ANOMALOUS RECURRENCE PROPERTIES OF MANY-DIMENSIONAL ZERO-DRIFT RANDOM WALKS

BY NICHOLAS GEORGIOU, MIKHAIL V. MENSHIKOV, ALEKSANDAR MIJATOVIĆ AND ANDREW R. WADE

Abstract

Famously, a d-dimensional, spatially homogeneous random walk whose increments are nondegenerate, have finite second moments, and have zero mean is recurrent if $d \in \{1, 2\}$, but transient if $d \geq 3$. Once spatial homogeneity is relaxed, this is no longer true. We study a family of zero-drift spatially nonhomogeneous random walks (Markov processes) whose increment covariance matrix is asymptotically constant along rays from the origin, and which, in any ambient dimension $d \geq 2$, can be adjusted so that the walk is either transient or recurrent. Natural examples are provided by random walks whose increments are supported on ellipsoids that are symmetric about the ray from the origin through the walk's current position; these *elliptic random walks* generalize the classical homogeneous Pearson–Rayleigh walk (the spherical case). Our proof of the recurrence classification is based on fundamental work of Lamperti.

Keywords: Nonhomogeneous random walk; elliptic random walk; zero drift; recurrence; transience

2010 Mathematics Subject Classification: Primary 60J05
Secondary 60J10; 60G42; 60G50

1. Introduction

A d-dimensional random walk that proceeds via a sequence of unit-length steps, each in an independent and uniformly random direction, is sometimes called a *Pearson–Rayleigh* random walk (PRRW), after the exchange in the letters pages of *Nature* between Karl Pearson and Lord Rayleigh in 1905 [17]. Pearson was interested in two dimensions and questions of migration of species (such as mosquitoes) [16], although Carazza has speculated that Pearson was a golfer [3, p. 419]; Rayleigh had earlier considered the acoustic 'random walks' in phase space produced by combinations of sound waves of the same amplitude and random phases.

The PRRW can be represented via partial sums of sequences of independent and identically distributed (i.i.d.) random vectors that are uniformly distributed on the unit sphere $\mathbb{S}^{d-1}$ in $\mathbb{R}^d$. Clearly, the increments have mean zero, i.e. the PRRW has *zero drift*. The PRRW has received some renewed interest recently as a model for microbe locomotion [1], [14], [15]. Hughes [8, Chapter 2] gave a general discussion of these walks, which have been well understood for many years. In particular, it is well known that the PRRW is recurrent for $d \in \{1, 2\}$ and transient if $d \geq 3$.

Suppose that we replace the spherically symmetric increments of the PRRW by increments that instead have some *elliptical* structure, while retaining the zero drift. For example, one could take the increments to be uniformly distributed on the surface of an ellipsoid of fixed shape and orientation, as represented by the right-hand picture of Figure 1. More generally, one

*doi:*10.1017/apr.2016.44 © *Applied Probability Trust* 2016

https://doi.org/10.1017/apr.2016.44 Published online by Cambridge University Press

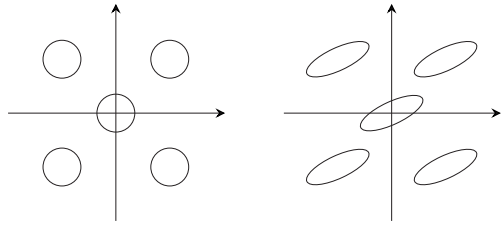


FIGURE 1: Pictorial representation of spatially homogeneous random walks with increments distributed on a fixed circle (*left*) and a fixed ellipse (*right*).

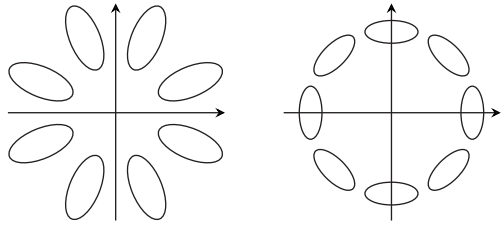


FIGURE 2: Pictorial representation of spatially nonhomogeneous random walks with increments distributed on a radially-aligned ellipse with major axis aligned in the radial sense (*left*) and in the transverse sense (*right*).

should view the ellipses in Figure 1 as representing the *covariance* structure of the increments of the walk (we will give a concrete example later; the uniform distribution on the ellipse is actually not the most convenient for calculations).

A little thought shows that the walk represented by the right-hand picture of Figure 1 is essentially no different to the PRRW: an affine transformation of $\mathbb{R}^d$ will map the walk back to a walk whose increments have the same covariance structure as the PRRW. To obtain genuinely different behaviour, it is necessary to abandon spatial homogeneity.

In this paper we consider a family of spatially *nonhomogeneous* random walks with zero drift. These include generalizations of the PRRW in which the increments are not i.i.d. but have a distribution supported on an ellipsoid of fixed size and shape, and whose orientation depends upon the current position of the walk. In Figure 2 we give representations of two important types of example, in which the ellipsoid is aligned so that its principal axes are parallel or perpendicular to the vector of the current position of the walk, which sits at the centre of the ellipse.

The random walks represented by Figure 2 are no longer sums of i.i.d. variables. These modified walks can behave very differently to the PRRW. For instance, one of the two-dimensional random walks represented in Figure 2 is *transient* while the other (as in the classical case) is recurrent. The reader who has not seen this kind of example before may take a moment to identify which is which. It is this *anomalous recurrence behaviour* that is the main subject of the present paper. In the next section, we give a formal description of our model and state our main results.

We end this introduction with a brief comment on motivation. In biology, the PRRW is more natural than a lattice-based walk for modelling the motion of microscopic organisms, such as certain bacteria, on a surface. Experiment suggests that the locomotion of several kinds of cells consists of roughly straight-line segments linked by discrete changes in direction;

https://doi.org/10.1017/apr.2016.44 Published online by Cambridge University Press

see, e.g. [14] and [15]. The generalization to elliptically distributed increments studied here represents movement on a surface on which either radial or transverse motion is inhibited. In chemistry and physics, the trajectory of a finite-step PRRW (also called a 'random chain') is an idealized model of the growth of weakly interacting polymer molecules; see, e.g. Section 2.6 of [8]. The modification to ellipsoid-supported jumps represents polymer growth in a biased medium.

2. Model and main results

We work in $\mathbb{R}^d$, where $d \geq 1$. Our main interest is in $d \geq 2$, as we shall explain shortly. Write $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_d$ for the standard orthonormal basis vectors in $\mathbb{R}^d$. Write $\mathbf{0}$ for the origin in $\mathbb{R}^d$, and let $\|\cdot\|$ denote the Euclidean norm and $\langle \cdot, \cdot \rangle$ the Euclidean inner product on $\mathbb{R}^d$. Write $\mathbb{S}^{d-1} := \{\boldsymbol{u} \in \mathbb{R}^d : \|\boldsymbol{u}\| = 1\}$ for the unit sphere in $\mathbb{R}^d$. For $\boldsymbol{x} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$, set $\hat{\boldsymbol{x}} := \boldsymbol{x}/\|\boldsymbol{x}\|$; also, set $\hat{\mathbf{0}} := \boldsymbol{e}_1$, for convenience. For definiteness, vectors $\boldsymbol{x} \in \mathbb{R}^d$ are viewed as column vectors throughout.

We now define $X = (X_n,\ n \in \mathbb{Z}_+)$, a discrete-time, time-homogeneous Markov process on a (nonempty, unbounded) subset $\mathbb{X}$ of $\mathbb{R}^d$. Formally, $(\mathbb{X}, \mathcal{B}_{\mathbb{X}})$ is a measurable space, $\mathbb{X}$ is a Borel subset of $\mathbb{R}^d$, and $\mathcal{B}_{\mathbb{X}}$ is the σ-algebra of all $B \cap \mathbb{X}$ for B a Borel set in $\mathbb{R}^d$. Suppose that X_0 is some fixed (i.e. nonrandom) point in $\mathbb{X}$. Write

$$\Delta_n := X_{n+1} - X_n, \qquad n \in \mathbb{Z}_+,$$

for the increments of X. By assumption, given $X_0, \ldots, X_n$, the law of Δ_n depends only on X_n (and not on n); so often we ease notation by taking $n = 0$ and writing just Δ for Δ_0. We also use the shorthand $\mathbb{P}_{\boldsymbol{x}}[\cdot] = \mathbb{P}[\cdot \mid X_0 = \boldsymbol{x}]$ for probabilities when the walk is started from $\boldsymbol{x} \in \mathbb{X}$; similarly, we use $\mathbb{E}_{\boldsymbol{x}}$ for the corresponding expectations.

We make the following moments assumption.

(A0) There exists $p > 2$ such that $\sup_{\boldsymbol{x} \in \mathbb{X}} \mathbb{E}_{\boldsymbol{x}}[\|\Delta\|^p] < \infty$.

Assumption (A0) ensures that Δ has a well-defined mean vector $\mu(\boldsymbol{x}) := \mathbb{E}_{\boldsymbol{x}}[\Delta]$, and we suppose that the random walk has *zero drift*:

(A1) $\mu(\boldsymbol{x}) = \mathbf{0}$ for all $\boldsymbol{x} \in \mathbb{X}$.

Assumption (A0) also ensures that Δ has a well-defined covariance matrix, which we denote by $M(\boldsymbol{x}) := \mathbb{E}_{\boldsymbol{x}}[\Delta\Delta^\top]$, where Δ is viewed as a column vector. To rule out pathological cases, we assume that Δ is *uniformly nondegenerate*, in the following sense.

(A2) There exists $v > 0$ such that $\operatorname{tr}(M(\boldsymbol{x})) = \mathbb{E}_{\boldsymbol{x}}[\|\Delta\|^2] \geq v$ for all $\boldsymbol{x} \in \mathbb{X}$.

Note that assumption (A2) is weaker than *uniform ellipticity*, which in this context usually means, for some $\varepsilon > 0$, $\mathbb{P}_{\boldsymbol{x}}[\langle \Delta, \boldsymbol{u} \rangle \geq \varepsilon] \geq \varepsilon$ for all $\boldsymbol{u} \in \mathbb{S}^{d-1}$ and all $\boldsymbol{x} \in \mathbb{X}$.

Our main interest is in a recurrence classification. First, we state the following basic 'nonconfinement' result.

Proposition 2.1. *Suppose that X satisfies assumptions (A0), (A1), and (A2). Then*

$$\limsup_{n \to \infty} \|X_n\| = +\infty \quad \textit{a.s. (almost surely).} \tag{2.1}$$

We give the proof of Proposition 2.1 in Section 4; we actually prove more, namely that the hypotheses of Proposition 2.1 ensure that a version of Kolmogorov's 'other' inequality holds.

https://doi.org/10.1017/apr.2016.44 Published online by Cambridge University Press

Fact (2.1) ensures that questions of the escape of trajectories to infinity are nontrivial. Indeed, we will give conditions under which one or other of the following two behaviours (which are not *a priori* exhaustive) occurs:

- $\lim_{n\to\infty}\|X_n\| = +\infty$ a.s., in which case we say that X is *transient*;
- $\liminf_{n\to\infty}\|X_n\| \le r_0$ a.s. for some constant $r_0 \in \mathbb{R}_+$, in which case we say that X is *recurrent*.

If X is an irreducible time-homogeneous Markov chain on a locally finite state space, these definitions reduce to the usual notions of transience and recurrence; in general state spaces, our approach allows us to avoid unnecessary technicalities concerning irreducibility.

In dimension $d = 1$, it is a consequence of the classical Chung–Fuchs theorem (see [4] or Chapter 9 of [9]) that a spatially *homogeneous* random walk with zero drift is necessarily recurrent. However, this is *not* true for a spatially nonhomogeneous random walk: as observed by Rogozin and Foss [19], a counterexample is provided by a version of the 'oscillating random walk' of Kemperman [10] in which the increment law is one of two distributions (with mean zero but infinite second moment) depending on the walk's present sign. Our conditions exclude these heavy-tailed phenomena, so that in $d = 1$ recurrence is assured in our setting.

