DOI: 10.1214/EJP.v12-408

Electronic Journal of Probability

Vol. 12 (2007), Paper no. 20, pages 591–612.

Journal URL
http://www.math.washington.edu/~ejpecp/

On the range of the simple random walk bridge on groups

Itai Benjamini
The Weizmann Institute of Science
76100 Rehovot
Israel
itai.benjamini@gmail.com

Roey Izkovsky
The Weizmann Institute of Science
76100 Rehovot
Israel
roey.izkovsky@gmail.com

Harry Kesten
Department of Mathematics
Malott Hall
Cornell University
Ithaca NY 14853, USA
kesten@math.cornell.edu

Abstract

Let $\mathcal{G}$ be a vertex transitive graph. A study of the range of simple random walk on $\mathcal{G}$ and of its bridge is proposed. While it is expected that on a graph of polynomial growth the sizes of the range of the unrestricted random walk and of its bridge are the same in first order, this is not the case on some larger graphs such as regular trees. Of particular interest is the case when $\mathcal{G}$ is the Cayley graph of a group G. In this case we even study the range of a general symmetric (not necessarily simple) random walk on G. We hope that the few examples for which we calculate the first order behavior of the range here will help to discover some relation between the group structure and the behavior of the range. Further problems regarding bridges are presented.

Key words: range of random walk; range of a bridge.

AMS 2000 Subject Classification: Primary 60K35.

Submitted to EJP on October 30 2006, final version accepted April 30 2007.

1 Introduction.

A simple random walk bridge of length n on a graph, is a simple random walk (SRW) conditioned to return to the starting point of the walk at time n. In this note we initiate a study of bridges on vertex transitive graphs, concentrating mainly on the range of a bridge. There is a considerable literature (see for instance (7), (28), (29), (15), (5), (6), (13)) on the range of a random walk on $\mathbb{Z}^d$ and on more general graphs. The first result in this area seems to be the following strong law of large numbers from (7), (28), Theorem 4.1 : Let $\{S_n\}_{n\geq 0}$ be a random walk on $\mathbb{Z}$ and let $R_n := \big|\{S_0, S_1, \dots, S_{n-1}\}\big|$ be its range at time n ($|A|$ denotes the cardinality of the set A). Then

$$\lim_{n\to\infty} \frac{1}{n} R_n \to 1 - F \text{ a.s.,} \tag{1.1}$$

where

$$F := P\{S_n = S_0 \text{ for some } n \geq 1\}. \tag{1.2}$$

This result is for an unrestricted random walk, that is, for $S_n = \sum_{i=1}^n X_i$ with the X_i i.i.d. $\mathbb{Z}$-valued random variables. It was extended in (29), (5) to the case when the $\{X_i\}$ form a stationary ergodic sequence. The proof is a simple application of Kingman's subadditive ergodic theorem. It can even be extended to a simple random walk on a vertex transitive graph (see below for a definition). In this paper we are interested in comparing the limit in (1.1), (1.2) with the limit of $(1/n)R_n$ when $\{S_0, \dots, S_n\}$ is conditioned on the event $\mathcal{E}_n := \{S_n = S_0\}$. In this conditioned case, which has the condition varying with n, we can only speak of the limit in probability, since an almost sure limit is meaningless. In a number of examples we shall calculate this limit in probability of $(1/n)R_n$ and see that it equals $1 - F$ in some cases and differs from $1 - F$ in other cases. "Usually" the limit of $(1/n)R_n$ under the condition $\mathcal{E}_n$ is less than or equal to the limit for unrestricted random walk. The idea is that conditioning on $\mathcal{E}_n$ will pull in S_i closer to its starting point than in the unconditioned case, and that this may diminish the range. We shall be particularly interested in the case when $\mathcal{G}$ is the Cayley graph of a finitely generated, infinite group. One would hope that in this case the values of the different limits for $(1/n)R_n$ give some information about the size or structure of the group. Even though it is unclear to what extent such group properties influence R_n, it is likely that the volume growth of the group plays a role (see also Open Problem 1 later in this section). It will also be apparent from our calculations that the behavior of the return probability $P\{S_n = S_0\} = P\{\mathcal{E}_n\}$ is significant.

Various other papers have discussed bridges of random walks on graphs and in particular Cayley graphs. (22) and (2) prove invariance principles for such bridges. (8) discusses the graph distance between the starting point of a bridge and its "midpoint" (to be more specific, if the bridge returns to its starting point at time $2n$, then by its midpoint we mean the position of the bridge at time n); see also the disussion preceding and following (1.27) below). (32) studies still other aspects of bridges of random walks on Cayley graphs and their relation to group structure. In particular, this reference considers the expected value of the so-called Dehn's function of a bridge. We shall mention some further aspects of the range and bridges, as well as some open problems towards the end of this introduction. In fact, some of those remarks served as motivations for the present study.

Here is a formal description of our set up. A countable graph $\mathcal{G}$ is *vertex transitive* if for any two of its vertices v' and v'', there is a graph automorphism $\Phi(\cdot) = \Phi(\cdot; v', v'')$ which maps v' to v''. Throughout we let $\mathcal{G}$ be a countably infinite, connected vertex transitive graph, all of whose vertices have degree $D < \infty$ and let e be a specific (but arbitrary) vertex of $\mathcal{G}$. *Simple random walk* on $\mathcal{G}$ is the Markov chain $\{S_n\}_{n\geq 0}$ which moves from a vertex v to any one of the neighbors of v with probability $1/D$. More formally, its transition probabilities are $P\{S_{n+1} = w|S_n = v\} = 1/D$ if w is a neighbor of v, and 0 otherwise. Unless stated otherwise, we assume that $S_0 = e$. Of particular interest is the case when $\mathcal{G}$ is the *Cayley graph* of a finitely generated infinite group G. Let G be generated by the finite set $\mathcal{S} = \{g_1, \cdots, g_s, g_1^{-1}, \dots, g_s^{-1}\}$ of its elements and their inverses. We can then take for $\mathcal{G}$ the graph whose vertices are the elements of G and with an edge between v' and v'' if and only if $v'' = v'g_i$ or $v'' = v'g_i^{-1}$ for some $1 \leq i \leq s$. This graph $\mathcal{G}$ depends on $\mathcal{S}$ and it will be denoted by $(G, \mathcal{S})$. $\mathcal{G}$ is called the Cayley graph of G corresponding to the generating set $\mathcal{S}$.

