

ANOMALOUS RECURRENCE OF MARKOV CHAINS ON NEGATIVELY CURVED MANIFOLDS

JOHN ARMSTRONG* ** AND
TIM KING , *** *King's College London*

Abstract

We present a recurrence–transience classification for discrete-time Markov chains on manifolds with negative curvature. Our classification depends only on geometric quantities associated to the increments of the chain, defined via the Riemannian exponential map. We deduce that a recurrent chain that has zero average drift at every point cannot be uniformly elliptic, unlike in the Euclidean case. We also give natural examples of zero-drift recurrent chains on negatively curved manifolds, including on a stochastically incomplete manifold.

Keywords: Non-homogeneous random walk; uniform ellipticity.

2020 Mathematics Subject Classification: Primary 60J05
Secondary 60D05

1. Introduction

It is a classical result [12] that Brownian motion in hyperbolic space is transient in dimensions two and higher, in contrast to the Euclidean case [15], where it is recurrent in dimension two (meaning that, almost surely, it visits any given open set at arbitrarily large times). In this paper we study more general random walks on negatively curved manifolds. We focus our attention on cases where the process respects the geometry of the manifold. Specifically, we consider discrete-time Markov processes that have martingale-like properties. To define a martingale on a manifold, one needs some geometric structure. We will be interested in processes where, in the chart induced by the Riemannian exponential map, each increment has zero mean. Such processes are called zero-drift processes.

Even in dimension three or more, a zero-drift Markov chain in Euclidean space need not be transient. Examples of recurrent zero-drift chains include what could be termed the ‘maximal symmetric random walk’ of [25, Theorem 1.5], and the ‘elliptic random walk’ of [8, Section 3]. These examples are light-tailed (the conditional increment of the chain has a finite covariance matrix at every point), and the latter is uniformly elliptic, meaning that there exists $\varepsilon > 0$ such that, for any fixed direction, there is a probability of at least ε that the chain will move a distance at least ε in that direction (see Section 2 for a precise definition).

Our main result (Theorem 1) is a recurrence–transience criterion for Markov chains on negatively curved manifolds. The criterion is phrased in terms of certain geometric quantities defined in the tangent bundle of the manifold. We deduce from our result that, unlike in the

Received 20 September 2021; revision received 15 March 2022.

* Postal address: Department of Mathematics, Strand Building, Strand, London, WC2R 2LS

** Email address: john.armstrong@kcl.ac.uk

*** Email address: tim.king@kcl.ac.uk

© The Author(s), 2022. Published by Cambridge University Press on behalf of Applied Probability Trust.

Euclidean case, zero-drift recurrent walks on negatively curved manifolds cannot be uniformly elliptic. More generally, we quantify the extent to which uniform ellipticity must fail, in terms of the asymptotic behaviour of the curvature of the manifold, if a zero-drift chain is to be recurrent. This allows us to write down recurrent chains on a large class of manifolds, including some that are *stochastically incomplete*, meaning that Brownian motion is not merely transient but explosive, in the sense that it may go to infinity within finite time. Another contrast we observe is that in Euclidean space it is possible to give a simple recurrence criterion using the growth of quantities calculated from the covariance matrices. We give an example (Proposition 6) to show that the corresponding results do not hold in hyperbolic space for any polynomial growth condition.

Our proof strategy is to combine the methods in [8], which uses methods of Lamperti type [19], with differential geometric comparison theorems. Whereas there is a well-established literature on the recurrence and ergodicity of random walks on manifolds with a Lie group or homogeneous space structure (see, for example, [26] and the references therein), comparison theorems allow us to study manifolds that do not have these structures, by reducing certain computations to the constant-curvature case. The use of comparison theorems is standard within the study of Brownian motion on manifolds [10, 11], but is less known within the Markov chains literature. The technical details are different from the Euclidean case, and the main technical novelty in this paper is Proposition 3, which gives an asymptotic approximation to the moments of the increments of the process measuring distance from an origin, in terms of geometrically meaningful quantities.

In the existing literature, much attention has been paid to the rate of escape of (continuous and discrete) Markov processes on manifolds, and the question of the ultimate fate of the angular process (defined using geodesic polar coordinates). The most basic example, Brownian motion in hyperbolic space (of arbitrary dimension ≥ 2), escapes to infinity at linear speed, and, unlike in the Euclidean case, its angular process almost surely converges to a limiting direction. These facts can be proved using a variety of techniques, including ergodic theory and group-theoretic methods [16], harmonic function theory [17, 29], or by applying integral tests for one-dimensional processes [27]. We show in Section 5 that (under certain assumptions) a uniformly elliptic Markov chain will always escape at linear speed. We also give an example where reducing the movements that a chain makes in the transverse (as opposed to radial) direction reduces the rate of escape. We return to some of these concepts at the end of the paper as avenues of exploration in future work.

2. Notation and main results

Throughout, we adopt the convention that $0 \in \mathbb{N}$, and denote by $X = (X_n)_{n \in \mathbb{N}}$ a discrete-time, time-homogeneous Markov chain whose state space is a Riemannian manifold M , with the Borel sigma algebra. We review the geometric concepts that we need; the reader unfamiliar with this material might consult, for example, [20]. Suppose that M has dimension d . If $x \in M$ is a point, then the *tangent space* at x , denoted $T_x M$, is a d -dimensional real vector space whose elements may be viewed as ‘vectors tangent to M at x ’. The *tangent bundle* TM is a manifold of dimension $2d$ whose points consist of pairs (x, v) , where $v \in T_x M$. The space $T_x M$ inherits an inner product space structure, denoted $\langle \cdot \rangle_x$, from the Riemannian metric on M . Let $\text{Planes}(M, x)$ denote the collection of two-dimensional subspaces of $T_x M$. Then the *sectional curvature* of M at x with respect to the plane $\pi \in \text{Planes}(M, x)$, which we denote by $\text{sec}(x, \pi)$, is a real number that may be calculated using the metric on M . If M is a sphere of radius r (in any dimension), then $\text{sec}(x, \pi) = r^{-2}$ for all choices of x and π . We are interested in *hyperbolic* manifolds, where the sectional curvature is everywhere negative. Finally, the

Riemannian exponential map $\exp_x : T_x M \rightarrow M$ (when it exists) sends the vector $v \in T_x M$ to the point in M given by starting at x and travelling along the geodesic determined by v a Riemannian distance of $\sqrt{\langle v, v \rangle_x}$. We denote by $\text{Dist}_M(x, y)$ the Riemannian distance between $x \in M$ and $y \in M$.

Assumption 1. *M is complete and simply connected. Also, there is a constant $\kappa > 0$ such that $\sec(x, \pi) \leq -\kappa^2$ for all $x \in M$ and all planes π .*

Under Assumption 1, the Riemannian exponential map $\exp_x : T_x M \rightarrow M$ exists and is a diffeomorphism for every $x \in M$ [14, Lemma 2.1.4]. For technical reasons, we make an arbitrary (but fixed) choice of origin $O \in M$, and work with local notions of recurrence and transience. See the discussion following Definition 1 for how these local notions relate to (global) recurrence and transience, which do not depend on a choice of origin. Having chosen O , the function

$$e_{\text{rad}}(x) = \frac{\exp_x^{-1}(O)}{\sqrt{\langle \exp_x^{-1}(O), \exp_x^{-1}(O) \rangle_x}} \quad (x \neq O)$$

is well defined. We also define some sequences of random variables, indexed by $n \in \mathbb{N}$:

$$\begin{aligned} V^{(n)} &= \exp_{X_n}^{-1}(X_{n+1}), \\ D_{\text{tot}}^{(n)} &= \sqrt{\langle V^{(n)}, V^{(n)} \rangle_{X_n}}, \\ D_{\text{rad}}^{(n)} &= \begin{cases} -\langle V^{(n)}, e_{\text{rad}}(X_n) \rangle_{X_n} & \text{if } X_n \neq O, \\ D_{\text{tot}}^{(n)} & \text{if } X_n = O, \end{cases} \\ \Phi^{(n)} &= \begin{cases} D_{\text{rad}}^{(n)} / D_{\text{tot}}^{(n)} & \text{if } D_{\text{tot}}^{(n)} > 0, \\ 0 & \text{if } D_{\text{tot}}^{(n)} = 0. \end{cases} \end{aligned}$$

The TM -valued process $(X_n, V^{(n)})$ is known as a *geodesic random walk* (GRW). The excursion theory of a particular GRW on hyperbolic space, and the large deviation theory of GRWs on general Riemannian manifolds, have been respectively studied in [3, 18]. Geometrically, $D_{\text{tot}}^{(n)}$ is the total Riemannian distance between X_n and X_{n+1} , $e_{\text{rad}}(x)$ is the unit vector that ‘points from x to the origin’, $D_{\text{rad}}^{(n)}$ is the length of the radial component of $V^{(n)}$, and $\Phi^{(n)}$ is the cosine of the angle between $V^{(n)}$ and $(-e_{\text{rad}}(X_n))$ in $T_{X_n} M$. We also define the *radial process* $(R_n)_{n \in \mathbb{N}}$ by $R_n = \text{Dist}_M(X_n, O)$.

We introduce another piece of notation. If H is a (suitable) function from a subset of \mathbb{R}^3 to \mathbb{R} , then we write $\mathbb{E}_x[H(D_{\text{tot}}, \Delta R, \Phi)]$ as shorthand for $\mathbb{E}[H(D_{\text{tot}}^{(n)}, \Delta_n R, \Phi^{(n)}) | X_n = x]$, where $\Delta_n R := R_{n+1} - R_n$. This notation is unambiguous because X is Markov, and so the expression does not depend on n . We make the following assumptions on the chain X .