Theorem 2.1. *Suppose that $d = 1$. Suppose that X satisfies assumptions (A0), (A1), and (A2). Then X is recurrent.*

Theorem 2.1 is essentially contained in a result of Lamperti [12, Theorem 3.2]; we give a self-contained proof below. Theorem 2.1 shows that in $d = 1$, under mild conditions, the classical Chung–Fuchs recurrence classification for homogeneous zero-drift random walks extends to zero-drift nonhomogeneous random walks. The purpose of the present paper is to demonstrate a natural family of examples in dimension $d \ge 2$ where this extension fails, and hence exhibit the following.

Fact. *There exist spatially nonhomogeneous random walks whose increments are nondegenerate, have uniformly bounded second moments, and have zero mean, which are*

- *transient in $d = 2$;*
- *recurrent in $d \ge 3$.*

Although certainly appreciated by experts, this fact is perhaps not as widely known as it might be. Zeitouni [20, pp. 91–92] described an example of a transient zero-drift random walk on $\mathbb{Z}^2$, and stated that the idea 'goes back to Krylov (in the context of diffusions)'. Peres *et al.* [18] investigated the minimal number of different increment distributions required for anomalous recurrence behaviour.

We now introduce our family of nonhomogeneous random walks. Write $\|\cdot\|_{\mathrm{op}}$ for the matrix (operator) norm given by $\|M\|_{\mathrm{op}} = \sup_{\boldsymbol{u}\in\mathbb{S}^{d-1}}\|M\boldsymbol{u}\|$. The following assumption on the asymptotic stability of the covariance structure of the process along rays is central.

(A3) Suppose that there exists a positive-definite matrix-valued function σ^2 with domain $\mathbb{S}^{d-1}$ such that, as $r \to \infty$,

$$\varepsilon(r) := \sup_{\{\boldsymbol{x}\in\mathbb{X}:\ \|\boldsymbol{x}\|\ge r\}} \|M(\boldsymbol{x}) - \sigma^2(\hat{\boldsymbol{x}})\|_{\mathrm{op}} \to 0.$$

A little informally, (A3) says that $M(\boldsymbol{x}) \to \sigma^2(\hat{\boldsymbol{x}})$ as $\|\boldsymbol{x}\| \to \infty$; in what follows, we will often make similar statements, formal versions of which may be cast as in (A3).

https://doi.org/10.1017/apr.2016.44 Published online by Cambridge University Press

Note that (A2) and (A3) together imply that $\mathrm{tr}(\sigma^2(\boldsymbol{u})) \geq v > 0$; next we impose a key assumption on the form of σ^2 that is considerably stronger. To describe this, it is convenient to introduce the notation $\langle \cdot, \cdot \rangle_{\boldsymbol{u}}$ that defines, for each $\boldsymbol{u} \in \mathbb{S}^{d-1}$, an inner product on $\mathbb{R}^d$ via

$$\langle \boldsymbol{y}, \boldsymbol{z} \rangle_{\boldsymbol{u}} := \boldsymbol{y}^\top \cdot \sigma^2(\boldsymbol{u}) \cdot \boldsymbol{z} = \langle \boldsymbol{y}, \sigma^2(\boldsymbol{u}) \cdot \boldsymbol{z} \rangle \quad \text{for } \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^d.$$

(A4) Suppose that there exist constants U and V with $0 < U < V < \infty$ such that, for all $\boldsymbol{u} \in \mathbb{S}^{d-1}$, $\langle \boldsymbol{u}, \boldsymbol{u} \rangle_{\boldsymbol{u}} = U$ and $\mathrm{tr}(\sigma^2(\boldsymbol{u})) = V$.

Informally, V quantifies the total variance of the increments, while U quantifies the variance in the radial direction; necessarily $U \leq V$. The assumption that $0 \neq U \neq V$ excludes some degenerate cases. As we will see, one possible way to satisfy condition (A4) is to suppose that the eigenvectors of $\sigma^2(\boldsymbol{u})$ are all parallel or perpendicular to the vector $\boldsymbol{u}$, and that the corresponding eigenvalues are all constant as $\boldsymbol{u}$ varies; the level sets of the corresponding quadratic forms $q_{\boldsymbol{u}}(\boldsymbol{x}) := \langle \boldsymbol{x}, \boldsymbol{x} \rangle_{\boldsymbol{u}}$ for $\boldsymbol{u} \in \mathbb{S}^{d-1}$ are then ellipsoids like those depicted in Figure 2.

Our main result is the following, which shows that both transience and recurrence are possible for *any* $d \geq 2$, depending on parameter choices; as seen in Theorem 2.1, this possibility of anomalous recurrence behaviour is a genuinely multidimensional phenomenon under our regularity conditions.

Theorem 2.2. *Suppose that X satisfies (A0)–(A4), with constants $0 < U < V$ as defined in (A4). The following recurrence classification is valid.*

(i) *If $2U < V$ then X is transient.*

(ii) *If $2U > V$ then X is recurrent.*

(iii) *If $2U = V$ and (A3) holds with $\varepsilon(r) = \mathrm{O}(r^{-\delta_0})$ for some $\delta_0 > 0$, then X is recurrent.*

Moreover, we show that in any of the above cases, X is *null* in the following sense.

Theorem 2.3. *Suppose that X satisfies (A0)–(A4), with constants $0 < U < V$ as defined in (A4). Then in either case (i)–(iii) of Theorem 2.2, for any bounded $A \subset \mathbb{R}^d$,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbf{1}\{X_k \in A\} = 0 \quad \text{a.s. and in } L^q \text{ for any } q \geq 1. \tag{2.2}$$

Remark 2.1. Theorems 2.2 and 2.3 both remain valid if we permit $V = U > 0$ in (A4); indeed, the condition $U < V$ is not used in the proof of Theorem 2.2 given below, so this case is recurrent, by Theorem 2.2(ii). The condition $U < V$ is used at one point to simplify the proof of Theorem 2.3 given below, but a small modification of the argument also works in the case $U = V$. (See Remark 6.1 below for more details.)

The remainder of the paper is organised as follows. In Section 3 we describe a specific family of examples called *elliptic random walk models* that satisfy assumptions (A0)–(A4), and exhibit both transient and recurrent behaviour dependent on the parameters of the model. We also present some simulated data that depict the random walks in both cases. In Section 4 we prove a d-dimensional martingale version of Kolmogorov's other inequality and use that to prove the nonconfinement result (Proposition 2.1). In Section 5 we prove the recurrence classification (Theorem 2.2), and in Section 6 we prove Theorem 2.3. In the appendix we prove recurrence in the one-dimensional case (Theorem 2.1).

https://doi.org/10.1017/apr.2016.44 Published online by Cambridge University Press

Finally, we remark that in work in progress we investigate diffusive scaling limits for random walks of the type described in the present paper; the diffusions that appear as scaling limits possess certain pathologies from the point of view of diffusion theory that make them interesting in their own right.

3. Example: elliptic random walk model

Let $d \geq 2$. We describe a specific model on $\mathbb{X} = \mathbb{R}^d$ where the jump distribution at $\boldsymbol{x} \in \mathbb{R}^d$ is supported on an ellipsoid having one distinguished axis aligned with the vector $\boldsymbol{x}$. The model is specified by two constants $a, b > 0$. Construct Δ as follows. Given $X_0 = \boldsymbol{x}$, take $\boldsymbol{\zeta}$ uniform on $\mathbb{S}^{d-1}$ and set

$$\Delta = Q_{\hat{\boldsymbol{x}}} D\boldsymbol{\zeta} \tag{3.1}$$

for $Q_{\hat{\boldsymbol{x}}}$ an orthogonal matrix representing a transformation of $\mathbb{R}^d$ mapping $\boldsymbol{e}_1$ to $\hat{\boldsymbol{x}}$, and $D = \sqrt{d}\,\mathrm{diag}(a, b, \ldots, b)$. See Figure 3. (Recall that $\hat{\mathbf{0}} = \boldsymbol{e}_1$, so, for $X_0 = \mathbf{0}$, we can take $Q_{\hat{\boldsymbol{x}}} = I$ and $\Delta = D\boldsymbol{\zeta}$.) Thus, Δ is a random point on an ellipsoid that has one distinguished semi-axis, of length $a\sqrt{d}$, aligned in the $\hat{\boldsymbol{x}}$ direction, and all other semi-axes of length $b\sqrt{d}$. Note that the law of Δ is well defined owing to the spherical symmetry of the uniform distribution on $\mathbb{S}^{d-1}$ and the fact that only one axis of the ellipsoid is distinguished (for this reason, it is enough to take any $Q_{\hat{\boldsymbol{x}}}$ satisfying $Q_{\hat{\boldsymbol{x}}}\boldsymbol{e}_1 = \hat{\boldsymbol{x}}$ in order to define Δ; see also Remark 3.1 below).

Note also that Δ is not chosen to be uniformly distributed on the surface of the ellipsoid; this does not affect the range of asymptotic behaviour exhibited by the family of walks as a and b vary, but it does simplify the calculation of $M(\boldsymbol{x})$. Indeed, we have

$$M(\boldsymbol{x}) = \mathbb{E}_{\boldsymbol{x}}[\Delta\Delta^\top] = \mathbb{E}[Q_{\hat{\boldsymbol{x}}} D\boldsymbol{\zeta}\boldsymbol{\zeta}^\top D Q_{\hat{\boldsymbol{x}}}^\top] = Q_{\hat{\boldsymbol{x}}} D\,\mathbb{E}[\boldsymbol{\zeta}\boldsymbol{\zeta}^\top] D Q_{\hat{\boldsymbol{x}}}^\top = \frac{1}{d} Q_{\hat{\boldsymbol{x}}} D^2 Q_{\hat{\boldsymbol{x}}}^\top,$$

by linearity of expectation, and using the fact that $\mathbb{E}[\boldsymbol{\zeta}\boldsymbol{\zeta}^\top] = d^{-1} I$ for $\boldsymbol{\zeta}$ uniformly distributed on $\mathbb{S}^{d-1}$. Also, a calculation similar to the above confirms that $\mu(\boldsymbol{x}) = \mathbf{0}$ for all $\boldsymbol{x} \in \mathbb{R}^d$, since $\mathbb{E}[\boldsymbol{\zeta}] = \mathbf{0}$.

Since $\|\Delta\|$ is bounded above by $\sqrt{d}\max\{a, b\}$, assumption (A0) holds. Clearly, (A1) and (A3) hold, with $\sigma^2(\boldsymbol{u}) = d^{-1} Q_{\boldsymbol{u}} D^2 Q_{\boldsymbol{u}}^\top$ for $\boldsymbol{u} \in \mathbb{S}^{d-1}$. It is also a simple matter to check that (A2) and (A4) hold: the matrix $\sigma^2(\boldsymbol{u})$ represented in coordinates for the orthonormal basis $\{Q_{\boldsymbol{u}}\boldsymbol{e}_1 = \boldsymbol{u},\ Q_{\boldsymbol{u}}\boldsymbol{e}_2, \ldots, Q_{\boldsymbol{u}}\boldsymbol{e}_d\}$ is diagonal with entries $a^2, b^2, \ldots, b^2$. Indeed,

$$\sigma^2(\boldsymbol{u}) = \frac{1}{d} Q_{\boldsymbol{u}} D^2 Q_{\boldsymbol{u}}^\top = Q_{\boldsymbol{u}}[b^2 I + (a^2 - b^2)\boldsymbol{e}_1\boldsymbol{e}_1^\top] Q_{\boldsymbol{u}}^\top = a^2 \boldsymbol{u}\boldsymbol{u}^\top + b^2(I - \boldsymbol{u}\boldsymbol{u}^\top),$$

and, therefore, $\langle \boldsymbol{u}, \boldsymbol{u} \rangle_{\boldsymbol{u}} = \langle \boldsymbol{u}, \sigma^2(\boldsymbol{u}) \cdot \boldsymbol{u} \rangle = a^2 > 0$ for all $\boldsymbol{u} \in \mathbb{S}^{d-1}$, and $\mathrm{tr}\,(M(\boldsymbol{x})) = \mathrm{tr}\,(\sigma^2(\hat{\boldsymbol{x}})) = a^2 + (d-1)b^2 > 0$ for all $\boldsymbol{x} \in \mathbb{R}^d$.