If $\alpha_i \geq 0, \sum_{i=1}^{2s} \alpha_i = 1$, we can define a random walk $\{S_n\}_{n\geq 0}$ as follows: Let $X, X_1, X_2, \dots$ be i.i.d. G-valued random variables with the distribution

$$P\{X = g_i\} = \alpha_i, \quad P\{X = g_i^{-1}\} = \alpha_{s+i}, \quad 1 \leq i \leq s. \tag{1.3}$$

Take e to be the identity element of G and set $S_n = X_1X_2\cdots X_n$. This so-called right random walk on G has transition probabilities

$$P\{S_{n+1} = w\big|S_n = v\} = \sum_{i:\, vg_i = w} \alpha_i + \sum_{i:\, vg_i^{-1} = w} \alpha_{s+i}. \tag{1.4}$$

We shall restrict ourselves here to the symmetric case in which

$$\alpha_i = \alpha_{s+i} \text{ or } P\{X = g_i\} = P\{X = g_i^{-1}\}. \tag{1.5}$$

We shall further assume that

$$\alpha_i > 0 \text{ for all } i \text{ and } \mathcal{S} \text{ generates } G. \tag{1.6}$$

(Note that this condition is harmless. If it does not hold from the start we can simply replace G by the group generated by the g_i and g_i^{-1} with $\alpha_i > 0$.)

Throughout we use the following notation (this does not require $\mathcal{G}$ to be a Cayley graph): $\mathcal{E}_n = \{S_n = e\}$,

$$u_n = P\{S_n = e\} = P\{\mathcal{E}_n\},$$

$$f_n = P\{S_k \neq e, 1 \leq k \leq n-1, S_n = e\},\ F = \sum_{n=1}^{\infty} f_n. \tag{1.7}$$

f_n is the probability that $S_.$ returns to e for the first time at time n, and F is the probability that $S_.$ ever returns to e. Finally, $R_n = \big|\{S_0, S_1, \dots S_{n-1}\}\big|$.

A minor nuisance is possible periodicity of the random walk. The *period* is defined as

$$p = \text{g.c.d.}\{n : u_n > 0\}. \tag{1.8}$$

Since the random walk can move from a vertex v to a neighbor w at one step and then go back in the next step from w to v with positive probability, we always have $u_2 > 0$. Thus the period is either 1 or 2. In the latter case we have by definition $P\{\mathcal{E}_n\} = u_n = 0$ for all odd n. In this case it makes little sense to talk about conditioning on the occurrence of $\mathcal{E}_n$ for odd n. *If the period is 2 all statements which involve conditioning on $\mathcal{E}_n$ shall be restricted to even n.*

Our first result states that under the mild condition that u_n does not tend to 0 exponentially fast (see (1.9)), R_n conditioned on $\mathcal{E}_n$, is in some sense no bigger than R_n without the conditioning. In the second theorem we give sufficient conditions for the limit in probability of $(1/pn)R_{pn}$, conditioned on $\mathcal{E}_{pn}$, to equal $1 - F$, which is the same as the almost sure limit of $(1/n)R_n$ without any conditioning (recall (1.1)). In particular, Theorems 1 and 2 and the comments in Example (i) and Open problem 1 below show that for a random walk on a Cayley graph of a group G of polynomial growth which satisfy (1.5) and (1.6) it is the case that $(1/n)R_n$ conditioned on $\mathcal{E}_n$ converges in probability to $1 - F$. If the group G is merely amenable, then $\lim_{n\to\infty}(1/n)R_n \le 1 - F$. The last theorem gives another set of sufficient conditions for the existence of the limit in probability of $(1/pn)R_{pn}$, conditioned on $\mathcal{E}_{pn}$. However, under the conditions of Theorem 3 this limit will often differ from $1 - F$. Examples of random walks satisfying the conditions of Theorems 2 and 3 are given after the theorems.

Theorem 1. *Assume that $\{S_n\}$ is simple random walk on a vertex transitive graph $\mathcal{G}$ or any random walk on a Cayley graph for which* (1.5) *and* (1.6) *hold. Assume further that*

$$\limsup_{n\to\infty} [u_{2n}]^{1/n} = 1. \tag{1.9}$$

Then for all $\varepsilon > 0$

$$\lim_{n\to\infty, p|n} P\{\frac{1}{n}R_n > 1 - F + \varepsilon \big| \mathcal{E}_n\} = 0. \tag{1.10}$$

In particular, if $\{S_n\}$ is recurrent, then $(1/n)R_n$ conditioned on $\mathcal{E}_n$ tends to 0 in probability as $n \to \infty$ through multiples of the period p.

Theorem 2. *Assume that $\{S_n\}$ is a random walk on an infinite Cayley graph for which* (1.5) *and* (1.6) *hold. Assume further that*

$$\limsup_{n\to\infty} \frac{u_{2n}}{u_{4n}} < \infty. \tag{1.11}$$

Then for all $\varepsilon > 0$

$$\lim_{n\to\infty, p|n} P\{\big|\frac{1}{n}R_n - 1 + F\big| > \varepsilon \big| \mathcal{E}_n\} = 0. \tag{1.12}$$

(1.12) *is also valid if there exist two functions $g, h \ge 0$ on $\mathbb{Z}_+$ which satisfy*

$$g(n) \text{ is non-decreasing and tends to } \infty, \tag{1.13}$$

$$\lim_{n\to\infty} \frac{1}{n} g(n) = 0, \tag{1.14}$$

and

$$\lim_{n\to\infty} nh\big(\lfloor \frac{n}{g(n/2)} \rfloor\big) = 0, \tag{1.15}$$

and are such that for all large n

$$e^{-g(n)} \le u_{2n} \le h(n). \tag{1.16}$$

Note that (1.11) is stronger than (1.9) (see (2.26) and the lines preceding it). Note also that we do not require (1.11) in case when (1.13)-(1.16) hold.