Assumption 2. *There exist $p > 2$ and $B \in \mathbb{R}$ such that $\mathbb{E}_x[(D_{\text{tot}})^p] \leq B$ for all $x \in M$.*

Assumption 3. *It is almost surely the case that $\limsup_{n \rightarrow \infty} \text{Dist}_M(X_n, x) = \infty$ for some (equivalently for all) $x \in M$.*

Assumption 2 also appears in [8]. Without it, we can construct trivial examples of recurrent chains by having a probability of 10^{-6} (say) of jumping to the origin, regardless of the current location of the chain. Assumption 3 is global in nature, but Proposition 1 gives local conditions on X that are sufficient for Assumption 3 to hold. Proposition 1 is proved in Section 4.

Proposition 1. *Suppose that the manifold M and Markov chain X satisfy Assumptions 1 and 2. Suppose that $\mathbb{E}_x[D_{\text{rad}}] = 0$ for all $x \in M$, and that there exists $\varepsilon > 0$ such that $\mathbb{E}_x[D_{\text{rad}}^2] \geq \varepsilon$ for all $x \in M$. Then Assumption 3 holds.*

Definition 1. The chain X is called *O-recurrent* if there is some constant r_0 such that $\liminf_{n \rightarrow \infty} R_n \leq r_0$ almost surely. It is called *O-transient* if $R_n \rightarrow \infty$ almost surely.

The notions of *O*-recurrence and *O*-transience are local, and also appear in [8], although there they are simply called ‘recurrent’ and ‘transient’. For us, ‘recurrent’ means that, for any open $U \subset M$, X almost surely visits U infinitely often, and ‘transient’ means ‘not recurrent’. If X is *O*-transient for some choice of O (equivalently for all choices of O) then X is transient. If X is *O*-recurrent for some choice of O , and the chain is ‘irreducible’ (in some suitable sense), then we would expect X to be recurrent. For example, suppose that X is known to visit a neighbourhood N_0 of some origin O infinitely often almost surely, and has the property that, for every open neighbourhood N of M , there exist $m \in \mathbb{N}$ and $\delta > 0$ such that $\inf_{x \in N_0} \mathbb{P}_x[\tau_N \leq m] \geq \delta$, where $\tau_N := \min\{n \in \mathbb{N} : X_n \in N\}$. Then a Borel–Cantelli-type argument similar to [22, Example 2.3.20] reveals that X is recurrent. A full analysis of the relationship between *O*-recurrence and recurrence is technically involved since M is neither discrete nor countable; this is beyond the scope of this paper.

We introduce the remaining notation that we need to state our main result. Given $x \in M$ and a real number $d_{\text{tot}} \geq 0$ (we should think of d_{tot} as an observation of $D_{\text{tot}}^{(n)}$ for some n), the real-valued functions k_{min} and k_{max} are defined to measure the extremes of the sectional curvature within a distance d_{tot} of x . More precisely, we define

$$k_{\text{min}}(x, d_{\text{tot}}) = \inf_{\substack{y \in M : \text{Dist}_M(x, y) \leq d_{\text{tot}}, \\ \pi \in \text{Planes}(M, y)}} \sqrt{-\sec(y, \pi)},$$

$$k_{\text{max}}(x, d_{\text{tot}}) = \sup_{\substack{y \in M : \text{Dist}_M(x, y) \leq d_{\text{tot}}, \\ \pi \in \text{Planes}(M, y)}} \sqrt{-\sec(y, \pi)}.$$

By Assumption 1, $0 < \kappa \leq k_{\text{min}}(x, d_{\text{tot}}) \leq k_{\text{max}}(x, d_{\text{tot}})$ for all x, d_{tot} . Given $k > 0, d_{\text{tot}} > 0$, and $\phi \in [-1, 1]$, let G be the real-valued function

$$G(k, d_{\text{tot}}, \phi) = \frac{1}{k} \log(\cosh(kd_{\text{tot}}) + \phi \sinh(kd_{\text{tot}})). \tag{1}$$

We demonstrate in the proof of Proposition 3 that G is an asymptotic estimate, valid when R is much larger than D_{tot} , for the increment ΔR on a manifold of constant curvature $-k^2$. Finally, the notation $\mathbf{1}_E$, for an event E , refers to the indicator function that is equal to 1 on E and zero on the complement of E .

Theorem 1. *Let M and X satisfy Assumptions 1, 2, and 3. Let $S(r) = \{x \in M : \text{Dist}_M(O, x) = r\}$, and let*

$$\underline{\nu}_1(r) = \inf_{x \in S(r)} \mathbb{E}_x[G(k_{\text{min}}(x, D_{\text{tot}}), D_{\text{tot}}, \Phi)],$$

$$\bar{\nu}_1(r) = \sup_{x \in S(r)} \mathbb{E}_x[G(k_{\text{max}}(x, D_{\text{tot}}), D_{\text{tot}}, \Phi)],$$

$$\begin{aligned}\underline{\nu}_2(r) &= \inf_{x \in S(r)} \left\{ \mathbb{E}_x[G^2(k_{\min}(x, D_{\text{tot}}), D_{\text{tot}}, \Phi) \mathbf{1}_{\Delta R \geq 0}] \right. \\ &\quad \left. + \mathbb{E}_x[G^2(k_{\max}(x, D_{\text{tot}}), D_{\text{tot}}, \Phi) \mathbf{1}_{\Delta R < 0}] \right\}, \\ \bar{\nu}_2(r) &= \sup_{x \in S(r)} \left\{ \mathbb{E}_x[G^2(k_{\max}(x, D_{\text{tot}}), D_{\text{tot}}, \Phi) \mathbf{1}_{\Delta R \geq 0}] \right. \\ &\quad \left. + \mathbb{E}_x[G^2(k_{\min}(x, D_{\text{tot}}), D_{\text{tot}}, \Phi) \mathbf{1}_{\Delta R < 0}] \right\}.\end{aligned}$$

- (i) If $\liminf_{r \rightarrow \infty} (2r\underline{\nu}_1(r) - \bar{\nu}_2(r)) > 0$ then X is O -transient.
- (ii) If instead $\liminf_{r \rightarrow \infty} \underline{\nu}_2(r) > 0$ and there exist $r_0 \geq 0$, $\theta > 0$ such that

$$2r\bar{\nu}_1(r) \leq \left(1 + \frac{1 - \theta}{\log r}\right) \underline{\nu}_2(r) \quad \text{for all } r \geq r_0,$$

then X is O -recurrent.

Remark 1.

- (i) The intuition behind Theorem 1 is that R is a one-dimensional process, and so we expect (under suitable assumptions [6, 19]) the recurrence or transience of R to be determined by the behaviour of the mean and variance of the increments ΔR when R is large. The expressions $\bar{\nu}_i(r)$ and $\underline{\nu}_i(r)$, when r is large, are approximate upper and lower bounds for $\mathbb{E}_x[(\Delta R)^i]$ when $x \in S(r)$. They are analogous to the right-hand side of Equations (5.3) and (5.4) in [8].
- (ii) It is an immediate consequence of Theorem 1 that if, for $i = 1, 2$, we can find functions $\underline{\nu}'_i(r)$ and $\bar{\nu}'_i(r)$ such that $\underline{\nu}'_i \leq \underline{\nu}_i$ and $\bar{\nu}'_i \geq \bar{\nu}_i$, then Theorem 1 will hold with $\underline{\nu}_i$ and $\bar{\nu}_i$ replaced by $\underline{\nu}'_i$ and $\bar{\nu}'_i$ respectively. So, in applications we can find bounds for the $\underline{\nu}_i$ and $\bar{\nu}_i$ instead of evaluating them explicitly.
- (iii) Another immediate consequence of Theorem 1 is that the equation

$$\lim_{r \rightarrow \infty} r\underline{\nu}_1(r) = \infty \tag{2}$$

is sufficient for O -transience.

- (iv) If more information is known about the curvature of M , then the expressions for the $\underline{\nu}_i$ and $\bar{\nu}_i$ simplify. For example, if the sectional curvature is globally bounded by $-\kappa_1^2 \leq \text{sec} \leq -\kappa_2^2$, then k_{\min} , k_{\max} can be replaced by the constants κ_1 and κ_2 . Using the fact that $\mathbf{1}_{\Delta R \geq 0} + \mathbf{1}_{\Delta R < 0} = 1$, it follows that the indicator functions disappear from the formulae when M is a constant curvature manifold.
- (v) Corollary 1 gives sufficient conditions to ensure that $\Delta R \geq 0$ (and $\Delta R < 0$) for sufficiently large values of R_n , in terms of Φ and D_{tot} only. This allows the indicator functions $\mathbf{1}_{\Delta R \geq 0}$ and $\mathbf{1}_{\Delta R < 0}$ to be bounded in terms of purely local quantities, as claimed in the abstract.
- (vi) Not every chain that satisfies our assumptions can be classified by Theorem 1, since conditions (i) and (ii) of Theorem 1 are not exhaustive. We do not discuss the ambiguous case in any detail here, except to remark that if X has sufficient radial symmetry for R to be Markov, then we may be able to use the estimates in this paper, together with recurrence–transience results for one-dimensional processes (for example [6, Theorem 2.10]) to obtain a finer classification.

Definition 2. A chain X on M is called *zero drift* if $\mathbb{E}[\exp_{X_n}^{-1}(X_{n+1}) \mid \mathcal{F}_n] = 0$ almost surely for all $n \in \mathbb{N}$, where the conditional expectation is defined using the vector space structure of $T_{X_n}M$, and $\mathcal{F}_n = \sigma(X_1, X_2, \dots, X_n)$.

To say that a chain X is zero drift is to say that, for all $x \in M$, the conditional law of X_{n+1} , given that $X_n = x$, has x as its Riemannian centre of mass (or ‘barycentre’ [7]). This concept appears in the statistics and Monte Carlo literature as a way of taking means and medians on manifolds [2]. Zero-drift chains are also closely related to the notion of martingales on M . When $M = \mathbb{R}^d$ the two notions are equivalent, although for general M the story is slightly more complicated, as explained in [28].