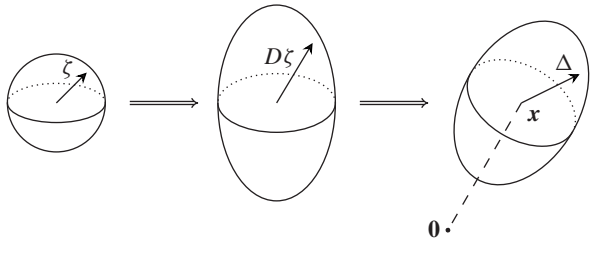


FIGURE 3: Definition of $\Delta = Q_{\hat{\boldsymbol{x}}} D\boldsymbol{\zeta}$.

https://doi.org/10.1017/apr.2016.44 Published online by Cambridge University Press

Remark 3.1. The seeming ambiguity in the definition of Δ due to the choice of $Q_{\hat{\boldsymbol{x}}}$ can be resolved by noting that Δ can be rewritten as

$$\Delta = Q_{\hat{\boldsymbol{x}}} D Q_{\hat{\boldsymbol{x}}}^\top Q_{\hat{\boldsymbol{x}}} \boldsymbol{\zeta} = Q_{\hat{\boldsymbol{x}}} D Q_{\hat{\boldsymbol{x}}}^\top \tilde{\boldsymbol{\zeta}},$$

where $\tilde{\boldsymbol{\zeta}} = Q_{\hat{\boldsymbol{x}}} \boldsymbol{\zeta}$ is also uniform on $\mathbb{S}^{d-1}$ (this follows from the spherical symmetry of the uniform distribution on $\mathbb{S}^{d-1}$). Moreover, the symmetric matrix $H_{\hat{\boldsymbol{x}}} := Q_{\hat{\boldsymbol{x}}} D Q_{\hat{\boldsymbol{x}}}^\top$ is determined explicitly in terms of $\hat{\boldsymbol{x}}$:

$$H_{\hat{\boldsymbol{x}}} = Q_{\hat{\boldsymbol{x}}} D Q_{\hat{\boldsymbol{x}}}^\top = Q_{\hat{\boldsymbol{x}}} (b\sqrt{d} I + (a-b)\sqrt{d} \boldsymbol{e}_1 \boldsymbol{e}_1^\top) Q_{\hat{\boldsymbol{x}}}^\top = b\sqrt{d} I + (a-b)\sqrt{d} \hat{\boldsymbol{x}} \hat{\boldsymbol{x}}^\top.$$

Consequently, we could choose to specify Δ explicitly as

$$\Delta = H_{\hat{\boldsymbol{x}}} \tilde{\boldsymbol{\zeta}} = b\sqrt{d} \tilde{\boldsymbol{\zeta}} + (a-b)\sqrt{d} \hat{\boldsymbol{x}} \langle \hat{\boldsymbol{x}}, \tilde{\boldsymbol{\zeta}} \rangle,$$

with $\tilde{\boldsymbol{\zeta}}$ taken to be uniform on $\mathbb{S}^{d-1}$. As before, we find that $\mathbb{E}_{\boldsymbol{x}}[\Delta] = H_{\hat{\boldsymbol{x}}} \mathbb{E}[\tilde{\boldsymbol{\zeta}}] = \mathbf{0}$ and

$$\mathbb{E}_{\boldsymbol{x}}[\Delta \Delta^\top] = H_{\hat{\boldsymbol{x}}} \mathbb{E}[\tilde{\boldsymbol{\zeta}} \tilde{\boldsymbol{\zeta}}^\top] H_{\hat{\boldsymbol{x}}} = \frac{1}{d} H_{\hat{\boldsymbol{x}}}^2 = a^2 \hat{\boldsymbol{x}} \hat{\boldsymbol{x}}^\top + b^2 (I - \hat{\boldsymbol{x}} \hat{\boldsymbol{x}}^\top).$$

Recall that we assume our random walk to be time homogeneous, so that (3.1) in fact determines the distribution of Δ_n for all $n \geq 0$. Formally, we define $\boldsymbol{\zeta}_0, \boldsymbol{\zeta}_1, \ldots$ to be a sequence of independent random variables uniformly distributed on $\mathbb{S}^{d-1}$, and, for each $n \geq 0$, we define Δ_n conditional on $\{X_n = \boldsymbol{x}\}$ via

$$\Delta_n = Q_{\hat{\boldsymbol{x}}} D \boldsymbol{\zeta}_n. \tag{3.2}$$

We call $X = (X_n,\ n \in \mathbb{Z}_+)$ defined in this way an *elliptic random walk*.

As a corollary to Theorems 2.2 and 2.3, we obtain the following recurrence classification for the elliptic random walk model. For this model, the $\varepsilon(r)$ in (A3) is identically zero, so we get a complete classification that includes the boundary case.

Corollary 3.1. *Let $d \geq 2$ and $a, b \in (0, \infty)$. Let X be an elliptic random walk on $\mathbb{R}^d$. Then X is transient if $a^2 < (d-1)b^2$ and null recurrent if $a^2 \geq (d-1)b^2$.*

In two dimensions we can explicitly describe the random walk as follows. For $\boldsymbol{x} \in \mathbb{R}^2$, $\boldsymbol{x} \neq \mathbf{0}$ with $\boldsymbol{x} = (x_1, x_2)$ in Cartesian components, set $\boldsymbol{x}^\perp := (-x_2, x_1)$. Fix $a, b \in (0, \infty)$. Let $E_{\boldsymbol{x}}(a, b)$ denote the ellipse with centre $\boldsymbol{x}$ and principal axes aligned in the $\boldsymbol{x}, \boldsymbol{x}^\perp$ directions, with lengths $2\sqrt{2}a$, $2\sqrt{2}b$, respectively, given in parametrized form by

$$E_{\boldsymbol{x}}(a, b) := \left\{ \boldsymbol{x} + \sqrt{2}a \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|} \cos\phi + \sqrt{2}b \frac{\boldsymbol{x}^\perp}{\|\boldsymbol{x}\|} \sin\phi \colon \phi \in (-\pi, \pi] \right\}, \tag{3.3}$$

and, for $\boldsymbol{x} = \mathbf{0}$, set $E_{\mathbf{0}}(a, b) := \{\sqrt{2}a\, \boldsymbol{e}_1 \cos\phi + \sqrt{2}b\, \boldsymbol{e}_2 \sin\phi \colon \phi \in (-\pi, \pi]\}$. The parameter ϕ in parametrization (3.3) should be interpreted with caution: it is *not*, in general, the central angle of the parametrized point on the ellipse.

Given $X_n = \boldsymbol{x} \in \mathbb{R}^2$, X_{n+1} is taken to be distributed on $E_{\boldsymbol{x}}(a, b)$, 'uniformly' with respect to parametrization (3.3). Precisely, let $\phi_0, \phi_1, \ldots$ be a sequence of independent random variables uniformly distributed on $(-\pi, \pi]$. Then, on $\{X_n \neq \mathbf{0}\}$,

$$X_{n+1} = X_n + \sqrt{2}a \frac{X_n}{\|X_n\|} \cos\phi_n + \sqrt{2}b \frac{X_n^\perp}{\|X_n\|} \sin\phi_n, \tag{3.4}$$

while, on $\{X_n = \mathbf{0}\}$, $X_{n+1} = (\sqrt{2}a \cos\phi_n, \sqrt{2}b \sin\phi_n)$.

https://doi.org/10.1017/apr.2016.44 Published online by Cambridge University Press

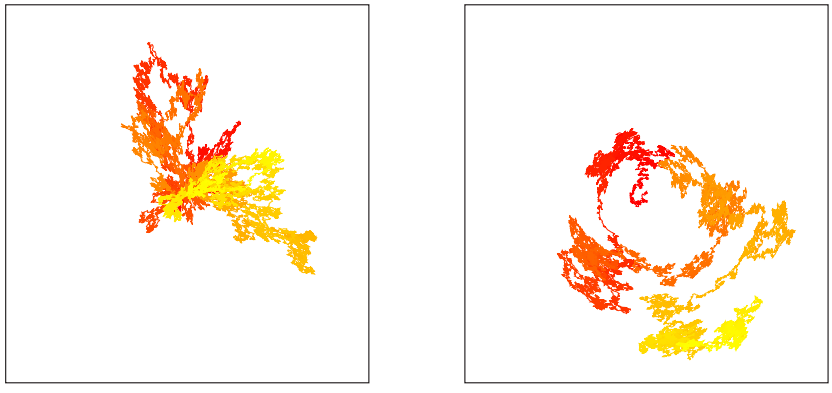


FIGURE 4: Simulation of the elliptic random walk in $\mathbb{R}^2$ for the recurrent case $a > b$ (*left*) and the transient case $a < b$ (*right*).

In Figure 4 we present two sample paths of a simulation of the elliptic random walk in $\mathbb{R}^2$ in the two cases of recurrence and transience. In each picture the walk starts at the origin at the centre of the picture; time is represented by the variation in colour (from red to yellow, or from dark to light if viewed in grey scale).

Remarks 3.2. (a) The process X reduces to the classical PRRW when $a = b$; in this case it is spatially homogeneous, i.e. the distribution of the increment $X_{n+1} - X_n$ does not depend on X_n. For $a \neq b$, the random walk is not spatially homogeneous, and the jump distribution depends upon the projection onto the unit sphere of the walk's current position.

(b) As mentioned earlier, we choose to take increments as defined at (3.4), rather than increments that are uniform on the ellipse with respect to the one-dimensional Lebesgue measure on $E_{\boldsymbol{x}}(a, b)$, purely for computational reasons. In fact, in two dimensions, since the Lebesgue measure on $E_{\boldsymbol{x}}(a, b)$ coincides with the measure induced by taking ϕ uniformly distributed on $(-\pi, \pi]$ when $a = b$, and the case $a = b$ is critically recurrent, the qualitative behaviour will be the same in either case: the walk will be transient for $a < b$ and recurrent for $a \geq b$. For higher dimensions, taking increments that are uniform with respect to the Lebesgue measure on $E^d_{\boldsymbol{x}}(a, b) := \{Q_{\hat{\boldsymbol{x}}} D\boldsymbol{u} : \boldsymbol{u} \in \mathbb{S}^{d-1}\}$ will still specify a family of models that exhibit a phase transition, from transience (for small a/b) to recurrence (for large a/b), but the exact shape of the ellipsoid in the critical case (i.e. the smallest ratio a/b for which the walk is recurrent) may be different.

(c) It follows from (3.2) that

$$\begin{aligned} \|X_{n+1}\|^2 &= \|X_n\|^2 + 2\|X_n\|\langle \widehat{X}_n, \Delta_n\rangle + \|\Delta_n\|^2 \\ &= \|X_n\|^2 + 2\|X_n\|\langle \boldsymbol{e}_1, D\boldsymbol{\zeta}_n\rangle + \langle \boldsymbol{\zeta}_n, D^2\boldsymbol{\zeta}_n\rangle \\ &= \|X_n\|^2 + 2a\sqrt{d}\|X_n\|\langle \boldsymbol{e}_1, \boldsymbol{\zeta}_n\rangle + (a^2 - b^2)d\langle \boldsymbol{e}_1, \boldsymbol{\zeta}_n\rangle^2 + b^2 d. \end{aligned} \tag{3.5}$$

In particular, for this family of models, $(\|X_n\|,\ n \in \mathbb{Z}_+)$ is itself a Markov process, since the distribution of $\|X_{n+1}\|$ depends only on $\|X_n\|$ and not X_n; however, in the general setting of Section 2, this need not be the case.

One-dimensional processes with evolutions reminiscent to those given by (3.5) have been studied previously by Kingman [11] and Bingham [2]. Those processes can be respectively viewed as the distance from the start point to the current position of a random walk in Euclidean

https://doi.org/10.1017/apr.2016.44 Published online by Cambridge University Press

space, and the geodesic distance from the start point to the current position of a random walk on the surface of a sphere, but in both cases the increments of the random walk have the property that the distribution of the jump vector is a product of the independent marginal distributions of the length and direction of the jump vector. In contrast, for the elliptic random walk, the laws of $\|\Delta_n\|$ and $\langle \widehat{X}_n, \widehat{\Delta}_n \rangle$ are *not* independent (except when $a = b$).

