journal of Applied Probability*, 14(4):850–856, 1977.

[20] E. Sperner. Ein Satz über Untermengen einer endlichen Menge. *Math. Z.*, 27(1):544–548, 1928.

[21] R. P. Stanley. Weyl groups, the hard Lefschetz theorem, and the Sperner property. *SIAM J. Algebraic Discrete Methods*, 1(2):168–184, 1980.

[22] D. W. Stroock. *Probability Theory: an analytic view.* Cambridge University Press, second edition, 2011.

[23] B. D. Sullivan. On a conjecture of Andrica and Tomescu. *J. Integer Seq.*, 16(3):Article 13.3.1, 6, 2013.