Recall from [8, Equation (1.11)] that a chain $(X_n)_{n \in \mathbb{N}}$ is called *uniformly elliptic* if there exists $\varepsilon > 0$ such that

$$\mathbb{P}[(\exp_{X_n}^{-1}(X_{n+1}), w)_{X_n} \geq \varepsilon \mid \mathcal{F}_n] \geq \varepsilon \tag{3}$$

almost surely for all unit vectors $w \in T_{X_n}M$ and all $n \in \mathbb{N}$. In Euclidean space, zero-drift recurrent uniformly elliptic chains exist. By contrast, in Section 5 we derive the following consequence of Theorem 1.

Theorem 2. *Let M satisfy Assumption 1. Let X be a Markov chain on M satisfying Assumptions 2 and 3. If X is uniformly elliptic and of zero drift, then X is O -transient.*

We deduce Theorem 2 from a stronger result, namely that, under our assumptions, if a zero-drift chain is to be recurrent, then the function

$$Q(x) := \mathbb{E}_x[D_{\text{tot}}^2 - D_{\text{rad}}^2] \tag{4}$$

cannot remain bounded above zero as $\text{Dist}_M(x, O) \rightarrow \infty$. Geometrically, Q is the total variance of the increment, conditional on the chain currently being at $x \in M$, in the transverse direction. Some intuition as to why this stronger result is true is given at the start of Section 5. In the Euclidean case, Q provides a great deal of information as to the recurrence of a zero-drift chain: if we ignore certain boundary or degenerate cases, and assume that Q and $\mathbb{E}_x[D_{\text{tot}}^2]$ tend to limiting values when far from the origin, then these limiting values alone are enough to deduce recurrence or transience [8, Theorem 2.3]. By contrast, in the hyperbolic setting, rapid decay of Q is insufficient to imply recurrence, unless very strong assumptions are placed on the tails of the chain.

We now proceed as follows. Sections 3, 4, and 5 respectively prove Theorem 1, Proposition 1, and Theorem 2. Finally, Section 6 gives some examples.

3. Geometric calculations and proof of Theorem 1

For real numbers $k > 0$, $d_{\text{tot}} \geq 0$, $\phi \in [-1, 1]$, and $x \in M$, define a function F by

$$F(k, d_{\text{tot}}, \phi, x) = \frac{1}{k} \operatorname{arccosh}(\cosh kr \cosh kd_{\text{tot}} + \phi \sinh kr \sinh kd_{\text{tot}}) - r, \tag{5}$$

where $r = r(x) = \text{Dist}_M(x, O)$. Proposition 2 implies that if the manifold has constant curvature $-k^2$, then $F(k, D_{\text{tot}}^{(n)}, \Phi^{(n)}, x)$ is the exact value of the increment $R_{n+1} - R_n$, given that $X_n = x$. We stress that Proposition 2 is a purely geometric result; its proof does not require any probabilistic information.

Proposition 2. *Let M be a manifold satisfying Assumption 1. For brevity, we write $k_{\min}^{(n)}$ (or $k_{\max}^{(n)}$) in place of $k_{\min}(X_n, D_{\text{tot}}^{(n)})$ (or $k_{\max}(X_n, D_{\text{tot}}^{(n)})$). Recall also that $\Delta_n R := R_{n+1} - R_n$. Then, for all $n \in \mathbb{N}$,*

$$\begin{aligned}
 F(k_{\min}^{(n)}, D_{\text{tot}}^{(n)}, \Phi^{(n)}, R_n) &\leq \Delta_n R \leq F(k_{\max}^{(n)}, D_{\text{tot}}^{(n)}, \Phi^{(n)}, R_n), \\
 F^2(k_{\min}^{(n)}, D_{\text{tot}}^{(n)}, \Phi^{(n)}, R_n) \mathbf{1}_{\Delta_n R \geq 0} &+ F^2(k_{\max}^{(n)}, D_{\text{tot}}^{(n)}, \Phi^{(n)}, R_n) \mathbf{1}_{\Delta_n R < 0} \\
 &\leq (\Delta_n R)^2 \leq \\
 F^2(k_{\max}^{(n)}, D_{\text{tot}}^{(n)}, \Phi^{(n)}, R_n) \mathbf{1}_{\Delta_n R \geq 0} &+ F^2(k_{\min}^{(n)}, D_{\text{tot}}^{(n)}, \Phi^{(n)}, R_n) \mathbf{1}_{\Delta_n R < 0}.
 \end{aligned}$$

Proof. The first inequality, in the constant curvature case, is a classical result equivalent to the hyperbolic law of cosines [1]; it follows from the constant curvature case together with Toponogov’s theorem (see, for example, [4, Chapter 10, Proposition 2.5]). The second inequality follows from the first, together with the elementary observation that, for real numbers x and y such that $x \leq y$, we have $x^2 \leq y^2$ if $x \geq 0$ whereas $x^2 \geq y^2$ if $y \leq 0$. \square

Some routine algebraic manipulation of the first inequality in Proposition 2 results in the following corollary.

Corollary 1. *A sufficient condition for $R_{n+1} - R_n < 0$ is that*

$$\Phi^{(n)} < \frac{\coth k_{\max}^{(n)} R_n (1 - \cosh k_{\max}^{(n)} D_{\text{tot}}^{(n)})}{\sinh k_{\max}^{(n)} D_{\text{tot}}^{(n)}}.$$

In particular, for all $\varepsilon > 0$ there exists a constant r_ε such that, if $R_n > r_\varepsilon$ and

$$\Phi^{(n)} < \frac{(1 - \varepsilon)(1 - \cosh k_{\max}^{(n)} D_{\text{tot}}^{(n)})}{\sinh k_{\max}^{(n)} D_{\text{tot}}^{(n)}},$$

then $R_{n+1} - R_n < 0$. A sufficient condition for $R_{n+1} - R_n \geq 0$ is that

$$\Phi^{(n)} \geq \frac{\coth k_{\min}^{(n)} R_n (1 - \cosh k_{\min}^{(n)} D_{\text{tot}}^{(n)})}{\sinh k_{\min}^{(n)} D_{\text{tot}}^{(n)}}.$$

In particular, for all $\varepsilon > 0$ there exists a constant r_ε such that, if $R_n > r_\varepsilon$ and

$$\Phi^{(n)} \geq \frac{(1 + \varepsilon)(1 - \cosh k_{\min}^{(n)} D_{\text{tot}}^{(n)})}{\sinh k_{\min}^{(n)} D_{\text{tot}}^{(n)}},$$

then $R_{n+1} - R_n \geq 0$.

Proposition 3. *Assume that M has constant curvature $-k^2$, and that X satisfies Assumption 2. Let $x \in M$ be a point of distance R_x from O . Then*

$$\mathbb{E}_x[\Delta R] = \frac{1}{k} \mathbb{E}_x[\log(\cosh k D_{\text{tot}} + \Phi \sinh k D_{\text{tot}})] + O(R_x^{1-p}), \tag{6}$$

$$\mathbb{E}_x[(\Delta R)^2 \mathbf{1}_{\Delta R \geq 0}] = \frac{1}{k^2} \mathbb{E}_x[\log^2(\cosh k D_{\text{tot}} + \Phi \sinh k D_{\text{tot}}) \mathbf{1}_{\Delta R \geq 0}] + O(R_x^{2-p}), \tag{7}$$

where the implicit constants in the remainder terms depend only on $k, p,$ and $B,$ and remain bounded as a function of k as $k \rightarrow \infty.$ Moreover, (7) remains true if $\mathbf{1}_{\Delta R \geq 0}$ is changed to $\mathbf{1}_{\Delta R < 0}$ throughout.

Proof. For brevity, let $\alpha = \cosh kR_x, \beta = \sinh kR_x, c = \cosh kD_{\text{tot}},$ and $s = \sinh kD_{\text{tot}}.$ Using Proposition 2, followed by some algebraic manipulation, we find that

$$\begin{aligned} \left| \Delta R - \frac{1}{k} \log(c + \phi s) \right| &= \left| \frac{1}{k} \operatorname{arccosh}(\alpha c + \Phi \beta s) - R_x - \frac{1}{k} \log(c + \Phi s) \right| \tag{8} \\ &= \left| \frac{1}{k} \log(2(\alpha c + \Phi \beta s)) + \frac{1}{k} \log\left(\frac{1 + \sqrt{1 - (\alpha c + \Phi \beta s)^{-2}}}{2}\right) - R_x - \frac{1}{k} \log(c + \Phi s) \right| \\ &= \left| \frac{1}{k} \log\left(\frac{2(\alpha c + \Phi \beta s)e^{-kR_x}}{c + \Phi s}\right) + \frac{1}{k} \log\left(\frac{1 + \sqrt{1 - (\alpha c + \Phi \beta s)^{-2}}}{2}\right) \right| \\ &= \left| \frac{1}{k} \log\left(\frac{c(1 + e^{-2kR_x}) + \Phi s(1 - e^{-2kR_x})}{c + \Phi s}\right) + \frac{1}{k} \log\left(\frac{1 + \sqrt{1 - (\alpha c + \Phi \beta s)^{-2}}}{2}\right) \right| \\ &= \left| \frac{1}{k} \log\left(1 + \frac{c - \Phi s}{c + \Phi s} e^{-2kR_x}\right) + \frac{1}{k} \log\left(\frac{1 + \sqrt{1 - (\alpha c + \Phi \beta s)^{-2}}}{2}\right) \right| \\ &\leq \left| \frac{1}{k} \log\left(1 + \frac{c - \Phi s}{c + \Phi s} e^{-2kR_x}\right) \right| + \left| \frac{1}{k} \log\left(\frac{1 + \sqrt{1 - (\alpha c + \Phi \beta s)^{-2}}}{2}\right) \right|. \end{aligned}$$

Note that, since $\alpha, \beta, c, s \geq 0$ and $\Phi \in [-1, 1],$

$$e^{-kD_{\text{tot}}} = c - s \leq c + s = e^{kD_{\text{tot}}}. \tag{9}$$

In particular, $c \pm \Phi s \geq 0.$ Further,

$$\alpha c + \Phi \beta s \geq \alpha c - \beta s = \cosh(kR_x - kD_{\text{tot}}) \geq \max\left(\frac{1}{2} \exp(kR_x - kD_{\text{tot}}), 1\right). \tag{10}$$