(d) The theory equally applies to the case where the ellipsoid specifying the jump distribution is oriented with some fixed angle $\alpha \in [0, \pi)$ with respect to the radial direction. If we define $\Delta = Q^{\alpha}_{\hat{\boldsymbol{x}}} D\boldsymbol{\zeta}$, where $Q^{\alpha}_{\hat{\boldsymbol{x}}}$ is an orthogonal matrix that maps $\boldsymbol{e}_\alpha := \boldsymbol{e}_1 \cos\alpha + \boldsymbol{e}_2 \sin\alpha$ to $\hat{\boldsymbol{x}}$, then we find that transience of X is equivalent to $(a^2 - b^2)\cos 2\alpha < (d-2)b^2$.

Note that, for $d = 2$, $Q^{\alpha}_{\hat{\boldsymbol{x}}}$ and therefore Δ are well defined, but this is not so for higher dimensions. Nevertheless, for *any* collection of matrices $(Q^{\alpha}_{\boldsymbol{u}} : \boldsymbol{u} \in \mathbb{S}^{d-1})$ satisfying $Q^{\alpha}_{\boldsymbol{u}} \boldsymbol{e}_\alpha = \boldsymbol{u}$ for all $\boldsymbol{u} \in \mathbb{S}^{d-1}$, we get the same recurrence classification. This is because the distribution of $\|X_{n+1}\|$ given X_n is determined through the angle α, since $\|X_{n+1}\|^2$ is equal to

$$\|X_n\|^2 + 2\sqrt{d}\|X_n\|(a\langle \boldsymbol{e}_1, \boldsymbol{\zeta}_n\rangle \cos\alpha + b\langle \boldsymbol{e}_2, \boldsymbol{\zeta}_n\rangle \sin\alpha) + (a^2 - b^2)d\langle \boldsymbol{e}_1, \boldsymbol{\zeta}_n\rangle^2 + b^2 d,$$

and therefore assumption (A4) holds with $U = a^2\cos^2\alpha + b^2\sin^2\alpha$ and $V = a^2 + (d-1)b^2$.

4. Nonconfinement

In this section we prove that assumptions (A0), (A1), and (A2) imply that $\limsup_{n\to\infty}\|X_n\| = +\infty$ a.s. We first present a general result for martingales on $\mathbb{R}^d$ satisfying a 'uniform dispersion' condition; the result can be viewed as a d-dimensional martingale version of *Kolmogorov's other inequality* (see, e.g. [5, pp. 123, 502]).

Lemma 4.1. *Let $d \in \mathbb{N}$. Suppose that $(Y_n,\ n \in \mathbb{Z}_+)$ is an $\mathbb{R}^d$-valued process adapted to a filtration $(\mathcal{G}_n,\ n \in \mathbb{Z}_+)$, with $\mathbb{P}[Y_0 = \mathbf{0} \mid \mathcal{G}_0] = 1$. Suppose that there exist $p > 2$, $v > 0$, and $B < \infty$ such that, for all $n \in \mathbb{Z}_+$, a.s.,*

$$\mathbb{E}[\|Y_{n+1} - Y_n\|^p \mid \mathcal{G}_n] \le B, \tag{4.1}$$

$$\mathbb{E}[\|Y_{n+1} - Y_n\|^2 \mid \mathcal{G}_n] \ge v, \tag{4.2}$$

$$\mathbb{E}[Y_{n+1} - Y_n \mid \mathcal{G}_n] = \mathbf{0}. \tag{4.3}$$

Then there exists $D < \infty$, depending only on B, p, and v, such that, for all $n \in \mathbb{Z}_+$ and all $x \in \mathbb{R}_+$,

$$\mathbb{P}\Big[\max_{0\le\ell\le n}\|Y_\ell\| \ge x \;\Big|\; \mathcal{G}_0\Big] \ge 1 - \frac{D(1+x)^2}{n} \quad \text{a.s.}$$

Proof. Let $x > 0$ and set $\tau = \min\{n \ge 0 : \|Y_n\| \ge x\}$; throughout the paper, we adopt the usual convention that $\min\varnothing := \infty$. In analogy with previous notation, write $\Delta_n = Y_{n+1} - Y_n$ for the jump distribution, set $\widehat{\Delta}_n = \Delta_n/\|\Delta_n\|$, and let

$$W_n = \begin{cases} Y_n & \text{if } \|Y_n\| \le A(1+x), \\ Y_{n-1} + \widehat{\Delta}_{n-1}(A-1)(1+x) & \text{if } \|Y_n\| > A(1+x), \end{cases}$$

where $A > 1$ is a constant to be specified later. Note that W_n is $\mathcal{G}_n$-measurable.

Now, on $\{\|Y_n\| \le x\}$, $W_n = Y_n$ and

$$\begin{aligned}\mathbb{E}[W_{n+1} - W_n \mid \mathcal{G}_n] &= \mathbb{E}[\Delta_n \mid \mathcal{G}_n] + \mathbb{E}[\widehat{\Delta}_n((A-1)(1+x) - \|\Delta_n\|)\mathbf{1}\{\|Y_{n+1}\| > A(1+x)\} \mid \mathcal{G}_n].\end{aligned}$$

https://doi.org/10.1017/apr.2016.44 Published online by Cambridge University Press

But $\{\|Y_{n+1}\| > A(1+x)\} \cap \{\|Y_n\| \le x\}$ implies that $\|\Delta_n\| > (A-1)(1+x)$, and, by (4.3), $\mathbb{E}[\Delta_n \mid \mathcal{G}_n] = \mathbf{0}$. Hence, on $\{\|Y_n\| \le x\}$,

$$\begin{aligned}\|\mathbb{E}[W_{n+1} - W_n \mid \mathcal{G}_n]\| &\le \mathbb{E}[\|\Delta_n\| \mathbf{1}\{\|\Delta_n\| > (A-1)(1+x)\} \mid \mathcal{G}_n] \\ &\le \frac{\mathbb{E}[\|\Delta_n\|^2 \mid \mathcal{G}_n]}{(A-1)(1+x)} \\ &\le \frac{B'}{(A-1)(1+x)} \quad \text{a.s.},\end{aligned}$$

where, by (4.1) and Lyapunov's inequality, $B' < \infty$ depends only on B and p. Hence, we can choose $A \ge A_0$ for some $A_0 = A_0(B, p, v)$ large enough so that

$$\|\mathbb{E}[W_{n+1} - W_n \mid \mathcal{G}_n]\| \le \frac{v}{8(1+x)} \quad \text{on } \{\|Y_n\| \le x\}. \tag{4.4}$$

Also, on $\{\|Y_n\| \le x\}$, by a similar argument, $\mathbb{E}[\|W_{n+1} - W_n\|^2 \mid \mathcal{G}_n]$ is equal to

$$\begin{aligned}&\mathbb{E}[\|\Delta_n\|^2 \mid \mathcal{G}_n] + \mathbb{E}[((A-1)^2(1+x)^2 - \|\Delta_n\|^2)\mathbf{1}\{\|Y_{n+1}\| > A(1+x)\} \mid \mathcal{G}_n] \\ &\quad \ge \mathbb{E}[\|\Delta_n\|^2 \mid \mathcal{G}_n] - \mathbb{E}[\|\Delta_n\|^2 \mathbf{1}\{\|\Delta_n\| > (A-1)(1+x)\} \mid \mathcal{G}_n] \\ &\quad \ge v - (A-1)^{2-p}(1+x)^{2-p}\,\mathbb{E}[\|\Delta_n\|^p \mid \mathcal{G}_n] \\ &\quad \ge \tfrac{1}{2}v\end{aligned} \tag{4.5}$$

for all $x \ge 0$ and $A \ge A_1$ for sufficiently large $A_1 = A_1(B, p, v)$, using (4.1) and (4.2).

Now, set $Z_n = \|W_{n\wedge\tau}\|^2$. Then, on $\{n < \tau\}$, by (4.4) and (4.5),

$$\begin{aligned}\mathbb{E}[Z_{n+1} - Z_n \mid \mathcal{G}_n] &= \mathbb{E}[\|W_{n+1}\|^2 - \|W_n\|^2 \mid \mathcal{G}_n] \\ &= \mathbb{E}[\|W_{n+1} - W_n\|^2 \mid \mathcal{G}_n] + 2\langle W_n, \mathbb{E}[W_{n+1} - W_n \mid \mathcal{G}_n]\rangle \\ &\ge \frac{v}{2} - \frac{2\|W_n\| v}{8(1+x)} \\ &\ge \frac{v}{2} - \frac{vx}{4(1+x)} \\ &\ge \tfrac{1}{4}v.\end{aligned}$$

Hence, $Z_n - \sum_{k=0}^{n-1} v_k$ is a $\mathcal{G}_n$-adapted submartingale, where

$$v_k = \tfrac{1}{4}v\mathbf{1}\{k < \tau\} \ge \tfrac{1}{4}v\mathbf{1}\{n < \tau\} \quad \text{for } 0 \le k < n.$$

By construction, $0 \le Z_n \le A^2(1+x)^2$, so

$$0 = \mathbb{E}[Z_0 \mid \mathcal{G}_0] \le \mathbb{E}[Z_n \mid \mathcal{G}_0] - \sum_{k=0}^{n-1} \mathbb{E}[v_k \mid \mathcal{G}_0] \le A^2(1+x)^2 - \sum_{k=0}^{n-1} \frac{v}{4}\mathbb{P}[n < \tau \mid \mathcal{G}_0],$$

which implies that $\frac{1}{4}nv\mathbb{P}[n < \tau \mid \mathcal{G}_0] \le A^2(1+x)^2$. In other words,

$$\mathbb{P}\Big[\max_{0\le\ell\le n} \|Y_\ell\| < x \;\Big|\; \mathcal{G}_0\Big] \le \frac{4A^2(1+x)^2}{vn} \quad \text{a.s.} \qquad \square$$

Proof of Proposition 2.1. It suffices to show that, for all $x \in \mathbb{R}_+$, the event $\{\|X_n\| \ge x\}$ occurs infinitely often. For a given x, we will apply Lemma 4.1 to $Y_n = X_{m+n} - X_m$ with $\mathcal{G}_n = \sigma(X_0, \dots, X_{m+n})$; this result is applicable since (A0), (A1), and (A2) respectively imply

https://doi.org/10.1017/apr.2016.44 Published online by Cambridge University Press

(4.1), (4.3), and (4.2). Thus, Lemma 4.1 shows that, for some finite $t = t(x)$,

$$\mathbb{P}\Big[\max_{0\le\ell\le t-1}\|X_{m+\ell} - X_m\| \ge 2x \;\Big|\; X_0, \dots, X_m\Big] \ge \frac{1}{2} \quad \text{a.s.} \tag{4.6}$$

for all $m \ge 0$. For $k = 1, 2, \dots$, define the event

$$A_k = \Big\{\max_{0\le\ell\le t-1}\|X_{(k-1)t+\ell} - X_{(k-1)t}\| \ge 2x\Big\},$$

and filtration $\mathcal{G}'_{k-1} = \sigma(X_0, \dots, X_{(k-1)t})$. Then $A_k \in \mathcal{G}'_k$, and, by (4.6), $\mathbb{P}[A_k \mid \mathcal{G}'_{k-1}] \ge \frac{1}{2}$ a.s. for all k. An application of Lévy's extension of the Borel–Cantelli lemma (see, e.g. [9, Corollary 7.20]) shows that A_k occurs infinitely often, a.s. For each k such that A_k occurs, either

- $\|X_{(k-1)t}\| \ge x$, or
- $\|X_{(k-1)t}\| \le x$ and $\|X_n\| \ge x$ for some $(k-1)t < n < kt$.

Since one of these cases must occur for infinitely many k, then $\{\|X_n\| \ge x\}$ occurs infinitely often, as required. □

5. Recurrence classification

In this section we study the random walk X_n and give the proof of the recurrence classification, Theorem 2.2. The method of proof is based on applying classical results of Lamperti [12] to the $\mathbb{R}_+$-valued radial process given by $R_n := \|X_n\|$. The method rests on an analysis of the increments $R_{n+1} - R_n$ given $X_n = \boldsymbol{x} \in \mathbb{X}$; in general, R_n is not itself a Markov process.