Examples.
(i) Let $\{S_n\}$ be a random walk on a Cayley graph $(G, \mathcal{S})$ for which (1.5) and (1.6) hold. If G has polynomial growth, then (1.12) holds. To specialize even further, (1.12) holds for simple random walk on $\mathbb{Z}^d$. To show this we apply Theorem 5.1 of (14). This tells us that if G has polynomial growth of order D, then

$$u_{2n} \asymp n^{-D/2}, \tag{1.17}$$

where $a(n) \asymp b(n)$ for positive $a(\cdot), b(\cdot)$ means that there exist constants $0 < C_1 \leq C_2 < \infty$ such that $C_1 a(n) \leq b(n) \leq C_2 b(n)$ for large n. (1.17) trivially implies (1.11) and hence (1.12) (by Theorem 2). We point out that for random walks on a Cayley graph $(G, \mathcal{S})$ which satisfy (1.5) and (1.6), (1.11) is actually equivalent to polynomial growth of G, or more precisely, polynomial growth of the volume function

$$\begin{aligned} \mathcal{V}(n) = \mathcal{V}(n; G, \mathcal{S}) :=& \text{ number of elements of } G \text{ which can be written} \\ & \text{as } h_1 \cdot h_2 \cdots h_k \text{ with } k \leq n \text{ and each } h_i \in \mathcal{S} \text{ or } h_i^{-1} \in \mathcal{S}. \end{aligned} \tag{1.18}$$

(see Lemma 4 in the next section for a proof).

(ii) As we saw at the end of the preceding example, Theorem 2 deals with random walks on Cayley graphs $(G, \mathcal{S})$ in which G has polynomial growth. As we shall see, Theorem 3 deals with some cases in which G has exponential growth. It is therefore of interest to also look at groups of so-called intermediate groups, as constructed by Grigorchuk in (12). These are finitely generated groups for which there exist constants $0 < \alpha \leq \beta < 1$ and constants $0 < C_3, C_4 < \infty$ such that

$$C_3 e^{n^\alpha} \leq \mathcal{V}(n) \leq C_4 e^{n^\beta}, \quad n \geq 1. \tag{1.19}$$

A random walk on a Cayley graph $(G, \mathcal{S})$ for such a group G and satisfying (1.5) and (1.6) will have

$$u_{2n} \leq C_5 \exp[-C_6 n^{\alpha/(\alpha+2)}],$$

by virtue of Theorem 4.1 in (14). Moreover, since S_n is always a product of at most n elements of $\mathcal{S}$ or inverses of such factors, it holds for some $v_n \in G$ that $P\{S_n = v_n\} \geq 1/\mathcal{V}(n)$ and consequently

$$u_{2n} \geq P\{S_n = v_n\} P\{S_n = v_n^{-1}\} \geq [\mathcal{V}(n)]^{-2} \geq C_4^{-2} \exp[-2n^\beta].$$

Thus Theorem 2 applies with the choices $g(n) = 2n^\beta + 2\log C_4$ and $h(n) = C_5 \exp[-C_6 n^{\alpha/(\alpha+2)}]$. Accordingly, (1.12) holds for such random walks.

(iii) Let G be a wreath product $K \wr \mathbb{Z}^d$ with K a finitely generated group of polynomial growth and $\{S_n\}$ a random walk on a Cayley graph $(G, \mathcal{S})$ for which (1.5) and (1.6) hold. A specific case is the traditional lamplighter group $\mathbb{Z}_2 \wr \mathbb{Z}$ (see (24) for a definition of a wreath product). Then (1.12) holds. Indeed, Theorem 3.11 in (24) shows that there exist constants $0 < C_i < \infty$ such that

$$C_7 \exp[-C_8 n^{1/3} (\log n)^{2/3}] \leq u_{2n} \leq C_9 \exp[-C_{10} n^{1/3} (\log n)^{2/3}]. \tag{1.20}$$

Thus, (1.13)-(1.16) hold with $g(n) = C_8 n^{1/3} (\log n)^{2/3} - \log C_7$ and $h(n) = C_9 \exp[-C_{10} n^{1/3} (\log n)^{2/3}]$. Note that this argument also works if $G = \mathbb{Z} \wr \mathbb{Z}$ or for $G = K \wr Z^d$ with K finite. In the latter case we have to use Theorem 3.5 in (24) instead of Theorem 3.11. The Remark on p. 968 of (24) leads to many more examples to which Theorem 2 applies.

The relation (1.12) is no longer true for simple random walk on a regular tree (which includes the case of random walk on the Cayley graph of a free group) as the following theorem shows.

Theorem 3. *Assume that $\{S_n\}$ is simple random walk on a vertex transitive graph $\mathcal{G}$ or a random walk on a Cayley graph for which* (1.5) *and* (1.6) *hold. Let ρ be the radius of convergence of the power series $U(z) := \sum_{n=0}^{\infty} u_n z^n$ and let $F(z) = \sum_{n=1}^{\infty} f_n z^n$. Then*

$$1 \le \rho < \infty, \ \lim_{n\to\infty} [u_{pn}]^{1/(pn)} = \frac{1}{\rho} \text{ and } F(\rho) \le 1. \tag{1.21}$$

If

$$\text{for all } 0 < \eta < 1, \ \limsup_{n\to\infty, p|n} \ \sup_{1 \le r < (1-\eta)n, p|r} \frac{u_{n-r}}{\rho^r u_n} < \infty, \tag{1.22}$$

and

$$\text{for each fixed } r \text{ with } p|r, \ \lim_{n\to\infty, p|n} \frac{u_{n-r}}{\rho^r u_n} = 1, \tag{1.23}$$

then for all $\varepsilon > 0$

$$\lim_{n\to\infty, p|n} P\{\big|\frac{1}{n}R_n - \big(1 - F(\rho)\big)\big| > \varepsilon \big| \mathcal{E}_n\} = 0. \tag{1.24}$$

Remark. A referee has pointed out to us that (1.23) is automatically satisfied if $\{S_n\}$ is a symmetric random walk on a Cayley graph of period $p = 2$; see (the proof of) Lemma 10.1 in (31)).

Examples. Let $\mathcal{G}_b$ be a regular b-ary tree (in which each vertex has degree $b + 1$). Then, if $\{S_n\}$ is simple random walk on $\mathcal{G}_b$, it holds

$$u_{2n} \sim C_{11} n^{-3/2} \rho^{-2n} \text{ as } n \to \infty \tag{1.25}$$

for some constant $C_{11} > 0$ (see (10) or (20)). We are grateful to a referee who pointed out these two references to us. Alternatively one can apply (19), (26) to the random walk $\{S_{2n}\}_{n\ge 0}$. Actually C_{11} can be explicitly given, but its precise value has no importance for us. Clearly (1.25) implies (1.22) and (1.23) and hence (1.24) with $p = 2$ (the period for $\{S_n\}$).