We can verify that if $u \geq 0$ then $0 \leq \log(1 + u) \leq u,$ and that if $u \geq 1$ then

$$\left| \log\left(\frac{1 + \sqrt{1 - u^{-2}}}{2}\right) \right| \leq \frac{1}{u^2}.$$

It follows that

$$\begin{aligned} \left| \Delta R - \frac{1}{k} \log(c + \Phi s) \right| &\leq \frac{1}{k} \frac{c - \Phi s}{c + \Phi s} e^{-2kR_x} + \frac{1}{k} \frac{1}{(\alpha c + \Phi \beta s)^2} \\ &\leq \frac{1}{k} e^{2k(D_{\text{tot}} - R_x)} + \frac{4}{k} e^{2k(D_{\text{tot}} - R_x)} \quad \text{[by (9) and (10)]} \\ &= \frac{5}{k} e^{2k(D_{\text{tot}} - R_x)}. \end{aligned} \tag{11}$$

Let E be the event that $D_{\text{tot}} \leq R/2,$ and E^c its complement. Then

$$\begin{aligned} \left| \mathbb{E}_x \left[\Delta R - \frac{1}{k} \log(c + \Phi s) \right] \right| &\leq \left| \mathbb{E}_x \left[\left(\Delta R - \frac{1}{k} \log(c + \Phi s) \right) \mathbf{1}_E \right] \right| \\ &\quad + \left| \mathbb{E}_x \left[\left(\Delta R - \frac{1}{k} \log(c + \Phi s) \right) \mathbf{1}_{E^c} \right] \right| =: Q_1 + Q_2. \end{aligned}$$

Using (11),

$$Q_1 \leq \mathbb{E} \left[\frac{5}{k} e^{2k(D_{\text{tot}} - R_x)} \mathbf{1}_E \right] \leq \mathbb{E} \left[\frac{5}{k} e^{-kR_x} \mathbf{1}_E \right] \leq \frac{5}{k} e^{-kR_x}.$$

To bound Q_2 , we use (9) and then Assumption 2:

$$Q_2 \leq \mathbb{E}_x[2D_{\text{tot}} \mathbf{1}_{E^c}] = \mathbb{E}_x[2D_{\text{tot}}^p D_{\text{tot}}^{1-p} \mathbf{1}_{E^c}] \leq \left(\frac{R_x}{2}\right)^{1-p} \mathbb{E}_x[2D_{\text{tot}}^p \mathbf{1}_{E^c}] \leq 2^p R_x^{1-p} B.$$

Combining these bounds establishes (6). For (7), we note, from the elementary observation that $a^2 - b^2 = (a - b)^2 + 2b(a - b)$, that

$$\begin{aligned} & \left| \mathbb{E}_x \left[(\Delta R)^2 \mathbf{1}_{\Delta R \geq 0} - \left(\frac{1}{k} \log(c + \Phi s)\right)^2 \mathbf{1}_{\Delta R \geq 0} \right] \right| \leq \\ & \mathbb{E}_x \left[\left| \left(\left(\Delta R - \frac{1}{k} \log(c + \Phi s)\right)^2 \mathbf{1}_E + \frac{2}{k} \log(c + \Phi s) \left(\Delta R - \frac{1}{k} \log(c + \Phi s)\right) \mathbf{1}_E \right) \mathbf{1}_{\Delta R \geq 0} \right| \right] \\ & + \mathbb{E}_x \left[\left| \left(\left(\Delta R - \frac{1}{k} \log(c + \Phi s)\right)^2 \mathbf{1}_{E^c} + \frac{2}{k} \log(c + \Phi s) \left(\Delta R - \frac{1}{k} \log(c + \Phi s)\right) \mathbf{1}_{E^c} \right) \mathbf{1}_{\Delta R \geq 0} \right| \right] \\ & =: Q_3 + Q_4. \end{aligned}$$

Using (11), we find that

$$\begin{aligned} Q_3 & \leq \mathbb{E}_x \left[\left(\left(\frac{5}{k} e^{2k(D_{\text{tot}} - R_x)}\right)^2 + 2D_{\text{tot}} \left(\frac{5}{k} e^{2k(D_{\text{tot}} - R_x)}\right) \right) \mathbf{1}_E \right] \\ & \leq \mathbb{E}_x \left[\frac{25}{k^2} e^{-2kR_x} + \frac{10}{k} e^{-kR_x} D_{\text{tot}} \right]. \end{aligned}$$

Assumption 2 together with Lyapunov’s inequality implies that $\mathbb{E}_x[D_{\text{tot}}]$ is bounded as a function of x , giving a bound on Q_3 of the required form. To bound Q_4 , Assumption 2 gives

$$Q_4 \leq \mathbb{E}_x[8D_{\text{tot}}^2 \mathbf{1}_{E^c}] = \mathbb{E}_x[8D_{\text{tot}}^p D_{\text{tot}}^{2-p} \mathbf{1}_{E^c}] \leq 2^{p+1} R^{2-p} \mathbb{E}_x[D_{\text{tot}}^p \mathbf{1}_{E^c}] \leq 2^{p+1} B R^{2-p},$$

from which we deduce (7). Finally, the same proof as above shows that (7) holds when $\mathbf{1}_{\Delta R \geq 0}$ is changed to $\mathbf{1}_{\Delta R < 0}$ throughout. \square

3.1. Proof of Theorem 1

Recall that the functions F and G are defined in (5) and (1), respectively.

Proof. In what follows, C is a constant that depends only on κ , p , and B , but may change from line to line. We note that

$$\begin{aligned} \mathbb{E}_x[\Delta R] & \leq \mathbb{E}_x[F(k_{\max}(x, D_{\text{tot}}), D_{\text{tot}}, \Phi, x)] \\ & \leq \mathbb{E}_x[G(k_{\max}(x, D_{\text{tot}}), D_{\text{tot}}, \Phi)] + Cr^{1-p}, \end{aligned} \tag{12}$$

where $r = \text{Dist}_M(x, O)$. The first inequality here is from Proposition 2. The second follows from the proof of Proposition 3, noting that the right-hand side of (8) is exactly $|F(k, D_{\text{tot}}, \Phi, x) - G(k, D_{\text{tot}}, \Phi)|$. Similarly, we see that

$$\begin{aligned} \mathbb{E}_x[(\Delta R)^2] &\leq \mathbb{E}_x[F^2(k_{\text{max}}(x, D_{\text{tot}}), D_{\text{tot}}, \Phi, x)\mathbf{1}_{\Delta R \geq 0}] \\ &\quad + \mathbb{E}_x[F^2(k_{\text{min}}(x, D_{\text{tot}}), D_{\text{tot}}, \Phi, R)\mathbf{1}_{\Delta R < 0}] \\ &\leq \mathbb{E}_x[G^2(k_{\text{max}}(x, D_{\text{tot}}), D_{\text{tot}}, \Phi)\mathbf{1}_{\Delta R \geq 0}] \\ &\quad + \mathbb{E}_x[G^2(k_{\text{min}}(x, D_{\text{tot}}), D_{\text{tot}}, \Phi)\mathbf{1}_{\Delta R < 0}] + Cr^{2-p}. \end{aligned} \tag{13}$$

Define, for $i = 1$ and $i = 2$,

$$\underline{\mu}_i(r) = \underline{\nu}_i(r) - Cr^{i-p}, \quad \bar{\mu}_i(r) = \bar{\nu}_i(r) + Cr^{i-p}. \tag{14}$$

Taking suprema of (12) and (13) over $S(r)$, and infima of the corresponding lower bounds over $S(r)$, it follows that

$$\underline{\mu}_i(R_n) \leq \mathbb{E}[(\Delta R_n)^i \mid \mathcal{F}_n] \leq \bar{\mu}_i(R_n) \tag{15}$$

almost surely for all n , where \mathcal{F}_n is the sigma algebra generated by R_1, \dots, R_n .

Suppose now that the assumptions in Theorem 1(i) hold. It follows from (14), together with the fact that $p > 2$, that $\limsup_{r \rightarrow \infty} \bar{\mu}_2(r) < \infty$ and $\liminf_{r \rightarrow \infty} (2r\underline{\mu}_1(r) - \bar{\mu}_2(r)) > 0$. Theorem 1(i) now follows from these two expressions and (15), together with [22, Theorem 3.5.1].

Suppose instead that the assumptions in Theorem 1(ii) hold for some constants $r_0 \geq 0$ and $\theta > 0$, where $\theta = 2\theta'$. Then, with the constant C as given in (14),

$$\begin{aligned} 2r\bar{\mu}_1 - \left(1 + \frac{1 - \theta'}{\log r}\right)\underline{\mu}_2 &= 2r\bar{\nu}_1 - \left(1 + \frac{1 - \theta'}{\log r}\right)\underline{\nu}_2 + C\left(3 + \frac{1 - \theta'}{\log r}\right)r^{2-p} \\ &= \left(2r\bar{\nu}_1 - \left(1 + \frac{1 - \theta}{\log r}\right)\underline{\nu}_2\right) - \frac{\theta'\underline{\nu}_2}{\log r} + C\left(3 + \frac{1 - \theta'}{\log r}\right)r^{2-p}. \end{aligned}$$

By assumption, $\liminf_{r \rightarrow \infty} \underline{\nu}_2 > 0$, and therefore the second term of the last equality decays more slowly than the third as $r \rightarrow \infty$. It follows that there exists r'_0 such that, for all $r \geq r'_0$,

$$2r\bar{\mu}_1 - \left(1 + \frac{1 - \theta'}{\log r}\right)\underline{\mu}_2 \leq 0.$$

Theorem 1(ii) follows from this and (15), together with [22, Theorem 3.5.2]. □

4. Non-confinement

In this section we prove Proposition 1. Our strategy is to use martingale arguments to deduce Proposition 1 from a similar result in Euclidean space [8, Proposition 2.1]. Compared to that result, ours applies to a wider class of processes (not just martingales), but at the price of being a little more restrictive: we require $\mathbb{E}[D_{\text{rad}}^2] \geq \varepsilon$ as opposed to $\mathbb{E}[D_{\text{tot}}^2] \geq \varepsilon$. Let us note also that other non-confinement criteria are known (one example is [22, Equation (3.10)]), and proving non-confinement is often straightforward in practice. Let \mathcal{F} be the filtration that is naturally generated by X .