First we state a general result on the increments of R_n for a Markov process X_n on $\mathbb{X}$. Recall that we write $\Delta = X_1 - X_0$, and given $\boldsymbol{x} \ne \mathbf{0}$, we set

$$\Delta_{\boldsymbol{x}} := \frac{\langle \Delta, \boldsymbol{x}\rangle}{\|\boldsymbol{x}\|} = \langle \Delta, \hat{\boldsymbol{x}}\rangle,$$

so that $\Delta_{\boldsymbol{x}}$ is the length of the projection of Δ in the $\hat{\boldsymbol{x}}$ direction. In other words, $\Delta_{\boldsymbol{x}}$ is the radial (scalar) component of Δ at $X_0 = \boldsymbol{x}$; no confusion should arise with our notation $\Delta_n = X_{n+1} - X_n$ introduced previously.

We make an important comment on notation. When we write $\mathrm{O}(\|\boldsymbol{x}\|^{-1-\delta})$, and similar expressions, these are understood to be uniform in $\boldsymbol{x}$. That is, if $f\colon \mathbb{R}^d \to \mathbb{R}$ and $g\colon \mathbb{R}_+ \to \mathbb{R}_+$, we write $f(\boldsymbol{x}) = \mathrm{O}(g(\|\boldsymbol{x}\|))$ to mean that there exist $C \in \mathbb{R}_+$ and $r \in \mathbb{R}_+$ such that

$$|f(\boldsymbol{x})| \le C g(\|\boldsymbol{x}\|) \quad \text{for all } \boldsymbol{x} \in \mathbb{X} \text{ with } \|\boldsymbol{x}\| \ge r. \tag{5.1}$$

Lemma 5.1. *Suppose that X is a discrete-time, time-homogeneous Markov process on $\mathbb{X} \subseteq \mathbb{R}^d$ satisfying (A0) for some $p > 2$. Then, for $R_n := \|X_n\|$, we have*

$$\sup_{\boldsymbol{x}\in\mathbb{X}} \mathbb{E}[|R_{n+1} - R_n|^p \mid X_n = \boldsymbol{x}] < \infty, \tag{5.2}$$

and the radial increment moment functions satisfy

$$\mu_1(\boldsymbol{x}) := \mathbb{E}[R_{n+1} - R_n \mid X_n = \boldsymbol{x}] = \mathbb{E}_{\boldsymbol{x}}[\Delta_{\boldsymbol{x}}] + \frac{\mathbb{E}_{\boldsymbol{x}}[\|\Delta\|^2 - \Delta_{\boldsymbol{x}}^2]}{2\|\boldsymbol{x}\|} + \mathrm{O}(\|\boldsymbol{x}\|^{-1-\delta}), \tag{5.3}$$

$$\mu_2(\boldsymbol{x}) := \mathbb{E}[(R_{n+1} - R_n)^2 \mid X_n = \boldsymbol{x}] = \mathbb{E}_{\boldsymbol{x}}[\Delta_{\boldsymbol{x}}^2] + \mathrm{O}(\|\boldsymbol{x}\|^{-\delta}), \tag{5.4}$$

as $\|\boldsymbol{x}\| \to \infty$ for some $\delta = \delta(p) > 0$.

https://doi.org/10.1017/apr.2016.44 Published online by Cambridge University Press

Proof. By time homogeneity, it suffices to consider the $n = 0$ case. By the triangle inequality, $|R_1 - R_0| = |\|X_0 + \Delta\| - \|X_0\|| \le \|\Delta\|$, so that (5.2) follows from (A0).

We prove (5.3) and (5.4) by approximating

$$\|\boldsymbol{x} + \Delta\| - \|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x} + \Delta, \boldsymbol{x} + \Delta\rangle} - \|\boldsymbol{x}\| = \|\boldsymbol{x}\|\left[\left(1 + \frac{2\Delta_{\boldsymbol{x}}}{\|\boldsymbol{x}\|} + \frac{\|\Delta\|^2}{\|\boldsymbol{x}\|^2}\right)^{1/2} - 1\right] \tag{5.5}$$

for large $\boldsymbol{x}$. Let $A_{\boldsymbol{x}} = \{\|\Delta\| \le \|\boldsymbol{x}\|^{\beta}\}$ for some $\beta \in (0, 1)$ to be determined later. On the event $A_{\boldsymbol{x}}$, we approximate (5.5) using Taylor's formula for $(1 + y)^{1/2}$, and on the event $A_{\boldsymbol{x}}^{\mathrm{c}}$, we bound (5.5) using (A0).

Indeed, for all $y > -1$, Taylor's theorem with Lagrange remainder shows that

$$(1 + y)^{1/2} = 1 + \tfrac{1}{2}y - \tfrac{1}{8}y^2(1 + \gamma y)^{-3/2}$$

for some $\gamma = \gamma(y) \in [0, 1]$, so on the event $A_{\boldsymbol{x}}$,

$$\begin{aligned}
&\|\boldsymbol{x} + \Delta\| - \|\boldsymbol{x}\| \\
&\quad = \|\boldsymbol{x}\|\left(\frac{\Delta_{\boldsymbol{x}}}{\|\boldsymbol{x}\|} + \frac{\|\Delta\|^2}{2\|\boldsymbol{x}\|^2} - \frac{1}{8}\left(\frac{2\Delta_{\boldsymbol{x}}}{\|\boldsymbol{x}\|} + \frac{\|\Delta\|^2}{\|\boldsymbol{x}\|^2}\right)^2(1 + \mathrm{O}(\|\boldsymbol{x}\|^{\beta-1}))\right) \\
&\quad = \Delta_{\boldsymbol{x}} + \frac{\|\Delta\|^2}{2\|\boldsymbol{x}\|} - \frac{\|\boldsymbol{x}\|}{8}\left(\frac{4\Delta_{\boldsymbol{x}}^2}{\|\boldsymbol{x}\|^2} + \frac{\|\Delta\|^2}{\|\boldsymbol{x}\|^2}\left(\frac{4\Delta_{\boldsymbol{x}}}{\|\boldsymbol{x}\|} + \frac{\|\Delta\|^2}{\|\boldsymbol{x}\|^2}\right)\right)(1 + \mathrm{O}(\|\boldsymbol{x}\|^{\beta-1})) \\
&\quad = \Delta_{\boldsymbol{x}} + \frac{\|\Delta\|^2}{2\|\boldsymbol{x}\|} - \frac{\Delta_{\boldsymbol{x}}^2}{2\|\boldsymbol{x}\|}(1 + \mathrm{O}(\|\boldsymbol{x}\|^{\beta-1})) + \frac{\|\Delta\|^2}{\|\boldsymbol{x}\|}\mathrm{O}(\|\boldsymbol{x}\|^{\beta-1}) \\
&\quad = \Delta_{\boldsymbol{x}} + \frac{\|\Delta\|^2 - \Delta_{\boldsymbol{x}}^2}{2\|\boldsymbol{x}\|}(1 + \mathrm{O}(\|\boldsymbol{x}\|^{\beta-1})),
\end{aligned} \tag{5.6}$$

where the error terms follow from the fact that $|\Delta_{\boldsymbol{x}}| \le \|\Delta\| \le \|\boldsymbol{x}\|^{\beta}$ for $\beta < 1$.

On the other hand,

$$|\|\boldsymbol{x} + \Delta\| - \|\boldsymbol{x}\||\mathbf{1}(A_{\boldsymbol{x}}^{\mathrm{c}}) \le \|\Delta\|\mathbf{1}(A_{\boldsymbol{x}}^{\mathrm{c}}) = \|\Delta\|^p\|\Delta\|^{1-p}\mathbf{1}(A_{\boldsymbol{x}}^{\mathrm{c}}) \le \|\Delta\|^p\|\boldsymbol{x}\|^{\beta(1-p)}, \tag{5.7}$$

by the triangle inequality and the fact that $\|\Delta\| > \|\boldsymbol{x}\|^{\beta}$ on $A_{\boldsymbol{x}}^{\mathrm{c}}$. Since

$$\|\boldsymbol{x} + \Delta\| - \|\boldsymbol{x}\| = (\|\boldsymbol{x} + \Delta\| - \|\boldsymbol{x}\|)\mathbf{1}(A_{\boldsymbol{x}}) + (\|\boldsymbol{x} + \Delta\| - \|\boldsymbol{x}\|)\mathbf{1}(A_{\boldsymbol{x}}^{\mathrm{c}}),$$

we can combine (5.6) and (5.7) to give

$$\begin{aligned}
&\left|\|\boldsymbol{x} + \Delta\| - \|\boldsymbol{x}\| - \left[\Delta_{\boldsymbol{x}} + \frac{\|\Delta\|^2 - \Delta_{\boldsymbol{x}}^2}{2\|\boldsymbol{x}\|}(1 + \mathrm{O}(\|\boldsymbol{x}\|^{\beta-1}))\right]\right| \\
&\quad = \left|\|\boldsymbol{x} + \Delta\| - \|\boldsymbol{x}\| - \left[\Delta_{\boldsymbol{x}} + \frac{\|\Delta\|^2 - \Delta_{\boldsymbol{x}}^2}{2\|\boldsymbol{x}\|}(1 + \mathrm{O}(\|\boldsymbol{x}\|^{\beta-1}))\right]\right|\mathbf{1}(A_{\boldsymbol{x}}^{\mathrm{c}}) \\
&\quad \le \|\Delta\|^p\|\boldsymbol{x}\|^{\beta(1-p)} + \left|\Delta_{\boldsymbol{x}} + 2\left(\frac{\|\Delta\|^2 - \Delta_{\boldsymbol{x}}^2}{2\|\boldsymbol{x}\|}\right)\right|\mathbf{1}(A_{\boldsymbol{x}}^{\mathrm{c}}) \\
&\quad \le 2\|\Delta\|^p\|\boldsymbol{x}\|^{\beta(1-p)} + \|\Delta\|^p\|\boldsymbol{x}\|^{\beta(2-p)-1}.
\end{aligned}$$

Therefore, taking expectations and using (A0), we obtain

$$\mu_1(\boldsymbol{x}) = \mathbb{E}_{\boldsymbol{x}}[\Delta_{\boldsymbol{x}}] + \frac{\mathbb{E}_{\boldsymbol{x}}[\|\Delta\|^2 - \Delta_{\boldsymbol{x}}^2]}{2\|\boldsymbol{x}\|} + \mathrm{O}(\|\boldsymbol{x}\|^{\beta-2}) + \mathrm{O}(\|\boldsymbol{x}\|^{\beta(1-p)}) + \mathrm{O}(\|\boldsymbol{x}\|^{\beta(2-p)-1}).$$

https://doi.org/10.1017/apr.2016.44 Published online by Cambridge University Press

Taking $\beta = 2/p$ makes all the error terms of size $\mathrm{O}(\|\boldsymbol{x}\|^{-1-\delta})$ for some $\delta = \delta(p) > 0$, namely for $\delta = (p-2)/p$.

For the second moment, we use the identity

$$\begin{aligned}(\|\boldsymbol{x}+\Delta\| - \|\boldsymbol{x}\|)^2 &= \|\boldsymbol{x}+\Delta\|^2 - \|\boldsymbol{x}\|^2 - 2\|\boldsymbol{x}\|(\|\boldsymbol{x}+\Delta\| - \|\boldsymbol{x}\|)\\ &= 2\|\boldsymbol{x}\|\Delta_{\boldsymbol{x}} + \|\Delta\|^2 - 2\|\boldsymbol{x}\|(\|\boldsymbol{x}+\Delta\| - \|\boldsymbol{x}\|),\end{aligned}$$

so that

$$\mu_2(\boldsymbol{x}) = 2\|\boldsymbol{x}\|\,\mathbb{E}_{\boldsymbol{x}}[\Delta_{\boldsymbol{x}}] + \mathbb{E}_{\boldsymbol{x}}[\|\Delta\|^2] - 2\|\boldsymbol{x}\|\mu_1(\boldsymbol{x}) = \mathbb{E}_{\boldsymbol{x}}[\Delta_{\boldsymbol{x}}^2] + \mathrm{O}(\|\boldsymbol{x}\|^{-\delta}),$$

as required. □

With the additional assumptions (A1), (A3), and (A4), we can use Lemma 5.1 to prove the following result.