We shall show after the proof of Theorem 3 that

$$F = \frac{1}{b}, \ \rho = \frac{b+1}{2\sqrt{b}} \text{ and } F(\rho) = \frac{b+1}{2b}. \tag{1.26}$$

Thus if $b \ge 2$, then in this example, the limit in probability of $(1/2n)R_{2n}$ conditioned on $\mathcal{E}_{2n}$ is strictly less than the almost sure limit of $(1/n)R_n$ for the unconditioned walk. (These limits are $(b-1)/(2b)$ and $(b-1)/b$, respectively.) Similar results hold for random walks on various free products (see (3)). For instance we again have (1.25), and hence (1.24), for the simple random walk on the free product of s copies of $\mathbb{Z}^q$ with $s \ge 2, q \ge 1$, if we use the natural set of generators consisting of the coordinate vectors in each factor $\mathbb{Z}^q$.

We end this section with some related remarks on the range and bridges and list a number of open problems.

Note that if $\mathcal{G}$ is a regular graph, that is all degrees are equal, then the distribution of a simple random walk bridge is just uniform measure on all n-step paths which return to the starting point at the n-th step. The question of *sampling* a bridge on a given Cayley graph seems hard in general. We don't even know how to sample a bridge on the lamplighter group. Or even simpler, consider a symmetric random walk on $\mathbb{Z}$ which can take jumps of size 1 or 2. How can one describe a bridge of such a random walk ?

Open problem 1. Find a necessary and sufficient condition for (1.12) for a symmetric random walk on a Cayley graph $(G, \mathcal{S})$. Theorems 1-3 suggest that perhaps amenability of G is such a necessary and sufficient condition (recall that under (1.5) and (1.6) G is amenable if and only if (1.9) holds; see (17) or (16), Section 5, or (23), Theorems 4.19, 4.20 and Problem 4.24).

Open problem 2. Does there exist a transient random walk on a Cayley graph or a transient simple random walk on a vertex transitive graph for which $(1/n)R_n$ conditioned on $\mathcal{E}_n$ tends to 0 in probability ?

Open problem 3. If the limit in probability of $(1/n)R_n$, conditioned on $\mathcal{E}_n$ exists, is it necessarily $\le 1-F$ (even if (1.9) fails) ?

Open problem 4. Does unrestricted random walk drift further away from the starting point than a bridge ? More precisely, let $\mathcal{G}$ be a vertex transitive graph, $\{S_k\}$ a simple random walk starting at a given vertex e, and $\{S_k^b\}$, a simple random walk bridge conditioned to be back at e at time n. Is it true that for every fixed $k < n$, $d(S_k) :=$ the (graph) distance of S_k to e, stochastically dominates $d(S_k^b)$? A related conjecture might be that on any vertex transitive graph the simple random walk bridge is at most diffusive, that is, $\max_{k\le n} d(S_k^b)$ is (stochastically) at most of order $\sqrt{n}$ or perhaps $\sqrt{n}(\log n)^q$ for some q. In a more quantitave way one may ask for limit laws and further bounds on the distribution of $\max_{k\le n} d(S_k^b)$ (see also Section 7 in (8)). Note that on any graph with a Gaussian off diagonal correction, i.e., for which there is a bound $P\{\max_{k\le n} d(S_k^b) = \ell\} \sim C_{12} n^{-q} \exp[-C_{13}\ell^2/n]$ for some positive q and constants C_{12}, C_{13}, the bridge is at least diffusive. Vertex transitivity of $\mathcal{G}$ seems to be crucial as it is not hard to build examples of graphs on which bridges have non-diffusive and more erratic behavior.

A weaker statement might be that on any vertex transitive graph there is a subexponential lower bound of the form $P\{d(S_k^b) \le \ell | \mathcal{E}_n\} \le \exp[-C\ell/n]$ for the conditional probability given $\mathcal{E}_n$, that the bridge is near the starting point at time k.

Another interesting problem in the same direction is to adapt the Varopoulos-Carne subgaussian estimate (see (21), Theorem 12.1) to bridges. That is, to prove that $P(S_k^b = v) \le e^{-d^2(v)/2n}$ or a weaker result of this type. In (18), Proposition 3.3, there is a proof that can give a subgaussian estimate for bridges on certain graphs.

We add two more observations which may be of interest, but do not merit the label "open problem." The first concerns a random walk on the lamplighter group $\mathbb{Z}_2 \wr \mathbb{Z}^d$. A generic element of the group is of the form (σ, y) with $y \in \mathbb{Z}^d$ and σ a function from $\mathbb{Z}^d$ into $\{0, 1\}$. Assume that the random walk starts in $(\sigma_0, \mathbf{0})$, where σ_0 is the zero function and $\mathbf{0}$ is the origin in $\mathbb{Z}^d$. Further, let σ_y be the configuration obtained from σ by changing $\sigma(y)$, the value of σ at the position y, to $\sigma(y) + 1 \mod 2$.

Assume that the random walk moves from (σ, y) to $(\sigma, y \pm e_i)$ or to $(\sigma_y, y \pm e_i)$, $1 \le i \le d$, with probability $1/(4d)$ for each of the possibilities. Here e_i is the i-th coordinate vector in $\mathbb{Z}^d$. Let us define $\sigma_k, y_k, \sigma_k^b, y_k^b$ by $S_k = (\sigma_k, y_k)$ and $S_k^b = (\sigma_k^b, y_k^b)$. Then, for the unrestricted random walk, y_k is simple random walk on $\mathbb{Z}^d$. For the bridge, all possible paths have equal weight. Let $(y_0 = \mathbf{0}, y_1, \dots, y_n = \mathbf{0})$ be a nearest neighbor path on $\mathbb{Z}^d$ from $\mathbf{0}$ to $\mathbf{0}$. How many sequences $\{\sigma_k, y_k\}_{0\le k\le n}$ are there which return to $(\sigma_0, \mathbf{0})$ at time n and which project onto the given sequence $\{y_k\}$? Such a sequence must have its first coordinate set to 0 at position y at the last visit to y by $\{y_k\}_{0\le k\le n}$. Moreover, σ_k can change only at the position y_k. If y is not one of the y_k there is no condition on the σ_k at position y at all. If N_n denotes $|\{y_0, y_1, \dots y_{n-1}\}|$ (i.e., the range of the projection on $\mathbb{Z}^d$ of the random walk on $\mathbb{Z}_2 \wr \mathbb{Z}^d$), then one sees from the above that the number of possible $\{\sigma_k, y_k\}_{0\le k\le n}$ which return to $(\sigma_0, \mathbf{0})$ at time n and which project onto the given sequence $\{y_k\}$ equals 2^{n-N_n}. Thus, if p_r denotes the probability that an n-step simple random walk bridge from $\mathbf{0}$ to $\mathbf{0}$ in $\mathbb{Z}^d$ has range r, then the probability that the projection of the random walk bridge on $\mathbb{Z}_2 \wr \mathbb{Z}^d$ has range r equals

$$P\{N_n = r\} = \frac{2^{n-r}p_r}{\sum_s 2^{n-s}p_s} = \frac{2^{-r}p_r}{\sum_s 2^{-s}p_s},$$