Proposition 4. *Let X be a Markov chain on a manifold M satisfying Assumption 1. Assume that $\mathbb{E}_x[D_{\text{rad}}] \geq 0$ for all $x \in M$. Then the radial process R is a non-negative \mathcal{F} -submartingale.*

Proof. By Toponogov’s theorem, it is enough to prove this in the case where M is Euclidean. Using the cosine rule for triangles in \mathbb{R}^2 , we can show that

$$R_{n+1} = R_n \sqrt{1 + \frac{2D_{\text{rad}}^{(n)}}{R_n} + \frac{(D_{\text{tot}}^{(n)})^2}{R_n^2}}.$$

Since $D_{\text{tot}}^n \geq |D_{\text{rad}}^n|$, we deduce that

$$R_{n+1} - R_n \geq R_n \left(\left[\left(1 + \frac{D_{\text{rad}}^{(n)}}{R_n} \right)^2 \right]^{1/2} - 1 \right) \geq D_{\text{rad}}^{(n)}, \tag{16}$$

and the result follows upon taking expectations. □

In the remainder of this section, $Y = (Y_n)_{n \in \mathbb{N}}$ is a process in \mathbb{R}^d adapted to \mathcal{F} , such that $Y_0 = 0$. The processes L and A are defined using the Doob decomposition of Y [30]. More precisely, the j th one-dimensional component of Y is given by $Y_n^j = L_n^j + A_n^j$, where, for each j , $1 \leq j \leq d$, L^j is an \mathcal{F} -martingale and A^j is a predictable process, with the property that if Y^j is a submartingale then A^j is non-negative and increasing.

Lemma 1. *Assume that $\mathbb{E}[|\Delta_n L|^p \mid \mathcal{F}_n] \leq B$ and $\mathbb{E}[|\Delta_n A|^p \mid \mathcal{F}_n] \leq B$ for all n almost surely for some $p > 2$, $B \in \mathbb{R}$. After enlarging the probability space if necessary, consider the process $(Z_n)_{n \in \mathbb{N}}$ given by $Z_0 = Y_0$ and $(\Delta_n Z) = (\Delta_n L) + \xi_{n+1}(\Delta_n A)$, where the ξ_n are equal to ± 1 with equal probability, independently of each other, L , or A . Let $\mathcal{G}_0 = \mathcal{F}_0$ and, for each $n \geq 1$, let \mathcal{G}_n be the sigma algebra generated by \mathcal{F}_n and (ξ_1, \dots, ξ_n) . Then:*

- (i) $\mathbb{E}[|\Delta_n Z|^p \mid \mathcal{G}_n] \leq B'$ for some B' depending only on B , p , and d ;
- (ii) Z is a \mathcal{G} -martingale;
- (iii) $\mathbb{E}[|\Delta_n Y|^2 \mid \mathcal{G}_n] = \mathbb{E}[|\Delta_n Z|^2 \mid \mathcal{G}_n]$.

Proof. (i) This follows from the bounds on $\mathbb{E}[|\Delta L|^p]$ and $\mathbb{E}[|\Delta A|^p]$, and the inequality $|x + y|^p \leq C_{d,p}(|x|^p + |y|^p)$ for vectors $x, y \in \mathbb{R}^d$, where $C_{d,p}$ is a constant.

(ii) We check that $\mathbb{E}[\Delta_n Z \mid \mathcal{G}_n] = \mathbb{E}[\Delta_n L \mid \mathcal{G}_n] + (\Delta_n A)\mathbb{E}[\xi_{n+1} \mid \mathcal{G}_n] = 0 + 0 = 0$, and that, by part (i) and Lyapunov’s inequality, there is a constant B'' such that $\mathbb{E}[|Z_n|] \leq \sum_{k=0}^{n-1} \mathbb{E}[|\Delta_k Z|] \leq nB'' < \infty$.

(iii) We calculate

$$\begin{aligned} \mathbb{E}[|\Delta_n Z|^2 \mid \mathcal{G}_n] &= \sum_{j=1}^d \mathbb{E} \left[(\Delta_n L^j + \xi_{n+1} \Delta_n A^j)^2 \mid \mathcal{G}_n \right] \\ &= \sum_{j=1}^d \mathbb{E} \left[(\Delta_n L^j)^2 + 2\xi_{n+1} \Delta_n A^j \Delta_n L^j + (\Delta_n A^j)^2 \mid \mathcal{G}_n \right] \\ &= \mathbb{E}[|\Delta_n L|^2 + |\Delta_n A|^2 \mid \mathcal{G}_n], \end{aligned}$$

and, similarly, $\mathbb{E}[(\Delta_n Y)^2 \mid \mathcal{G}_n] = \mathbb{E}[(\Delta_n L)^2 + (\Delta_n A)^2 \mid \mathcal{G}_n]$. □

Proposition 5. *Suppose that Y is an \mathbb{R} -valued submartingale. Assume that there exist $B \in \mathbb{R}_+$ and $p > 2$ such that $\mathbb{E}[|\Delta_n L|^p \mid \mathcal{F}_n] \leq B$ and $\mathbb{E}[|\Delta_n A|^p \mid \mathcal{F}_n] \leq B$ for all n . Assume also that there exists $\varepsilon > 0$ such that $\mathbb{E}[|\Delta_n Y|^2 \mid \mathcal{F}_n] \geq \varepsilon$ for all n , and that*

$$\mathbb{P}\left[\limsup_{n \rightarrow \infty} L_n > -\infty\right] = 1. \tag{17}$$

Then $\mathbb{P}[\limsup_{n \rightarrow \infty} |Y_n| = \infty] = 1$. In particular, if Y is bounded below almost surely then $\mathbb{P}[\limsup_{n \rightarrow \infty} Y_n = \infty] = 1$.

Proof. We work in the enlarged probability space described in Lemma 1. Let $Z = (Z_n)_{n \in \mathbb{N}}$ be as in Lemma 1. We claim that

$$\begin{aligned} \left\{ \omega \in \Omega : \limsup_{n \rightarrow \infty} |Y_n| = \infty \right\} &\supseteq \left\{ \omega \in \Omega : \limsup_{n \rightarrow \infty} |Z_n| = \infty \right\} \\ &\cap \left\{ \omega \in \Omega : \limsup_{n \rightarrow \infty} L_n > -\infty \right\}. \end{aligned} \tag{18}$$

To see this, first suppose that $\omega \in \Omega$ is such that A is bounded, say $A_n \leq W$ for all n . Then $Y_n \geq Z_n - 2W$ for all n . On the other hand, if A is unbounded, then, since A is positive and increasing, $\lim_{n \rightarrow \infty} A_n = \infty$ and so $\limsup_{n \rightarrow \infty} A_n + L_n$ will be infinity provided that there exists $W \in \mathbb{R}$ such that $L_n > W$ infinitely often. This establishes (18). By 17, it suffices to prove that $\limsup_{n \rightarrow \infty} |Z_n| = \infty$ almost surely. Lemma 1, combined with [8, Proposition 2.1], establishes this result. \square

Proof of Proposition 1. Decompose the radial process as $R_n = L_n + A_n$. It follows from Proposition 4 and the uniqueness of the Doob decomposition that $L_n = \sum_{i=1}^n D_{\text{rad}}^{(i)}$. It suffices to check that the assumptions of Proposition 5 hold when $Y_n = R_n$. First, it is almost surely the case that $|\Delta_n L| = |D_{\text{rad}}^{(n)}| \leq D_{\text{tot}}^{(n)}$ and $|\Delta_n L + \Delta_n A| \leq D_{\text{tot}}^{(n)}$, and hence $|\Delta_n A| \leq 2D_{\text{tot}}^{(n)}$. Assumption 2 then gives the required bounds on $\mathbb{E}[|\Delta_n L|^p \mid \mathcal{F}_n]$ and $\mathbb{E}[|\Delta_n A|^p \mid \mathcal{F}_n]$. Second, using (16), $\mathbb{E}[|\Delta_n R|^2 \mid \mathcal{F}_n] \geq \mathbb{E}[|\Delta_n L|^2 \mid \mathcal{F}_n] \geq \varepsilon$. Finally, [8, Theorem 2.1] shows that almost surely there is a bounded neighbourhood of the origin N such that $L_n \in N$ infinitely often, so (17) holds. \square

5. Uniform ellipticity and proof of Theorem 2

In this section we prove Theorem 2, which states that any uniformly elliptic zero-drift chain must be transient. We first give an intuitive explanation for why this is true. Let M be a manifold of constant curvature $-k^2$. Suppose that the chain is currently a distance $R \gg 1$ from the origin, and that it makes a purely transverse step of unit length (‘transverse’ means perpendicular to the geodesic joining the origin to the chain’s current location). Then the (Euclidean or hyperbolic) Pythagorean theorem reveals that the change in the chain’s distance from the origin is given by

$$\Delta R = \begin{cases} R(\sqrt{1 + R^{-2}} - 1) & \text{if } k = 0, \\ k^{-1} \operatorname{arccosh}(\cosh kR \cosh k) - R & \text{if } k < 0. \end{cases}$$

Expanding this to first order in R^{-1} , we see that as $R \rightarrow \infty$, ΔR tends to a limit that is zero if $k = 0$, but *positive* if $k < 0$. For this reason, if the variance in the transverse direction remains

bounded above zero, then we would expect a zero-drift chain to be transient. Theorem 2 makes this intuition precise.