Lemma 5.2. *Suppose that X is a discrete-time, time-homogeneous Markov process on $\mathbb{X} \subseteq \mathbb{R}^d$ satisfying (A0), (A1), (A3), and (A4). Then, with μ_1 and μ_2 defined at (5.3) and (5.4), and $\varepsilon(r)$ defined at (A3), there exists $\delta > 0$ such that, as $\|\boldsymbol{x}\| \to \infty$,*

$$\begin{aligned}2\|\boldsymbol{x}\|\mu_1(\boldsymbol{x}) &= V - U + \mathrm{O}(\varepsilon(\|\boldsymbol{x}\|)) + \mathrm{O}(\|\boldsymbol{x}\|^{-\delta}),\\ \mu_2(\boldsymbol{x}) &= U + \mathrm{O}(\varepsilon(\|\boldsymbol{x}\|)) + \mathrm{O}(\|\boldsymbol{x}\|^{-\delta}).\end{aligned} \tag{5.8}$$

Proof. By the definition of $\varepsilon(r)$ at (A3) we have $\|M(\boldsymbol{x}) - \sigma^2(\hat{\boldsymbol{x}})\|_{\mathrm{op}} = \mathrm{O}(\varepsilon(\|\boldsymbol{x}\|))$ as $\|\boldsymbol{x}\| \to \infty$. Then (A4) implies that

$$\mathbb{E}_{\boldsymbol{x}}[\|\Delta\|^2] = \mathrm{tr}\,(M(\boldsymbol{x})) = \mathrm{tr}\,(\sigma^2(\hat{\boldsymbol{x}})) + \mathrm{O}(\varepsilon(\|\boldsymbol{x}\|)) = V + \mathrm{O}(\varepsilon(\|\boldsymbol{x}\|))$$

and

$$\mathbb{E}_{\boldsymbol{x}}[\Delta_{\boldsymbol{x}}^2] = \langle \hat{\boldsymbol{x}}, M(\boldsymbol{x})\cdot\hat{\boldsymbol{x}}\rangle = \langle \hat{\boldsymbol{x}}, \sigma^2(\hat{\boldsymbol{x}})\cdot\hat{\boldsymbol{x}}\rangle + \mathrm{O}(\varepsilon(\|\boldsymbol{x}\|)) = U + \mathrm{O}(\varepsilon(\|\boldsymbol{x}\|)),$$

and (A1) implies that $\mathbb{E}_{\boldsymbol{x}}[\Delta_{\boldsymbol{x}}] = \mathbb{E}_{\boldsymbol{x}}[\langle \Delta, \hat{\boldsymbol{x}}\rangle] = \langle \mu(\boldsymbol{x}), \hat{\boldsymbol{x}}\rangle = 0$. Using these expressions in Lemma 5.1 yields (5.8). □

Proof of Theorem 2.2. We apply Lamperti's [12] recurrence classification to $R_n = \|X_n\|$, the radial process. Proposition 2.1 shows that $\limsup_{n\to\infty} R_n = +\infty$, and Lemma 5.1 tells us that (5.2) is satisfied.

Because the error terms in (5.8) are uniform in $\boldsymbol{x}$, Lemma 5.2 shows that, for all $\eta > 0$, there exists $C < \infty$ such that $2\|\boldsymbol{x}\|\mu_1(\boldsymbol{x}) - \mu_2(\boldsymbol{x}) \in [V - 2U - \eta, V - 2U + \eta]$ for all $\boldsymbol{x} \in \mathbb{X}$ with $\|\boldsymbol{x}\| \geq C$. Therefore, it follows from Theorem 3.2 of [12] that X is transient if $V - 2U > 0$ and recurrent if $V - 2U < 0$. For the boundary case, when $V - 2U = 0$, if $\varepsilon(r) = \mathrm{O}(r^{-\delta_0})$ then $2\|\boldsymbol{x}\|\mu_1(\boldsymbol{x}) - \mu_2(\boldsymbol{x}) = \mathrm{O}(\|\boldsymbol{x}\|^{-\delta_1})$ for $\delta_1 = \min\{\delta, \delta_0\}$, which implies that X is recurrent, again by Theorem 3.2 of [12]. □

6. Nullity

In this section we give the proof of Theorem 2.3. In the transient case, this is straightforward.

Lemma 6.1. *In case (i) of Theorem 2.2, for any bounded $A \subset \mathbb{R}^d$, as $n \to \infty$, the null property (2.2) holds.*

Proof. It is sufficient to prove (2.2) for the case in which $A = B_r := \{\boldsymbol{x} \in \mathbb{X}\colon \|\boldsymbol{x}\| \leq r\}$. In case (i), X is transient, meaning that $\|X_n\| \to \infty$ a.s., so that $\mathbf{1}\{X_n \in B_r\} \to 0$ a.s. for any $r \in \mathbb{R}_+$. Hence, the Cesàro limit in (2.2) is also 0 a.s., and the L^q convergence follows from the bounded convergence theorem. □

https://doi.org/10.1017/apr.2016.44 Published online by Cambridge University Press

It remains to consider cases (ii) and (iii) of Theorem 2.2, when X is recurrent. Thus, there exists $r_0 \in \mathbb{R}_+$ such that $\liminf_{n\to\infty}\|X_n\| \le r_0$ a.s. Let $\tau_r := \min\{n \in \mathbb{Z}_+ \colon X_n \in B_r\}$. It suffices to take $A = B_r$, where $r > r_0$, so $X_n \in B_r$ infinitely often. We make the following claim, whose proof is deferred until the end of this section, which states that if the walk has not yet entered a ball of radius R (for any large enough $R > r$), the time until it reaches the ball of radius r has tail bounded below as displayed.

Lemma 6.2. *In cases (ii) and (iii) of Theorem 2.2, there exists a finite $r_1 \ge r_0$ such that, for any $r > r_1$ and $R > r$, there exists a finite positive c such that*

$$\mathbb{P}[\tau_r \ge n + m \mid X_0, \dots, X_n] \ge cm^{-1/2} \quad \text{on } \{n < \tau_R\} \tag{6.1}$$

for all sufficiently large m.

Proof of Theorem 2.3. In case (i), the result is contained in Lemma 6.1. So consider cases (ii) and (iii). Fix r and R with $R > r > r_1$, where r_1 is as in Lemma 6.2. Note that $\liminf_{n\to\infty}\|X_n\| \le r_0 \le r_1$ a.s.

Set $\gamma_1 := 0$, and then define recursively, for $\ell \in \mathbb{N}$, the stopping times

$$\eta_\ell := \min\{n \ge \gamma_\ell \colon X_n \notin B_R\}, \qquad \gamma_{\ell+1} := \min\{n \ge \eta_\ell \colon X_n \in B_r\},$$

with the convention that $\min \varnothing := \infty$. Since $r > r_0$ and $\limsup_{n\to\infty}\|X_n\| = \infty$ (by Proposition 2.1), for all $\ell \in \mathbb{N}$, we have $\eta_\ell < \infty$ and $\gamma_\ell < \infty$ a.s., and

$$0 = \gamma_1 < \eta_1 < \gamma_2 < \eta_2 < \cdots.$$

In particular, $\lim_{\ell\to\infty} \gamma_\ell = \lim_{\ell\to\infty} \eta_\ell = \infty$ a.s.

We now write $\mathcal{F}_n := \sigma(X_0, \dots, X_n)$. We use Lemma 4.1 to show that the process must exit from B_R rapidly enough. Indeed, if κ is any finite stopping time, set $Y_n = X_{\kappa+n} - X_\kappa$ and $\mathcal{G}_n = \mathcal{F}_{\kappa+n}$. Then assumptions (A0), (A1), and (A2) show that the hypotheses of Lemma 4.1 are satisfied, since, for example,

$$\mathbb{E}[\|Y_{n+1} - Y_n\|^p \mid \mathcal{G}_n] = \mathbb{E}[\|X_{\kappa+n+1} - X_{\kappa+n}\|^p \mid \mathcal{F}_{\kappa+n}] = \mathbb{E}_{X_{\kappa+n}}[\|\Delta\|^p],$$

by the strong Markov property for X at the finite stopping time $\kappa + n$. In particular, another application of Lemma 4.1, similarly to (4.6), shows that we may choose $n = n(R) \in \mathbb{N}$ sufficiently large so that

$$\mathbb{P}\Big[\max_{0\le\ell\le n(R)} \|X_{\kappa+\ell} - X_\kappa\| \ge 2R \;\Big|\; \mathcal{F}_\kappa\Big] \ge \frac{1}{2} \quad \text{a.s.}, \tag{6.2}$$

an event whose occurrence ensures that if $X_\kappa \in B_R$, then X exits B_R before time $\kappa + n(R)$. Fix $k \in \mathbb{N}$. Then, an application of (6.2) at stopping time $\kappa = \gamma_k$ shows that

$$\mathbb{P}[\eta_k - \gamma_k > n(R) \mid \mathcal{F}_{\gamma_k}] \le \mathbb{P}\Big[\max_{0\le\ell\le n(R)} \|X_{\gamma_k+\ell} - X_{\gamma_k}\| < 2R \;\Big|\; \mathcal{F}_{\gamma_k}\Big] \le \frac{1}{2} \quad \text{a.s.}$$

Similarly,

$$\begin{aligned}
&\mathbb{P}[\eta_k - \gamma_k > 2n(R) \mid \mathcal{F}_{\gamma_k}] \\
&\quad = \mathbb{E}[\mathbf{1}\{\eta_k - \gamma_k > n(R)\}\,\mathbb{E}[\mathbf{1}\{\eta_k - \gamma_k > 2n(R)\} \mid \mathcal{F}_{\gamma_k+n(R)}] \mid \mathcal{F}_{\gamma_k}] \\
&\quad \le \tfrac{1}{2}\mathbb{P}[\eta_k - \gamma_k > n(R) \mid \mathcal{F}_{\gamma_k}] \\
&\quad \le \tfrac{1}{4},
\end{aligned}$$

https://doi.org/10.1017/apr.2016.44 Published online by Cambridge University Press

this time applying (6.2) at stopping time $\kappa = \gamma_k + n(R)$ as well. Iterating this argument, it follows that $\mathbb{P}[\eta_k - \gamma_k > m \cdot n(R) \mid \mathcal{F}_{\gamma_k}] \le 2^{-m}$ a.s. for all $m \in \mathbb{N}$. From here, it is straightforward to deduce that, for some constant $C < \infty$ and any $k \in \mathbb{N}$,

$$\mathbb{E}[\eta_k - \gamma_k \mid \mathcal{F}_{\gamma_k}] \le C \quad \text{a.s.} \tag{6.3}$$

On the other hand, the tail estimate (6.1) implies that

$$\mathbb{P}[\gamma_{k+1} - \eta_k \ge m \mid \mathcal{F}_{\eta_k}] \ge cm^{-1/2} \quad \text{a.s.} \tag{6.4}$$

for $c > 0$ and all sufficiently large m.

For any $n \in \mathbb{N}$, set $k(n) := \min\{k \ge 2 \colon \gamma_k > n\}$, so that $\gamma_{k(n)-1} \le n < \gamma_{k(n)}$ for $k(n) \in \{2, 3, \ldots\}$. Note that $k(n) < \infty$ and $\lim_{n\to\infty} k(n) = \infty$ a.s. Then we claim that

$$\frac{1}{n}\sum_{k=0}^{n-1} \mathbf{1}\{X_k \in B_r\} \le \frac{\sum_{k=1}^{k(n)-1}(\eta_k - \gamma_k)}{\sum_{k=1}^{k(n)-2}(\gamma_{k+1} - \eta_k)}. \tag{6.5}$$

This is easiest to see by considering two separate cases. First, if $\eta_{k(n)-1} < n < \gamma_{k(n)}$,

$$\frac{1}{n}\sum_{k=0}^{n-1} \mathbf{1}\{X_k \in B_r\} \le \frac{1}{\eta_{k(n)-1}} \sum_{k=0}^{\eta_{k(n)-1}} \mathbf{1}\{X_k \in B_r\},$$

which implies (6.5), since the set of k less than n for which $X_k \in B_r$ is contained in the set $\bigcup_{k=1}^{k(n)-1}[\gamma_k, \eta_k)$. On the other hand, if $\gamma_{k(n)-1} \le n \le \eta_{k(n)-1}$, using the elementary inequality

$$\frac{a}{b} \le \frac{a+c}{b+c}$$

for nonnegative a, b, and c with $a/b \le 1$, we have

$$\frac{1}{n}\sum_{k=0}^{n-1} \mathbf{1}\{X_k \in B_r\} \le \frac{1}{\eta_{k(n)-1}}\left(\sum_{k=0}^{n-1} \mathbf{1}\{X_k \in B_r\} + \eta_{k(n)-1} - n\right),$$

which again gives (6.5).