(this formula is very similar to equation (3.1) in (24)) and the expectation of the range of the projection is

$$E\{N_n\} = \frac{\sum_{1\le r\le n} r2^{-r}p_r}{\sum_{1\le s\le n} 2^{-s}p_s}. \tag{1.27}$$

We shall next argue that this expectation is only $O(n^{d/(d+2)})$, so that by time n the projection on $\mathbb{Z}^d$ of the bridge of the random walk only travels a distance $O(n^{d/(d+2)})$ from the origin. If $d = 1$ it can be shown that then also the bridge of the random walk on $\mathbb{Z}_2 \wr \mathbb{Z}$ itself only moves distance $O(n^{1/3})$ from the starting point by time n. More precisely, define for the bridge

$$D_n = \max_{0\le k\le n} d(S_k^b).$$

Then $n^{-1/3}D_n, n \ge 1$, is a tight family. Even though this maximal distance D_n for the bridge grows slower than n, the range of the bridge still grows in first order at the same rate as the range of the unrestricted random walk (by example (iii) to Theorem 2). (We note in passing that a similar, but weaker, comment applies if $\{S_n\}$ is simple random walk bridge on the regular b-ary tree $\mathcal{G}_b$. In this case it is known ((2; 22)) that $n^{-1/2}D_n, n \ge 1$, is a tight family. Now the range of the bridge is in first order still linear in n, but the ranges of the bridge and of the unrestricted simple random walk on $\mathcal{G}_b$ differ already in first order; see the example to Theorem 3). We point out that (8) argues that $n^{-1/3}E\{D_n\} \ge n^{-1/3}d(S_{n/2})$ is also bounded away from 0 for n even, $d = 1$ and a random walk which differs from ours only a little in the choice of the distribution of the X_i. Presumably such a lower bound for D_n also holds in our case, but is not needed for the present remark.

Here is our promised estimate for $E\{N_n\}$. It is known (see (6)) that for an unrestricted simple random walk $\{S_k\}$ on $\mathbb{Z}^d$,

$$\begin{aligned} P\{R_n \le (2L+1)^d\} &\ge P\{S_k \in [-L, L]^d, 0 \le k \le n\} \\ &\ge C_{14}\exp[-C_{15}nL^{-2}] \end{aligned}$$

for some constants $0 < C_i < \infty$. It can be shown from this that also

$$P\{S_k \in [-L, L]^d, 0 \le k \le n, S_n = \mathbf{0}\} \ge C_{16}L^{-d}\exp[-C_{15}nL^{-2}].$$

A fortiori

$$P\{S_k \in [-L, L]^d, 0 \le k \le n \big| S_n = \mathbf{0}\} \ge C_{16}L^{-d}\exp[-C_{15}nL^{-2}].$$

By taking $L = n^{1/(d+2)}$ we find that the denominator in (1.27) is for large n at least

$$2^{-(2L+1)^d} C_{16} L^{-d} \exp[-C_{15} n L^{-2}] \geq C_{17} \exp[-C_{18} n^{d/(d+2)}].$$

Consequently the right hand side of (1.27) is at most

$$\begin{aligned} &\frac{\sum_{r \leq K} r 2^{-r} p_r}{\sum_s 2^{-s} p_s} + \frac{n}{C_{17} \exp[-C_{18} n^{d/(d+2)}]} \sum_{K < r \leq n} 2^{-r} p_r \\ &\leq K + \frac{n}{C_{17}} \exp[C_{18} n^{d/(d+2)}] 2^{-K}. \end{aligned}$$

By choosing $K = 2C_{18} n^{d/(d+2)} / \log 2$ we obtain the promised bound $E\{N_n\} = O(n^{d/(d+2)})$.

Our last observation deals with bridges on special finite graphs. Let $\mathcal{G}$ be a D-regular vertex transitive *expander* of size N with girth (smallest cycle) of size $c \log N$. The mixing time for simple random walk on such a graph is $C \log N$, for some constant $C > c$, where mixing time is taken in the strong sense of the maximum relative deviation. That is, we take the mixing time to be the number of steps, k, a simple random walk has to take to make $\sup_{v,w \in \mathcal{G}} \big| P\{S_k = w | S_0 = v\} / \pi(w) - 1 \big|$ smaller than some prescribed number, where $\pi(\cdot)$ is the stationary measure for the random walk (see (27)). Now a bridge of length $n < c \log N$ is just a bridge on a $(D-1)$-regular tree, since there are no cycles of length $< c \log N$. Thus the range of such a bridge is in first order described by Theorem 3. On the other hand, a bridge of length $n > C' \log N$ for large enough C' may be expected to look like an unconstrained simple random walk at least for times in $[c \log N, n - c \log N]$, because of the short mixing time. For larger n, but still order $\log N$, we actually expect the range of such a bridge to be like the range of a n-step unconstrained random walk on a $(D-1)$-regular tree, that is $n(D-2)/(D-1)$ in first order. This suggests that maybe there is a critical C^*, so that a bridge of length $< C^* \log N$ (respecively $> C^* \log N$) looks as in the first case (respectively, second case).

Acknowledgement. We thank Laurent Saloff-Coste for several helpful conversations about the subject of this paper.

2 Proofs.

Proof of Theorem 1. We shall treat the aperiodic case (i.e., the case with $p = 1$) only. The case when the period equals 2 can be treated in the same way. One merely has to restrict the appropriate subscripts to even integers.