Lemma 2. For all real numbers $k > 0$, $d_{\text{tot}} \geq 0$, and $\phi \in [-1, 1]$,

$$d_{\text{rad}} + J_{\min}(k, d_{\text{tot}})(d_{\text{tot}}^2 - d_{\text{rad}}^2) \leq G(k, d_{\text{tot}}, \phi) \leq d_{\text{rad}} + J_{\max}(k, d_{\text{tot}})(d_{\text{tot}}^2 - d_{\text{rad}}^2),$$

where

$$\begin{aligned} d_{\text{rad}} &= \phi d_{\text{tot}}, \\ J_{\min}(k, d_{\text{tot}}) &= \frac{1}{2d_{\text{tot}}^2} \left(d_{\text{tot}} - \frac{\sinh(kd_{\text{tot}})}{k(\cosh(kd_{\text{tot}}) + \sinh(kd_{\text{tot}}))} \right), \\ J_{\max}(k, d_{\text{tot}}) &= \frac{1}{2d_{\text{tot}}^2} \left(-d_{\text{tot}} + \frac{\sinh(kd_{\text{tot}})}{k(\cosh(kd_{\text{tot}}) - \sinh(kd_{\text{tot}}))} \right). \end{aligned}$$

Moreover, J_{\min} is positive, increasing in k , and decreasing in d_{tot} , whereas J_{\max} is non-negative and increasing in both k and d_{tot} .

Proof. For fixed k and d_{tot} , consider the function

$$H(\phi) := \left(\frac{1}{k} \log(\cosh(kd_{\text{tot}}) + \phi \sinh(kd_{\text{tot}})) - \phi d_{\text{tot}} \right) (1 - \phi^2)^{-1}.$$

It is lengthy but elementary to check that H is decreasing on $\phi \in [-1, 1]$ and that its limits at $\phi = \pm 1$ are

$$\frac{1}{2} \left(\pm d_{\text{tot}} \mp \frac{\sinh(kd_{\text{tot}})}{k(\cosh(kd_{\text{tot}}) \pm \sinh(kd_{\text{tot}}))} \right).$$

The first part follows, and the remainder follows from a direct check. □

In the following theorem, the function Q is as defined in (4).

Theorem 3. Let X be a Markov chain on a manifold M , and suppose that Assumptions 1, 2, and 3 all hold. Suppose also that X is of zero drift and that there exist constants $d_{\min} \geq 0$ and $\varepsilon > 0$ such that $Q(x) \geq \varepsilon$ for every $x \in M$ such that $\text{Dist}_M(O, x) \geq d_{\min}$. Then X is O -transient.

Proof. Let $c = \cosh(\kappa D_{\text{tot}})$, $s = \sinh(\kappa D_{\text{tot}})$, and A be a constant to be chosen later. Using Lemma 2, we obtain

$$\begin{aligned} \mathbb{E}_x[G(\kappa, D_{\text{tot}}, \Phi)] &\geq \mathbb{E}_x[D_{\text{rad}}] + \frac{1}{2} \mathbb{E}_x \left[\frac{1}{D_{\text{tot}}^2} \left(D_{\text{tot}} - \frac{s}{\kappa(c+s)} \right) (D_{\text{tot}}^2 - D_{\text{rad}}^2) \right] \\ &\geq \frac{1}{2} \mathbb{E}_x \left[\frac{1}{D_{\text{tot}}^2} \left(D_{\text{tot}} - \frac{s}{\kappa(c+s)} \right) (D_{\text{tot}}^2 - D_{\text{rad}}^2) \mathbf{1}_{D_{\text{tot}} < A} \right] \\ &= \frac{1}{2} \mathbb{E}_x \left[\left(\frac{1}{D_{\text{tot}}} - \frac{1}{2\kappa D_{\text{tot}}^2} (1 - e^{-2\kappa D_{\text{tot}}}) \right) (D_{\text{tot}}^2 - D_{\text{rad}}^2) \mathbf{1}_{D_{\text{tot}} < A} \right] \\ &\geq \frac{1}{2} \mathbb{E}_x \left[\left(\frac{1}{A} - \frac{1}{2\kappa A^2} (1 - e^{-2\kappa A}) \right) (D_{\text{tot}}^2 - D_{\text{rad}}^2) \mathbf{1}_{D_{\text{tot}} < A} \right], \end{aligned}$$

where the last line follows from the fact that the function

$$\frac{1}{d_{\text{tot}}} - \frac{1}{2\kappa d_{\text{tot}}^2} (1 - e^{-2\kappa d_{\text{tot}}})$$

is positive and decreasing in d_{tot} . Therefore, there is a constant A_0 , depending only on κ , such that, if $A > A_0$ and $\text{Dist}_M(O, x) \geq d_{\text{min}}$, then

$$\begin{aligned} \mathbb{E}_x[G(\kappa, D_{\text{tot}}, \Phi)] &\geq \frac{1}{4A} \mathbb{E}_x[(D_{\text{tot}}^2 - D_{\text{rad}}^2) \mathbf{1}_{D_{\text{tot}} < A}] \\ &= \frac{1}{4A} (\mathbb{E}_x[D_{\text{tot}}^2 - D_{\text{rad}}^2] - \mathbb{E}_x[(D_{\text{tot}}^2 - D_{\text{rad}}^2) \mathbf{1}_{D_{\text{tot}} > A}]) \\ &\geq \frac{1}{4A} (\varepsilon - \mathbb{E}_x[D_{\text{tot}}^2 \mathbf{1}_{D_{\text{tot}} > A}]). \end{aligned}$$

By Assumption 2,

$$\mathbb{E}_x[D_{\text{tot}}^2 \mathbf{1}_{D_{\text{tot}} < A}] = \mathbb{E}_x[D_{\text{tot}}^p \cdot D_{\text{tot}}^{2-p} \mathbf{1}_{D_{\text{tot}} > A}] \leq \mathbb{E}_x[D_{\text{tot}}^p] A^{2-p} \leq BA^{2-p}.$$

Choose A sufficiently large that $BA^{2-p} \leq \varepsilon/2$. This then gives

$$\mathbb{E}_x[G(\kappa, D_{\text{tot}}, \Phi)] \geq \frac{\varepsilon}{2}. \tag{19}$$

Equation (2) therefore holds, and X is O -transient. □

Proof of Theorem 2. Conditional on $X_n = x$, view $v(x) = \exp_x^{-1}(X_{n+1})$ as a random vector in T_xM . Decompose $v(x)$ as $v = \lambda_{\text{rad}} e_{\text{rad}}(x) + w(x)$, where $\lambda_{\text{rad}} \in \mathbb{R}$ and $w(x)$ is a random vector orthogonal to $e_{\text{rad}}(x)$. For each $\omega \in \Omega$, $w(x)$ is uniquely defined and does not depend on a choice of basis for T_xM , because O is fixed and uniquely determines $e_{\text{rad}}(x)$. We may now interpret $Q(x)$ geometrically as $Q(x) = \mathbb{E}_x[\langle w(x), w(x) \rangle_x]$, the expected squared length of the transverse component of v .

Choose a unit-length vector $e_{\text{trans}}(x)$ perpendicular to $e_{\text{rad}}(x)$. By the Cauchy–Schwarz inequality, $\langle w, w \rangle_x \geq \langle w, e_{\text{trans}}(x) \rangle_x^2 = \langle v, e_{\text{trans}}(x) \rangle_x^2$, and hence

$$Q(x) \geq \mathbb{E}_x[\langle v, e_{\text{trans}}(x) \rangle_x^2]. \tag{20}$$

If X is uniformly elliptic in the sense of (3) then there exists $\varepsilon > 0$ such that $\mathbb{P}_x[\langle v, e_{\text{trans}}(x) \rangle_x^2 \geq \varepsilon^2] \geq \varepsilon$ for every $x \in M$. It follows from (20) that $Q(x) \geq \varepsilon^3$ for all x , and X is O -transient by Theorem 3. □

Remark 2. We finish this section with a comment on the rate of escape of transient processes. Assumption 2 and the triangle inequality imply that there is a constant C such that $\mathbb{E}_x[\Delta R] \leq C$ for all x . Therefore, under our Assumptions 1, 2, and 3, [23, Theorem 2.3] (with $\beta = 0$) shows that there exists a constant Λ such that, almost surely,

$$\limsup_{n \rightarrow \infty} \frac{R_n}{n} \leq \Lambda. \tag{21}$$

In the case of a zero-drift, uniformly elliptic chain, (19) together with the same theorem from [23] gives a constant λ such that $\lambda \leq \liminf_{n \rightarrow \infty} \frac{X_n}{n}$ almost surely, which in combination with (21) gives a linear rate of escape. As mentioned in the introduction, it is a general theme in the literature that ‘hyperbolic’ processes tend to escape at linear speed. However, as we shall see in the next section, by reducing the transverse component when far from the origin (thereby losing uniform ellipticity) we may reduce the rate of escape and eventually obtain recurrence.

6. Examples

Let V be a finite-dimensional inner product space of dimension d . Given $v \in V$ and $a, b > 0$, define $L_V(a, b, v) : V \rightarrow V$ to be the linear transformation that sends v to $av\sqrt{d}$ and any $w \in \langle v \rangle^\perp$ to $bw\sqrt{d}$. Define an elliptical measure $\xi_V(a, b, v) : \text{Borel}(V) \rightarrow \mathbb{R}_{\geq 0}$ by $\xi_V = \mu_V \circ L_V^{-1}(a, b, v)$, where μ_V is the uniform measure on the unit sphere in V . Thus, ξ_V is supported on an ellipsoid with principal axes of lengths $a\sqrt{d}, b\sqrt{d}, \dots, b\sqrt{d}$. Given a d -dimensional manifold M with origin $O \in M$, and functions $a, b : M \rightarrow \mathbb{R}_{\geq 0}$, define a measure $\mu_p : \text{Borel}(M) \rightarrow \mathbb{R}$ at each point $p \in M$ by $\mu_p = \xi_{T_p M}(a(p), b(p), e_{\text{rad}}(p)) \circ \exp_p^{-1}$, where $e_{\text{rad}}(O)$ is defined arbitrarily as some fixed unit-length vector in $T_O M$ (as far as recurrence and transience is concerned, this choice is unimportant). This defines what we refer to as the *elliptic Markov chain* with parameters a and b . In the case where a and b are constant, and M is Euclidean space, the elliptic Markov chain reduces to the example in [8, Section 3].