To estimate the growth rates of the numerator and denominator on the right-hand side of (6.5), we apply some results from [7]. First, writing $Z_m = \sum_{k=1}^{m-1}(\eta_k - \gamma_k)$ and $\mathcal{G}_m = \mathcal{F}_{\gamma_m}$, by (6.3) we can apply Theorem 2.4 of [7] to the $\mathcal{G}_m$-adapted process Z_m to show that, for any $\varepsilon > 0$, a.s., for all but finitely many m, $\sum_{k=1}^{m-1}(\eta_k - \gamma_k) \le m^{1+\varepsilon}$. On the other hand, writing $Z_m = \sum_{k=1}^{m-1}(\gamma_{k+1} - \eta_k)$ and $\mathcal{G}_m = \mathcal{F}_{\eta_m}$, by (6.4) we can apply Theorem 2.6 of [7] to the $\mathcal{G}_m$-adapted process Z_m to show that, for any $\varepsilon > 0$ and all sufficiently large m, $\sum_{k=1}^{m-1}(\gamma_{k+1} - \eta_k) \ge m^{2-\varepsilon}$. Now (6.5) gives the almost-sure version of result (2.2). The L^q version follows from the bounded convergence theorem. □

It remains to complete the proof of Lemma 6.2. A more general, two-sided version of the inequality in Lemma 6.2 is proved in [6, Theorem 2.4], but under slightly different assumptions. Because of this, we cannot apply that result directly; nevertheless, the proof techniques naturally transfer to our setting. In doing so, the arguments become simpler to apply, so we reproduce them here.

https://doi.org/10.1017/apr.2016.44 Published online by Cambridge University Press

Proof of Lemma 6.2. By the Markov property for X, it is enough to prove the statement for $n = 0$, namely that there exists finite $r_1 \geq r_0$ such that, for any $r > r_1$ and $R > r$, there exists a finite positive constant c such that, if $X_0 \notin B_R$ then

$$\mathbb{P}[\tau_r > m \mid X_0] \geq cm^{-1/2} \quad \text{for sufficiently large } m.$$

We outline the two intuitive steps in the proof. First we show that the probability that $\max_{0\leq k\leq \tau_r} \|X_k\|$ exceeds some large x is bounded below by a constant times $1/x$. Second, we show that if the latter event does occur, with probability at least $\frac{1}{2}$, it takes a process time of at least a constant times x^2 to reach B_r. Combining these two estimates will show that, with probability of order $1/x$, the walk takes a time of order x^2 to reach B_r, which gives the desired tail bound. Roughly speaking, the first estimate (reaching distance x) is provided by the optional stopping theorem and the fact that $\|X_k\|$ is a submartingale (cf. [6, Theorem 2.3]), and the second (taking quadratic time to return) is provided by a maximal inequality applied to an appropriate quadratic displacement functional (cf. [6, Lemma 4.11]). A technicality required for the first estimate is that to apply optional stopping, we need uniform integrability; so we actually work with a truncated version of $\|X_k\|$.

We now give the details. Recall that $R_k = \|X_k\|$, and let $\mathcal{F}_k = \sigma(X_0, \ldots, X_k)$. Lemmas 5.1 and 5.2, with the fact that $V > U$ by (A4), imply that

$$\mathbb{E}[R_{k+1} - R_k \mid \mathcal{F}_k] \geq \frac{2\varepsilon}{R_k} + \mathrm{o}(R_k^{-1}) \geq \frac{\varepsilon}{R_k} \tag{6.6}$$

for all $R_k > r_1$, for sufficiently large $r_1 \geq r_0$ and some positive constant ε. Now, suppose that r and R satisfy $R > r > r_1$, and fix x with $x \gg R$. Set $R_k^x := \min\{2x, R_k\}$ and $\sigma_x := \min\{k \geq 0 \colon R_k > x\}$. Since X_k is a martingale, R_k is a submartingale, as is the stopped process $Y_k := R_{k\wedge\tau_r\wedge\sigma_x}$. In order to achieve uniform integrability, we consider the truncated process $Y_k^x := R_{k\wedge\tau_r\wedge\sigma_x}^x$ and show that this is a submartingale.

For $k \geq \tau_r \wedge \sigma_x$, we have $Y_{k+1}^x - Y_k^x = 0$, so $\mathbb{E}[Y_{k+1}^x - Y_k^x \mid \mathcal{F}_k] = 0$. For $k < \tau_r \wedge \sigma_x$,

$$Y_{k+1}^x - Y_k^x = R_{k+1} - R_k + (2x - R_{k+1})\mathbf{1}\{R_{k+1} > 2x\},$$

and the last term can be bounded in absolute value, i.e.

$$\begin{aligned} |(2x - R_{k+1})\mathbf{1}\{R_{k+1} > 2x\}| &\leq |R_{k+1} - R_k|\mathbf{1}\{R_{k+1} > 2x\} \\ &\leq |R_{k+1} - R_k|\mathbf{1}\{|R_{k+1} - R_k| > x\} \\ &\leq |R_{k+1} - R_k|^p x^{1-p} \end{aligned}$$

for $p > 2$ as appearing in (A0), since on $\{k < \sigma_x\}$ we have $R_k < x$ and therefore $R_{k+1} > 2x$ implies that $|R_{k+1} - R_k| > x$. Applying (5.2) from Lemma 5.1 we obtain

$$\mathbb{E}[|(Y_{k+1}^x - Y_k^x) - (R_{k+1} - R_k)| \mid \mathcal{F}_k] \leq Bx^{1-p}$$

for some $B < \infty$ not depending on x. Combining this with (6.6) and again the fact that $R_k < x$ on $\{k < \sigma_x\}$, we have

$$\mathbb{E}[Y_{k+1}^x - Y_k^x \mid \mathcal{F}_k] \geq \frac{\varepsilon}{R_k} - Bx^{1-p} \geq \frac{\varepsilon}{x} - Bx^{1-p} \geq 0 \quad \text{for sufficiently large } x.$$

https://doi.org/10.1017/apr.2016.44 Published online by Cambridge University Press

Hence, for sufficiently large x, Y_k^x is a uniformly integrable submartingale and, therefore, given $X_0 \notin B_R$, by optional stopping,

$$\begin{aligned} R < R_0 = Y_0^x &\le \mathbb{E}[Y^x_{\sigma_x \wedge \tau_r} \mid X_0] \\ &= \mathbb{E}[Y^x_{\sigma_x} \mathbf{1}\{\sigma_x < \tau_r\} \mid X_0] + \mathbb{E}[Y^x_{\tau_r} \mathbf{1}\{\tau_r < \sigma_x\} \mid X_0] \\ &\le 2x\mathbb{P}[\sigma_x < \tau_r \mid X_0] + r. \end{aligned}$$

In other words, given $X_0 \notin B_R$,

$$\mathbb{P}\Big[\max_{0\le k\le \tau_r} R_k > x \;\Big|\; X_0\Big] \ge \frac{R-r}{2x} \quad \text{for all sufficiently large } x. \tag{6.7}$$

Now, consider $W_k := R_{\sigma_x+k} - R_{\sigma_x}$, adapted to $\mathcal{G}_k := \mathcal{F}_{\sigma_x+k}$. We have

$$W_{k+1}^2 - W_k^2 = R^2_{\sigma_x+k+1} - R^2_{\sigma_x+k} - 2R_{\sigma_x}(R_{\sigma_x+k+1} - R_{\sigma_x+k}).$$

Using the fact that R_k is a submartingale together with the strong Markov property for X at the stopping time $\sigma_x + k$ yields $\mathbb{E}[R_{\sigma_x+k+1} - R_{\sigma_x+k} \mid \mathcal{F}_{\sigma_x+k}] \ge 0$ a.s., and Lemmas 5.1 and 5.2 again with the strong Markov property imply that $\mathbb{E}[R^2_{\sigma_x+k+1} - R^2_{\sigma_x+k} \mid \mathcal{F}_{\sigma_x+k}] \le C$ a.s. for some constant $C < \infty$; hence, $\mathbb{E}[W_{k+1}^2 - W_k^2 \mid \mathcal{G}_k] \le C$ a.s. for some constant $C < \infty$. Then a maximal inequality [13, Lemma 3.1] similar to Doob's submartingale inequality implies that, on $\{\sigma_x < \infty\}$,

$$\mathbb{P}\Big[\max_{0\le k\le n} W_k^2 \ge y \;\Big|\; \mathcal{G}_0\Big] \le \frac{Cn}{y} \quad \text{for any } y > 0.$$

In particular, we may choose $\varepsilon > 0$ small enough so that

$$\mathbb{P}\Big[\max_{0\le k\le \varepsilon x^2} |R_{\sigma_x+k} - R_{\sigma_x}| \ge \frac{x}{2} \;\Big|\; \mathcal{F}_{\sigma_x}\Big] \le \frac{1}{2} \quad \text{on } \{\sigma_x < \infty\}. \tag{6.8}$$

Combining inequalities (6.7) and (6.8), we find that, given $X_0 \notin B_R$,

$$\begin{aligned} &\mathbb{P}\Big[\Big\{\max_{0\le k\le \tau_r} R_k > x\Big\} \cap \Big\{\max_{0\le k\le \varepsilon x^2} |R_{\sigma_x+k} - R_{\sigma_x}| < \frac{x}{2}\Big\} \;\Big|\; X_0\Big] \\ &\quad = \mathbb{E}\Big[\mathbf{1}\{\sigma_x < \tau_r\}\mathbb{P}\Big[\max_{0\le k\le \varepsilon x^2} |R_{\sigma_x+k} - R_{\sigma_x}| < \frac{x}{2} \;\Big|\; \mathcal{F}_{\sigma_x}\Big] \;\Big|\; X_0\Big] \\ &\quad \ge \frac{1}{2}\mathbb{P}\Big[\max_{0\le k\le \tau_r} R_k > x \;\Big|\; X_0\Big] \\ &\quad \ge \frac{R-r}{4x} \end{aligned}$$

for sufficiently large x, where the equality here uses the fact that $\{\sigma_x < \tau_r\} \in \mathcal{F}_{\sigma_x}$.

If both of the events $\{\max_{0\le k\le \tau_r} R_k > x\}$ and $\{\max_{0\le k\le \varepsilon x^2} |R_{\sigma_x+k} - R_{\sigma_x}| < x/2\}$ occur, then the process X_k leaves the ball B_x before time τ_r and takes more than εx^2 steps to return to the ball $B_{x/2} \subset B_r$, and, therefore, $\tau_r > \varepsilon x^2$. Setting $m = \varepsilon x^2$ and $c = (R-r)\sqrt{\varepsilon}/4$ yields the claimed inequality. □

Remark 6.1. It is only in the proof of Lemma 6.2 that we use the condition $U < V$ from (A4). In the $U = V$ case, inequality (6.6) holds only for (any) $\varepsilon < 0$, and not $\varepsilon > 0$; thus, to obtain a submartingale, one should look at $(Y_k^x)^\gamma$ for $\gamma > 1$. The modified argument yields a weaker version of (6.1), with $m^{-1/2}$ replaced by $m^{-(1/2)-\delta}$ for any $\delta > 0$, but, as stated in Remark 2.1, this is still comfortable enough to give Theorem 2.3 (any exponent greater than -1 in the tail bound will do). We omit these additional technical details, as the $U = V$ case is outside our main interest.

https://doi.org/10.1017/apr.2016.44 Published online by Cambridge University Press

Appendix A. Recurrence in one dimension

We use a Lyapunov-function method with function $f(x) = \log(1 + |x|)$.