Define the random variables

$$Y(k, M) := I\big[S_{k+r} \neq S_k \text{ for } 1 \leq r \leq M\big].$$

By a last exit decomposition we have for any positive integer M

$$\begin{aligned} R_n &= \sum_{k=0}^{n-1} I\big[S_{k+r} \neq S_k \text{ for } 1 \leq r \leq n-k\big] \\ &\leq M + \sum_{k=0}^{n-M} I\big[S_{k+r} \neq S_k \text{ for } 1 \leq r \leq M\big] \\ &= M + \sum_{k=0}^{n-M} Y(k, M) \end{aligned} \tag{2.1}$$

(compare proof of Theorem 4.1 in (28), which uses a first entry decomposition). Now by the Markov property of the random walk $\{S_k\}$ and the transitivity of $\mathcal{G}$,

$$\begin{aligned} &P\{Y(k,M) = 1 \big| S_0, S_1, \ldots, S_k\} \\ &= P\{S_{k+r} \neq S_k \text{ for } 1 \leq r \leq M \big| S_0, S_1, \ldots, S_k\} \\ &= P\{S_r \neq S_0 \text{ for } 1 \leq r \leq M\} = P\{Y(0, M) = 1\}. \end{aligned} \tag{2.2}$$

This relation says that any collection $\{Y(k_i, M)\}$ of these random variables with $|k_i - k_j| > M$ for $i \neq j$ consists of i.i.d. random variables. In fact, each $Y(k, M)$ can take only the values 0 or 1. This will allow us to use exponential bounds for binomial random variables.

First observe that for given $\varepsilon > 0$ we can choose M such that

$$\begin{aligned} E\{R_n\} &\geq \sum_{k=0}^{n-1} P\{S_{k+r} \neq S_k \text{ for all } r \geq 1\} \\ &\geq \sum_{k=0}^{n-1} \big[E\{Y(k,M)\} - P\{\text{first return to } S_k \text{ occurs at time} \\ &\qquad\qquad k + r \text{ for some } r > M\}\big] \\ &= \sum_{k=0}^{n-1} \big[E\{Y(k,M)\} - \sum_{r>M} f_r\big] \geq \sum_{k=0}^{n-1} E\{Y(k,M)\} - \varepsilon n. \end{aligned} \tag{2.3}$$

We now rewrite (2.1) as

$$R_n \leq M + \sum_{a=0}^{M} \sum_{\substack{k \equiv a \bmod (M+1) \\ 0 \leq k \leq n-1}} Y(k, M). \tag{2.4}$$

Moreover, by (2.1) we have for large enough M and $n \geq n_0$ for some $n_0 = n_0(M)$,

$$\begin{aligned} E\{R_n\} &\leq M + \sum_{k=0}^{n-M} E\{Y(k,M)\} \\ &\leq M + nP\{S_r \neq S_0 \text{ for } 1 \leq r \leq M\} \leq n[1 - F + \varepsilon]. \end{aligned} \tag{2.5}$$

Thus, if for given $\varepsilon > 0$, M is chosen so that (2.3) and (2.5) hold, then for all large n

$$\begin{aligned} &P\{R_n \geq n(1-F) + 4\varepsilon n\} \leq P\big\{R_n - E\{R_n\} \geq 3\varepsilon n\big\} \\ &\leq \sum_{a=0}^{M} P\Big\{ \sum_{\substack{k \equiv a \bmod (M+1) \\ 0 \leq k \leq n-1}} \big[Y(k,M) - E\{Y(k,M)\}\big] \geq \frac{\varepsilon}{M+1} n \Big\}. \end{aligned} \tag{2.6}$$

The right hand side here tends to 0 exponentially fast as $n \to \infty$ by standard exponential bounds for large deviations in binomial distributions (e.g. by Bernstein's inequality (4), Exercise 4.3.14). We now remind the reader of (1.9). It is easy to see from the definition of u_k that

$$u_{2k+2\ell} \geq u_{2k} u_{2\ell}. \tag{2.7}$$

From this and the fact that $u_{2k} \geq [u_2]^k > 0$ one obtains that $\lim_{n\to\infty} [u_{2n}]^{1/n}$ exists, and then by (1.9) that this limit must equal 1. This holds regardless of whether $p = 1$ or 2. If $p = 1$, then even $u_m > 0$ for some odd m and $u_n \geq u_{n-m} u_m$ then shows that also $[u_{2k+1}]^{1/(2k+1)} \to 1$ as $k \to \infty$. Consequently,

$$\frac{1}{u_n} P\{\frac{1}{n} R_n > 1 - F + 4\varepsilon\} \to 0 \text{ as } n \to \infty.$$

(in fact, exponentially fast). But then also

$$P\{\frac{1}{n} R_n > 1 - F + 4\varepsilon \big| \mathcal{E}_n\} \to 0.$$

This proves (1.10). If the random walk $\{S_n\}$ is recurrent, then $F = 1$ and (1.10) says that $(1/n)R_n$ conditioned on $\mathcal{E}_n$ tends to 0 in probability. ■

Proof of Theorem 2. We begin this subsection with some lemmas on monotonicity and smoothness of the u_n (as functions of n). Then we give a general sufficient condition in terms of the u_n and f_n for the limit in probability of $(1/n)R_n$ conditioned on $\mathcal{E}_n$ to equal $1 - F$. Finally we show that this sufficient condition holds under the conditions of Theorem 2.

Lemma 1. *Let $\{S_n\}$ be a random walk on a Cayley graph $(G, \mathcal{S})$ and assume that the symmetry property* (1.5) *holds. Then there exists a probability measure μ on $[-1, 1]$ such that*

$$u_n = \int_{[-1,1]} x^n \mu(dx),\ n \geq 0. \tag{2.8}$$

Consequently

$$u_{2n} \text{ is non-increasing in } n, \tag{2.9}$$

$$u_{2n+1} \leq u_{2n}, \tag{2.10}$$

and also $u_{2r} u_{2n-2r}$ is non-increasing in r for $0 \leq r \leq (n-1)/2$, that is,

$$u_{2r} u_{2n-2r} \geq u_{2r+2} u_{2n-2r-2} \text{ for } 0 \leq r \leq (n-1)/2. \tag{2.11}$$

Proof. We write

$$P(v,w) = P\{X = v^{-1}w\} = P\{S_{n+1} = w | S_n = v\}, \quad v, w \in G,$$

for the transition probabilities of the random walk $\{S_n\}$. Then the k-th power of P gives the k-step transition probability. That is

$$P^k(v,w) = P\{S_{k+n} = w | S_n = v\}.$$

P defines a linear operator on $\ell^2(G)$ by means of

$$Pf(v) = \sum_{w \in G} P(v,w) f(w).$$

This linear operator takes $\ell^2(G)$ into itself and is self-adjoint. By the spectral theorem (see (25), Theorems 12.23, 12.24; see also (16), Section 5) there therefore exists a measure μ on the Borel sets of $\mathbb{R}$ such that for $f_0(v) = I[v = e]$, and $\langle \cdot, \cdot \rangle$ the inner product on $\ell^2(\mathcal{G})$,

$$u_n = \langle f_0, P^n f_0 \rangle = \int_{\mathbb{R}} x^n \mu(dx), \; n \geq 0.$$

From $|Pf(v)|^2 \leq \sum_w P(v,w)|f|^2(w) \sum_w P(v,w) = \sum_w P(v,w)|f|^2(w)$ we see that $\|P\| \leq 1$, so that the support of μ must be contained in $[-1,1]$ and (2.8) must hold. (This can also be see directly from $|u_n| \leq 1$ for all n.)