We claim that, for the elliptic Markov chain,

$$\mathbb{E}_p[D_{\text{tot}}^2] = a(p)^2 + (d - 1)b(p)^2, \quad \mathbb{E}_p[D_{\text{rad}}^2] = a(p)^2. \tag{22}$$

To prove this, note that the computation in [8, p.7] establishes this result when V has the Euclidean inner product. The general result follows from the definition of ξ_V together with the fact that any two inner product spaces of dimension d whose inner products are positive definite are isometric.

We study a special case that allows us to give a recurrence criterion that relates the asymptotic behaviour of b and the curvature of the manifold. Choose $a(r) = a$ for all r , where $a > 0$ is a positive constant. For constants $c > 0$ and $\gamma \geq 0$, choose the curvature of M such that, if p is a point at a distance r from O , then

$$\text{sec}(p, \pi) = \begin{cases} -c^2 & \text{if } r \leq 1, \\ -(cr^\gamma)^2 & \text{if } r \geq 1 \end{cases}$$

for all $\pi \in \text{Planes}(M, p)$. Since $a > 0$, it follows from Proposition 1 that Assumption 3 holds. Further, assume that $b(r) \leq b$ for all r . Then Assumption 2 holds, because, almost surely,

$$D_{\text{tot}}^{(n)} \leq d_{\text{max}} \tag{23}$$

for all $n \in \mathbb{N}$, where $d_{\text{max}} := \sqrt{d} \max(a, b)$. To apply Theorem 1, we need to compute $\underline{\nu}_1, \underline{\nu}_2, \bar{\nu}_1$, and $\bar{\nu}_2$. Although this could be done by choosing coordinates and writing down a multidimensional integral, we instead obtain estimates using only (22) and (23). This better enables comparison with the results in [8].

Since M is radially symmetric, we may unambiguously write \mathbb{E}_r to mean \mathbb{E}_x for any $x \in S(r)$. Using the fact that $\mathbb{E}_r[D_{\text{rad}}] = 0$, together with (22) and (23), we obtain that, for all sufficiently large r ,

$$\begin{aligned} \underline{\nu}_1(r) &= \mathbb{E}_r[G(k_{\text{min}}(r, D_{\text{tot}}), D_{\text{tot}}, \Phi)] \\ &\geq \mathbb{E}_r[J_{\text{min}}(k_{\text{min}}(r, D_{\text{tot}}), D_{\text{tot}})(D_{\text{tot}}^2 - D_{\text{rad}}^2)] \\ &\geq J_{\text{min}}(k_{\text{min}}(r, d_{\text{max}}), d_{\text{max}})\mathbb{E}_r[(D_{\text{tot}}^2 - D_{\text{rad}}^2)] \\ &= J_{\text{min}}(k_{\text{min}}(r, d_{\text{max}}), d_{\text{max}})(d - 1)b(r)^2 \\ &= J_{\text{min}}(c(r - d_{\text{max}})^\gamma, d_{\text{max}})(d - 1)b(r)^2. \end{aligned}$$

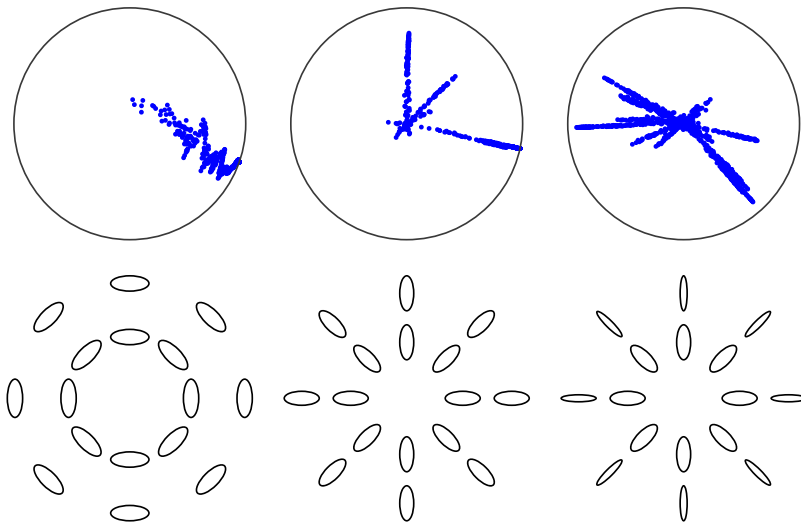


FIGURE 1. Simulations of the elliptic Markov chain with parameters $a(r)$ and $b(r)$ in the hyperbolic plane. The law of each simulation in the upper row is given schematically by the corresponding picture in the lower row. We take $a(r) = a$ and $b(r) = \frac{b}{r^\beta}$, where the constants (a, b, β) respectively take the values $(0.01, 0.2, 0)$, $(0.2, 0.01, 0)$, and $(0.2, 0.01, 1.1)$.

Similarly, $\bar{\nu}_1(r) \leq J_{\max}(c(r + d_{\max})^\gamma, d_{\max})(d - 1)b(r)^2$. To bound $\bar{\nu}_2$ and ν_2 , we simply observe that, for all $k > 0$, $\mathbb{E}_r[D_{\text{rad}}^2 \mathbf{1}_{D_{\text{rad}} > 0}] \leq \mathbb{E}_r[G(k, D_{\text{tot}}, \Phi)^2] \leq \mathbb{E}_r[D_{\text{tot}}^2]$ and hence, by symmetry, $\frac{1}{2}a^2 \leq \nu_2(r) \leq \bar{\nu}_2(r) \leq a^2 + (d - 1)b(r)^2$. We discuss the cases $\gamma = 0$ and $\gamma > 0$ separately.

Suppose that $\gamma = 0$ and that, for sufficiently large r , $b(r) = b/r^\beta$ for some constant b . Then the elliptic Markov chain is recurrent if $\beta > \frac{1}{2}$ and transient if $\beta < \frac{1}{2}$. If $\beta = \frac{1}{2}$, then the chain is transient provided $2J_{\min}(c, d_{\max})(d - 1)b^2 > a^2$, and recurrent provided $2J_{\max}(c, d_{\max})(d - 1)b^2 < \frac{1}{2}a^2$. Now suppose that $\gamma > 0$. Note that, for fixed d , $J_{\min}(d, k) \rightarrow 1/2d$ as $k \rightarrow \infty$, and $J_{\max}(d, k) \leq e^{2kd}$ for all sufficiently large k , and so $J_{\max}(d, (r + d)^\gamma) \leq e^{3dr^\gamma}$ for all sufficiently large r . Accordingly, we obtain transience if $(1/d_{\max})r(d - 1)b^2(r) > a^2$ and recurrence if $2re^{3d_{\max}r^\gamma}(d - 1)b^2(r) < \frac{1}{2}a^2$ for all sufficiently large r .

Tighter estimates of ν_i and $\bar{\nu}_i$ would give sharper criteria than those stated above, but we do not pursue this here as our main intention is to contrast the Euclidean and hyperbolic cases. Figure 1 shows numerical simulations of this example in the hyperbolic plane ($c = 1$ and $\gamma = 0$). Only the third simulation, where the transverse component decays to zero, shows recurrence. If analogues of these chains were constructed in Euclidean space, both the second and the third would be recurrent, since results in [8] show that, even if we take a and b to be constant, we can still obtain recurrence provided $2a > b$.

A result in [21] shows that if $k_{\min}(r) \geq Cr^{2+\varepsilon}$ for constants $C, \varepsilon > 0$ then M is stochastically incomplete, and so the discussion above gives a recurrent chain on such a manifold, as promised. At an intuitive level, we might say that the chain resembles Brownian motion on the manifold *locally* (for example, it has zero drift), but has very different global properties.

Finally, we note that in the previous example, for a constant-curvature manifold (i.e. $\gamma = 0$), we obtain recurrence provided $\sup_{x \in S(r)} Q(x)$ decays faster than $O(1/r)$. We stress that this is not sufficient in general, as exemplified below.

Proposition 6. *There is a zero-drift O-transient chain in the hyperbolic plane such that, for every $\delta \in (0, \frac{1}{2})$, $\lim_{r \rightarrow \infty} \exp(r^\delta) \sup_{x \in S(r)} Q(x) = 0$.*

Proof. We give an example of such a chain. Take the probability density of D_{tot} , conditional on the chain being at $x \in M$, to be the same for every x and given by $f_{\text{tot}}(y | x) = \frac{m-1}{y^m}$, $1 \leq y < \infty$, where m is a constant; it is necessary to choose $m > 3$ in order for Assumption 2 to hold. For some function $\lambda(r)$ (to be chosen later), let

$$\varepsilon(y) = \frac{1 - \cosh(y) + \sinh(y)}{\sinh(y)} \cdot \mathbf{1}_{y \geq \lambda(r)}$$

and, conditional on $D_{\text{tot}} = y$, let Φ be distributed as

$$\Phi(\cdot | D_{\text{tot}} = y) = \begin{cases} 1 & \text{with probability } \alpha(y), \\ -1 + \varepsilon(y) & \text{with probability } 1 - \alpha(y), \end{cases}$$

where

$$\alpha(y) = \frac{1 - \varepsilon(y)}{2 - \varepsilon(y)} = \begin{cases} \frac{1 - \cosh(y) + \sinh(y)}{2} & \text{if } y \geq \lambda(r), \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