Lemma A.1. *Suppose that X is a discrete-time, time-homogeneous Markov process on $\mathbb{X} \subseteq \mathbb{R}$. Suppose that, for some $p > 2$ and $v > 0$,*

$$\sup_{x\in\mathbb{X}} \mathbb{E}[(X_{n+1} - X_n)^p \mid X_n = x] < \infty, \qquad \inf_{x\in\mathbb{X}} \mathbb{E}[(X_{n+1} - X_n)^2 \mid X_n = x] \geq v.$$

Suppose also that, for some bounded set $A \subset \mathbb{R}$, $\mathbb{E}[X_{n+1} - X_n \mid X_n = x] = 0$ for all $x \in \mathbb{X}\setminus A$. Then there exists a bounded set $A' \subset \mathbb{R}$ for which $\mathbb{E}[f(X_{n+1}) - f(X_n) \mid X_n = x] \leq 0$ for all $x \in \mathbb{X} \setminus A'$.

Proof. Write $\Delta = X_1 - X_0$ and $E_x = \{|\Delta| < |x|\}$. We compute

$$\mathbb{E}[f(X_{n+1}) - f(X_n) \mid X_n = x] = \mathbb{E}_x[(f(x+\Delta) - f(x))\mathbf{1}(E_x)] + \mathbb{E}_x[(f(x+\Delta) - f(x))\mathbf{1}(E_x^{\mathrm{c}})].$$

On $\{|\Delta| < |x|\}$, x and $x + \Delta$ have the same sign, so

$$\begin{aligned}\mathbb{E}_x[(f(x + \Delta) - f(x))\mathbf{1}(E_x)] &= \mathbb{E}_x\left[\log\left(\frac{1 + |x + \Delta|}{1 + |x|}\right)\mathbf{1}(E_x)\right] \\ &= \mathbb{E}_x\left[\log\left(1 + \frac{\Delta\,\mathrm{sgn}(x)}{1 + |x|}\right)\mathbf{1}(E_x)\right] \\ &\leq \frac{\mathrm{sgn}(x)}{1 + |x|}\,\mathbb{E}_x[\Delta\mathbf{1}(E_x)] - \frac{1}{6}\frac{\mathbb{E}_x[\Delta^2\mathbf{1}(E_x)]}{(1 + |x|)^2},\end{aligned}$$

using the inequality $\log(1 + y) \leq y - \frac{1}{6}y^2$ for all $-1 < y \leq 1$. Here, since $\mathbb{E}_x[\Delta] = 0$ for $x \notin A$,

$$|\mathbb{E}_x[\Delta\mathbf{1}(E_x)]| \leq \mathbb{E}_x[|\Delta|\mathbf{1}(E_x^{\mathrm{c}})] \leq \mathbb{E}_x[|\Delta|^p|x|^{1-p}] = \mathrm{o}(|x|^{-1}).$$

Similarly,

$$\mathbb{E}_x[\Delta^2\mathbf{1}(E_x)] \geq v - \mathbb{E}_x[\Delta^2\mathbf{1}(E_x^{\mathrm{c}})] \geq v - \mathrm{o}(1).$$

(Note that here, and in what follows, our notation follows the convention as described by (5.1); consequently, in one dimension the error terms are understood to be uniform as either $x \to +\infty$ or $x \to -\infty$.) Finally, we estimate the term

$$|\mathbb{E}_x[(f(x + \Delta) - f(x))\mathbf{1}(E_x^{\mathrm{c}})]| \leq \mathbb{E}_x[(\log(1 + |\Delta|) + \log(1 + 2|\Delta|))\mathbf{1}(E_x^{\mathrm{c}})].$$

Here,

$$\log(1 + 2|\Delta|)\mathbf{1}(E_x^{\mathrm{c}}) = \log(1 + 2|\Delta|)|\Delta|^p|\Delta|^{-p}\mathbf{1}(E_x^{\mathrm{c}}) \leq |x|^{-p}\log(1 + 2|x|)|\Delta|^p$$

for all x with $|x|$ greater than some sufficiently large x_0, using the fact that $y \mapsto y^{-p}\log(1+2y)$ is eventually decreasing. It follows that

$$|\mathbb{E}_x[(f(x + \Delta) - f(x))\mathbf{1}(E_x^{\mathrm{c}})]| \leq 2|x|^{-p}\log(1 + 2|x|)\,\mathbb{E}_x[|\Delta|^p] = \mathrm{o}(|x|^{-2}).$$

Combining these calculations we obtain

$$\begin{aligned}\mathbb{E}[f(X_{n+1}) - f(X_n) \mid X_n = x] &\leq \frac{\mathrm{sgn}(x)}{1 + |x|}\mathrm{o}(|x|^{-1}) - \frac{1}{6}\frac{v - \mathrm{o}(1)}{(1 + |x|)^2} + \mathrm{o}(|x|^{-2}) \\ &\leq -\frac{v}{6(1 + |x|)^2} + \mathrm{o}(|x|^{-2}),\end{aligned}$$

which is negative for all x with sufficiently large $|x|$. □

https://doi.org/10.1017/apr.2016.44 Published online by Cambridge University Press

Proof of Theorem 2.1. Under assumptions (A0), (A1), and (A2), the hypotheses of Lemma A.1 are satisfied, so that, for some $x_0 \in \mathbb{R}_+$, $\mathbb{E}[f(X_{n+1}) - f(X_n) \mid X_n = x] \le 0$ for all $x \in \mathbb{X}$ with $|x| \ge x_0$.

We note that assumption (A0) implies that $\mathbb{E}[|X_n|] < \infty$ for all n, and, therefore, $\mathbb{E}[f(X_n)] < \infty$ for all n. Let $n_0 \in \mathbb{N}$, and set $\tau = \min\{n \ge n_0 \colon |X_n| \le x_0\}$. Let $Y_n = f(X_{n\wedge\tau})$. Then $(Y_n,\ n \ge n_0)$ is a nonnegative supermartingale, and, hence, there exists a random variable $Y_\infty \in \mathbb{R}_+$ with $\lim_{n\to\infty} Y_n = Y_\infty$ a.s. In particular, this means that $\limsup_{n\to\infty} f(X_n) \le Y_\infty$ on $\{\tau = \infty\}$. Setting $\zeta = \sup\{|x| \colon x \in \mathbb{X},\ f(x) \le Y_\infty\}$, which satisfies $\zeta < \infty$ a.s., since $f(x) \to \infty$ as $|x| \to \infty$, it follows that $\limsup_{n\to\infty} |X_n| \le \zeta$ on $\{\tau = \infty\}$. However, under assumptions (A0), (A1), and (A2), Proposition 2.1 implies that $\limsup_{n\to\infty} |X_n| = +\infty$ a.s., so to avoid contradiction, we must have $\tau < \infty$ a.s. In other words, $\mathbb{P}[\inf_{n\ge n_0} |X_n| \le x_0] = 1$, and since n_0 was arbitrary, it follows that

$$\mathbb{P}\Bigg[\bigcap_{n_0\in\mathbb{N}} \Big\{\inf_{n\ge n_0} |X_n| \le x_0\Big\}\Bigg] = 1,$$

which gives the result. □

Acknowledgements

Part of this work was supported by the Engineering and Physical Sciences Research Council (grant number EP/J021784/1).

An antecedent of this work, concerning only the elliptic random walk in two dimensions, was written down in 2008–9 by MM and AW, who benefited from stimulating discussions with Iain MacPhee (7/11/1957–13/1/2012). The present authors also thank Stas Volkov for a comment that inspired Remark 3.2(d).

This work was carried out while AM was at the Department of Mathematics, Imperial College London.

References

[1] Berg, H. C. (1993). *Random Walks in Biology*, expanded edn. Princeton University Press.
[2] Bingham, N. H. (1972). Random walk on spheres. *Z. Wahrscheinlichkeitsth.* **22,** 169–192.
[3] Carazza, B. (1977). The history of the random-walk problem: considerations on the interdisciplinarity in modern physics. *Rivista del Nuovo Cimento Serie 2* **7,** 419–427.
[4] Chung, K. L. and Fuchs, W. H. J. (1951). On the distribution of values of sums of random variables. *Mem. Amer. Math. Soc.* **6,** 12pp.
[5] Gut, A. (2005). *Probability: A Graduate Course* (Springer Texts Statist. **75**). Springer, New York.
[6] Hryniv, O., Menshikov, M. V. and Wade, A. R. (2013). Excursions and path functionals for stochastic processes with asymptotically zero drifts. *Stoch. Process. Appl.* **123,** 1891–1921.
[7] Hryniv, O., MacPhee, I. M., Menshikov, M. V. and Wade, A. R. (2012). Non-homogeneous random walks with non-integrable increments and heavy-tailed random walks on strips. *Electron. J. Prob.* **17,** article 59, 28pp.
[8] Hughes, B. D. (1995). *Random Walks and Random Environments*, Vol. 1. Oxford University Press.
[9] Kallenberg, O. (2002). *Foundations of Modern Probability*, 2nd edn. Springer, New York.
[10] Kemperman, J. H. B. (1974). The oscillating random walk. *Stoch. Process. Appl.* **2,** 1–29.
[11] Kingman, J. F. C. (1963). Random walks with spherical symmetry. *Acta Math.* **109,** 11–53.
[12] Lamperti, J. (1960). Criteria for the recurrence or transience of stochastic process, I. *J. Math. Anal. Appl.* **1,** 314–330.
[13] Menshikov, M. V., Vachkovskaia, M. and Wade, A. R. (2008). Asymptotic behaviour of randomly reflecting billiards in unbounded tubular domains. *J. Statist. Phys.* **132,** 1097–1133.
[14] Nossal, R. (1983). Stochastic aspects of biological locomotion. *J. Statist. Phys.* **30,** 391–400.
[15] Nossal, R. J. and Weiss, G. H. (1974). A generalized Pearson random walk allowing for bias. *J. Statist. Phys.* **10,** 245–253.

https://doi.org/10.1017/apr.2016.44 Published online by Cambridge University Press

[16] Pearson, K. and Blakeman, J. (1906). *A Mathematical Theory of Random Migration* (Drapers' Company Research Memoirs, Biometric Series **3**). Dulau and Co., London.
[17] Pearson, K. and Lord Rayleigh (1905). The problem of the random walk. *Nature* **72,** pp. 294, 318, 342.
[18] Peres, Y., Popov, S. and Sousi, P. (2013). On recurrence and transience of self-interacting random walks. *Bull. Brazilian Math. Soc. (N.S.)* **44,** 841–867.
[19] Rogozin, B. A. and Foss, S. G. (1978). Recurrency of an oscillating random walk. *Theory Prob. Appl.* **23,** 155–162. Translated from *Teor. Veroyatn. Primen.* **23** (1978) 161–169.
[20] Zeitouni, O. (2006). *Lecture Notes on Random Walks in Random Environment*. 9th May 2006 version. Available at http://www.wisdom.weizmann.ac.il/∼zeitouni/.

NICHOLAS GEORGIOU, *Durham University and Heilbronn Institute for Mathematical Research*
Department of Mathematical Sciences, Durham University, South Road, Durham DH1 3LE, UK.
Email address: nicholas.georgiou@durham.ac.uk

MIKHAIL V. MENSHIKOV, *Durham University*
Department of Mathematical Sciences, Durham University, South Road, Durham DH1 3LE, UK.
Email address: mikhail.menshikov@durham.ac.uk

ALEKSANDAR MIJATOVIĆ, *King's College London*
Department of Mathematics, King's College London, Strand, London WC2R 2LS, UK.
Email address: aleksandar.mijatovic@kcl.ac.uk

ANDREW R. WADE, *Durham University*
Department of Mathematical Sciences, Durham University, South Road, Durham DH1 3LE, UK.
Email address: andrew.wade@durham.ac.uk

https://doi.org/10.1017/apr.2016.44 Published online by Cambridge University Press