This proves the existence of some measure μ for which (2.8) is satisfied. In fact, the spectral theorem tells us that for A a Borel set of $[-1,1]$, $\mu(A) = \langle f_0, E(A) f_0 \rangle$ for some resolution of the identity $\{E(\cdot)\}$. In particular, $E([-1,1])$ is the identity I on $\ell^2(\mathcal{G})$ so that $\mu([-1,1]) = \langle f_0, I f_0 \rangle = 1$. Thus μ is a probability measure as claimed.

(2.9) and (2.10) are immediate consequences of (2.8). As for (2.11), we have from (2.8) that

$$\begin{aligned}
&u_{2r}u_{2n-2r} - u_{2r+2}u_{2n-2r-2} \\
&= \int \mu(dx) \int \mu(dy) \big[x^{2r}y^{2n-2r} - x^{2r+2}y^{2n-2r-2}\big] \\
&= \frac{1}{2} \int \mu(dx) \int \mu(dy) \big[x^{2r}y^{2n-2r} - x^{2r+2}y^{2n-2r-2} \\
&\qquad\qquad + y^{2r}x^{2n-2r} - y^{2r+2}x^{2n-2r-2}\big] \\
&= \frac{1}{2} \int \mu(dx) \int \mu(dy) x^{2r}y^{2r} \big[y^{2n-4r-2} - x^{2n-4r-2}\big]\big[y^2 - x^2\big] \geq 0,
\end{aligned} \tag{2.12}$$

because $[y^{2n-4r-2} - x^{2n-4r-2}][y^2 - x^2] \geq 0$ for all x, y if $2n - 4r - 2 \geq 0$. ∎

Lemma 2. *Let $\{S_n\}$ be a random walk on a Cayley graph $(G, \mathcal{S})$ which satisfies the symmetry assumption* (1.5). *Assume that*

$$u_{2n} \geq e^{-g(n)} \tag{2.13}$$

for some function $g(\cdot)$ which satisfies (1.13), (1.14). *Then*

$$\frac{n}{n+r}\Big(\frac{r}{n+r}\Big)^{r/n} e^{[-rg(n)/n]} \leq \frac{u_{2n+2r}}{u_{2n}} \leq 1. \tag{2.14}$$

for all $r \geq 1$. *In particular,*

$$\lim_{n\to\infty} \frac{u_{2n+2}}{u_{2n}} = 1. \tag{2.15}$$

Moreover, if $p = 1$,

$$\lim_{n\to\infty} \frac{u_{n+1}}{u_n} = 1. \tag{2.16}$$

Proof. The right hand inequality in (2.14) is part of Lemma 1.

To find a lower bound for u_{2n+2r}/u_{2n} we again appeal to (2.8). This tells us that for any $\gamma \geq 0$,

$$\begin{aligned}
e^{-g(n)} \leq u_{2n} &= \int_{[-1,1]} x^{2n}\mu(dx) \\
&= \int_{|x|\leq \exp[-(\gamma+g(n))/(2n)]} x^{2n}\mu(dx) + \int_{|x|> \exp[-(\gamma+g(n))/(2n)]} x^{2n}\mu(dx) \\
&\leq \exp[-(\gamma+g(n))] + \int_{|x|> \exp[-(\gamma+g(n))/(2n)]} x^{2n}\mu(dx) \\
&\leq e^{-\gamma}u_{2n} + \int_{|x|> \exp[-(\gamma+g(n))/(2n)]} x^{2n}\mu(dx).
\end{aligned}$$

Consequently,

$$[1-e^{-\gamma}]u_{2n} \leq \int_{|x|> \exp[-(\gamma+g(n))/(2n)]} x^{2n}\mu(dx).$$

It follows that

$$\begin{aligned}
u_{2n+2r} &\geq \int_{|x|> \exp[-(\gamma+g(n))/(2n)]} x^{2n+2r}\mu(dx) \\
&\geq \exp\big[-\frac{r}{n}(\gamma+g(n))\big] \int_{|x|> \exp[-(\gamma+g(n))/(2n)]} x^{2n}\mu(dx) \\
&\geq \exp\big[-\frac{r}{n}(\gamma+g(n))\big][1-e^{-\gamma}]u_{2n}.
\end{aligned}$$

The left hand inequality of (2.14) follows by taking $e^{-\gamma} = r/(n+r)$.

The limit relation (2.16) is proven in (11) and in (1). ■

Lemma 3. *Assume that* $\{S_n\}$ *is a random walk on an infinite Cayley graph for which* (1.5), (1.6) *and* (1.9) *hold. If, in addition, for each* $\eta > 0$

$$\lim_{M\to\infty} \limsup_{n\to\infty, p|n} \sum_{\substack{M\leq r\leq(1-\eta)n \\ p|r}} f_r \frac{u_{n-r}}{u_n} = 0, \tag{2.17}$$

then for all $\varepsilon > 0$

$$\lim_{n\to\infty, p|n} P\{\big|\frac{1}{n}R_n - 1 + F\big| > \varepsilon \big| \mathcal{E}_n\} = 0. \tag{2.18}$$

Proof. We claim that it suffices to show

$$\liminf_{n\to\infty, p|n} E\{\frac{1}{n}R_n \big| \mathcal{E}_n\} \geq 1 - F. \tag{2.19}$$