Notice that ε and α depend on the point $x \in M$ via its distance from O , although for brevity we omit this from our notation. We can check that $0 \leq \alpha(y) \leq 1$ for all $y \geq 1$, so that this definition makes sense. The choice of α ensures that $\mathbb{E}[\Phi | D_{\text{tot}} = y] = 0$ for all y , and hence that $\mathbb{E}_x[D_{\text{rad}}] = \mathbb{E}_x[\Phi D_{\text{tot}}] = 0$ for all $x \in M$. The choice of ε is made to simplify some of the forthcoming expectation calculations. Having specified the distributions of D_{rad} and D_{tot} , it is straightforward to choose the transverse components to give a zero-drift chain. We compute

$$\begin{aligned} \mathbb{E}_x[D_{\text{tot}}^2(1 - \Phi^2) | D_{\text{tot}} = y] &= y^2 \mathbb{E}_x[1 - \Phi^2 | D_{\text{tot}} = y] \\ &= y^2(0 + (1 - \alpha)(1 - (-1 + \varepsilon)^2)) = y^2 \varepsilon. \end{aligned}$$

From now on, assume that $\lambda(r) \geq 1$ for all r . Then

$$\begin{aligned} \mathbb{E}_x[D_{\text{tot}}^2(1 - \Phi^2)] &= \int_{y=1}^\infty f(y) \mathbb{E}_x[D_{\text{tot}}^2(1 - \Phi^2) | D_{\text{tot}} = y] dy \\ &= \int_{y=\lambda}^\infty \frac{(m-1)}{\sinh(y)y^{m-2}} (1 - \cosh(y) + \sinh(y)) dy \\ &\leq \int_{y=\lambda}^\infty \frac{2m}{e^y y^{m-2}} dy \\ &\leq \int_{y=\lambda}^\infty 2me^{-y} dy = 2me^{-\lambda}. \end{aligned} \tag{24}$$

On the other hand, we find that, if $y \geq \lambda(r)$, $\mathbb{E}_x[G(1, D_{\text{tot}}, \Phi) | D_{\text{tot}} = y] = \alpha \log(\cosh y + \sinh y) + (1 - \alpha) \log(1) = y\alpha(y)$, whereas if $y < \lambda(r)$ then $\mathbb{E}[G(1, D_{\text{tot}}, \Phi) | D_{\text{tot}} = y] = 0$. So,

$$\begin{aligned} \mathbb{E}_x[G(1, D_{\text{tot}}, \Phi)] &= \int_{y=\lambda}^\infty y\alpha(y)f(y) dy \\ &= \int_{y=\lambda}^\infty \frac{(m-1)(1 + \sinh(y) - \cosh(y))}{2y^{m-1}} dy \\ &\geq \int_{y=\lambda}^\infty \frac{m-1}{4y^{m-1}} dy = \frac{m-1}{4(m-2)} \frac{1}{\lambda^{m-2}}. \end{aligned} \tag{25}$$

Choose $\lambda(r) = r^{1/(m-1)}$. Then (24) tells us that $\sup_{x \in S(r)} Q(x)$ has the required rate of decay, and (25) tells us that (2) holds and so the chain is transient. \square

7. Future work

In this section we briefly outline some further questions that might be considered. We might allow the curvature of the manifold M to be asymptotically zero (but always negative). Since recurrence and transience typically only depends on what happens very far from the origin, it is a priori not obvious whether chains will typically behave in a hyperbolic manner, in a Euclidean manner, or somewhere in between. There has been recent interest in random processes on manifolds whose metric (and so curvature) changes in time [5, 24]. The latter article contains geometric conditions on the evolution of M for Brownian motion on M to be stochastically complete, and it would be interesting to see what effect (if any) these conditions have on the range of behaviours observed in discrete chains defined on M .

We have considered here only the radial process that measures distance from the origin, but for a full understanding we should also consider the angular process. As mentioned in the introduction, there is a general theme that transient processes on hyperbolic manifolds converge to a limiting angle. A concrete question is whether a zero-drift transient Markov chain in hyperbolic space, under suitable assumptions on the moments of its increments, *must* converge to a limiting angle. In cases where it does, it is natural to ask for a characterisation of the law of that angle.

In Euclidean space, the radial and angular processes of the scaling limits of a class of Markov chains similar to those found in this paper have been considered in [9]. The authors presented a stochastic differential equation satisfied by the limit, and described the behaviour of both the radial and angular components of the limit in detail. The notion of diffusive scaling generalises easily to the manifold setting [13], and so we expect that in our case we would obtain manifold-valued diffusions with similar properties to those in [9]. In view of the qualitative differences between the Euclidean and hyperbolic cases, it would be of interest to compare the limits obtained for hyperbolic and Euclidean manifolds.

Funding information

This work was supported by the Engineering and Physical Sciences Research Council [EP/L015234/1], The EPSRC Centre for Doctoral Training in Geometry and Number Theory (The London School of Geometry and Number Theory), University College London. In addition, the authors thank King's College London for its support.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

- [1] ANDERSON, J. W. (2005). *Hyperbolic Geometry*, 2nd edn. Springer, London.
- [2] ARNAUDON, M., BARBARESCO, F. AND YANG, L. (2011). Medians and means in Riemannian geometry: Existence, uniqueness and computation. In *Matrix Information Geometry*, eds F. Nielsen and R. Bhatia. Springer, Berlin, pp. 169–197.
- [3] CAMMAROTA, V. AND ORSINGER, E. (2008). Travelling randomly on the Poincaré half-plane with a Pythagorean compass. *J. Statist. Phys.* **130**, 455–482.
- [4] CARMO, M. P. DO (1992). *Riemannian Geometry*, 2nd edn. Birkhäuser, Boston, MA.

- [5] COULIBALY-PASQUIER, K. A. (2011). Brownian motion with respect to time-changing Riemannian metrics, applications to Ricci flow. *Ann. Inst. Henri Poincaré Prob. Statist.* **47**, 515–538.
- [6] DENISOV, D., KORSHUNOV, D. AND WACHTEL, V. (2016). At the edge of criticality: Markov chains with asymptotically zero drift. Preprint, [arXiv:1612.01592](https://arxiv.org/abs/1612.01592).
- [7] ÉMERY, M. AND MOKOBODZKI, G. (1991). Sur le barycentre d'une probabilité dans une variété. *Séminaire de probabilités de Strasbourg* **25**, 220–233.
- [8] GEORGIU, N., MENSHIKOV, M. V., MIJATOVIĆ, A. AND WADE, A. R. (2016). Anomalous recurrence properties of many-dimensional zero-drift random walks. *Adv. Appl. Prob.* **48**, 99–118.
- [9] GEORGIU, N., MIJATOVIĆ, A. AND WADE, A. R. (2019). Invariance principle for non-homogeneous random walks. *Electron. J. Prob.* **24**, 1–38.
- [10] GRIGOR'YAN, A. (1999). Analytic and geometric background of recurrence and non-explosion of the Brownian motion on Riemannian manifolds. *Bull. Amer. Math. Soc.* **36**, 135–249.
- [11] HSU, E. P. (2002). *Stochastic Analysis on Manifolds*. American Mathematical Society, Providence, RI.
- [12] ICHIHARA, K. (1982). Curvature, geodesics and the Brownian motion on a Riemannian manifold (I): Recurrence properties. *Nagoya Math. J.* **87**, 101–114.
- [13] JØRGENSEN, E. (1975). The central limit problem for geodesic random walks. *Z. Wahrscheinlichkeitsth.* **32**, 1–64.
- [14] JOST, J. (2012). *Nonpositive Curvature: Geometric and Analytic Aspects*. Birkhäuser, Basel.
- [15] KAKUTANI, S. (1944). On Brownian motions in n -space. *Proc. Imperial Acad.* **20**, 648–652.
- [16] KARLSSON, A. (2004). Linear rate of escape and convergence in direction. In *Random Walks and Geometry*, ed. V. Kaimanovich. De Gruyter, Berlin, pp. 459–472.
- [17] KENDALL, W. S. (1984). Brownian motion on a surface of negative curvature. *Séminaire de probabilités de Strasbourg* **18**, 70–76.
- [18] KRAAIJ, R. C., REDIG, F. AND VERSENDAAL, R. (2019). Classical large deviation theorems on complete Riemannian manifolds. *Stoch. Process. Appl.* **129**, 4294–4334.
- [19] LAMPERTI, J. (1960). Criteria for the recurrence or transience of a stochastic process (I). *J. Math. Anal. Appl.* **1**, 314–330.
- [20] LEE, J. M. (1997). *Riemannian Manifolds: An Introduction to Curvature*. Springer, New York.
- [21] LENZ, D., SOBIECZKY, F. AND WOESS, W. (2011). *Random Walks, Boundaries and Spectra*. Springer, Basel.
- [22] MENSHIKOV, M., POPOV, S. AND WADE, A. (2016). *Non-homogeneous Random Walks: Lyapunov Function Methods for Near-Critical Stochastic Systems*. Cambridge University Press.
- [23] MENSHIKOV, M. V. AND WADE, A. R. (2010). Rate of escape and central limit theorem for the supercritical Lamperti problem. *Stoch. Process. Appl.* **120**, 2078–2099.
- [24] PAENG, S.-H. (2011). Brownian motion on manifolds with time-dependent metrics and stochastic completeness. *J. Geom. Phys.* **61**, 940–946.
- [25] PERES, Y., POPOV, S. AND SOUSI, P. (2013). On recurrence and transience of self-interacting random walks. *Bull. Brazil. Math. Soc. New Ser.* **44**, 841–867.
- [26] PROHASKA, R., SERT, C. AND SHI, R. (2021). Expanding measures: Random walks and rigidity on homogeneous spaces. Preprint, [arXiv:2104.09546](https://arxiv.org/abs/2104.09546).
- [27] SHIOZAWA, Y. (2017). Escape rate of the Brownian motions on hyperbolic spaces. *Proc. Japan Acad. Ser. A* **93**, 27–29.
- [28] STURM, K.-T. (2002). Nonlinear martingale theory for processes with values in metric spaces of nonpositive curvature. *Ann. Prob.* **30**, 1195–1222.
- [29] SULLIVAN, D. (1983). The Dirichlet problem at infinity for a negatively curved manifold. *J. Differential Geom.* **18**, 723–732.
- [30] WILLIAMS, D. (1991). *Probability with Martingales*. Cambridge University Press.