Indeed, we already know from Theorem 1 that (1.10) holds. Together with $(1/n)R_n \le 1$ and (2.19) this implies for $\varepsilon > 0$ and n a large multiple of p that

$$\begin{aligned}1 - F - \varepsilon \le E\{\frac{1}{n}R_n|\mathcal{E}_n\} &\le \big(1 - F - \sqrt{\varepsilon}\big)P\{\frac{1}{n}R_n \le 1 - F - \sqrt{\varepsilon}|\mathcal{E}_n\}\\ &\quad + \big(1 - F + \varepsilon\big)P\{1 - F - \sqrt{\varepsilon} < \frac{1}{n}R_n \le 1 - F + \varepsilon|\mathcal{E}_n\}\\ &\quad + P\{\frac{1}{n}R_n > 1 - F + \varepsilon|\mathcal{E}_n\}\\ &\le \big(1 - F - \sqrt{\varepsilon}\big)P\{\frac{1}{n}R_n \le 1 - F - \sqrt{\varepsilon}|\mathcal{E}_n\}\\ &\quad + \big(1 - F + \varepsilon\big)\big[1 - P\{\frac{1}{n}R_n \le 1 - F - \sqrt{\varepsilon}|\mathcal{E}_n\}\big] + \frac{1}{2}\varepsilon.\end{aligned} \tag{2.20}$$

By simple algebra this is equivalent to

$$P\{\frac{1}{n}R_n \le 1 - F - \sqrt{\varepsilon}|\mathcal{E}_n\} \le \frac{5\varepsilon}{2(\sqrt{\varepsilon} + \varepsilon)}.$$

This justifies our claim that we only have to prove (2.19).

We turn to the proof of (2.19). Again by a last exit decomposition

$$\begin{aligned}R_n \ge R_{\lfloor(1-\eta)n\rfloor} &\ge \sum_{0 \le k < \lfloor(1-\eta)n\rfloor} I[S_{k+r} \ne S_k \text{ for } 1 \le r \le (1-\eta)n - k]\\ &= \sum_{0 \le k < \lfloor(1-\eta)n\rfloor} \Big[1 - \sum_{r=1}^{\lfloor(1-\eta)n\rfloor - k} I[\text{first return to } S_k \text{ occurs at time } k + r]\Big]\end{aligned}$$

Now multiply this inequality by $I_n := I[\mathcal{E}_n]$ and take expectations. This yields

$$\begin{aligned}\frac{1}{n}E\{R_n|\mathcal{E}_n\} = \frac{1}{nu_n}E\{R_nI_n\} &\ge \frac{(1-\eta)n}{n}\\ - \frac{1}{nu_n}\sum_{0 \le k < \lfloor(1-\eta)n\rfloor}\ \sum_{1 \le r \le (1-\eta)n - k}\ \sum_{v \in G} &P\{S_k = v, \text{first return to } v\\ &\text{after } k \text{ is at time } k + r \text{ and } S_n = e\}.\end{aligned} \tag{2.21}$$

The inner sum of the triple sum in the right hand side here equals

$$\begin{aligned}&\sum_{v \in G} P\{S_k = v\}f_rP\{S_n = e|S_{k+r} = v\}\\ &\quad = f_r\sum_{v \in G} P\{S_k = v\}P\{S_{n-k-r} = e|S_0 = v\} = f_ru_{n-r}.\end{aligned} \tag{2.22}$$

We substitute this into (2.21) and use the assumption (2.17) to obtain for any $\eta \in (0,1)$

$$\liminf_{n\to\infty, p|n} \frac{1}{n}E\{R_n|\mathcal{E}_n\} \ge 1 - \eta - \limsup_{M\to\infty}\ \limsup_{n\to\infty, p|n} \sum_{r=1}^{M} f_r\frac{u_{n-r}}{u_n}. \tag{2.23}$$

Note that $f_r \le u_r = 0$ if $p \nmid r$. Thus, if we prove that

$$\lim_{n\to\infty, p|n} \frac{u_{n-r}}{u_n} = 1 \text{ for fixed } r \text{ with } p|r, \tag{2.24}$$

then the desired (2.19), and hence also (2.18) will follow.

We shall now deduce (2.24) from Lemma 2. As we saw in (2.7) and the lines following it, (1.9) implies that

$$\lim_{n\to\infty} [u_{2n}]^{1/n} = 1. \tag{2.25}$$

Define

$$g(n) = \log \frac{1}{u_{2n}}.$$

$g(\cdot)$ is non-decreasing by virtue of (2.9), and its limit as $n \to \infty$ must be ∞. To see this note that the Markov chain $\{S_n\}$ cannot be positive recurrent, because the measure which puts mass 1 at each vertex of $\mathcal{G}$ is an infinite invariant measure if G is infinite (see (9), Theorem in Section XV.7 and Theorem XV.11.1). Thus it must be the case that $u_n \to 0$. It follows that $g(\cdot)$ satisfies (1.13). Moreover, (1.14) for $g(\cdot)$ is implied by (2.25), and (2.13) holds by definition. (2.24) then follows from (2.15) and (2.16). ■

We are finally ready to give the proof of Theorem 2. By virtue of Lemma 3 it suffices to prove that (1.9) and (2.17) hold. Assume first that (1.11) holds. (1.9) then follows easily. Indeed, (1.11) implies that there exists some n_0 and a constant $C_{14} > 0$ such that

$$u_{2^r n_0} \ge C_{14}^{r-1} u_{2n_0}. \tag{2.26}$$

From the existence of $\lim_{n\to\infty}[u_{2n}]^{1/n}$ (see (2.7) and the lines following it) we then see that (1.9) holds. As for (2.17), we have from (1.11) and the monotonicity in q of u_{2q} that for fixed $\eta > 0$ there exists some constant C_{15} depending on η only such that $u_{n-r}/u_n \le C_{15}$ for even n and even $r \le (1-\eta)n$. For the case when $p = 2$ this shows for even n that

$$\sum_{\substack{M \le r \le (1-\eta)n \\ r \text{ even}}} f_r \frac{u_{n-r}}{u_n} \le C_{15} \sum_{r \ge M} f_r.$$

In this case (2.17) is therefore immediate. If $p = 1$ we can use essentially the same argument, since for odd $r \le (1-\eta)n$, $u_r \le u_{r-1}$ (see (2.10)), while for odd n, $u_n \ge u_{L_0} u_{n-L_0}$, where L_0 is a fixed odd integer for which $u_{L_0} > 0$. These simple observations show that in case $p = 1$ we still have u_{n-r}/u_n bounded by some $C_{15}(\eta)$ for all $r \le (1-\eta)n$, and hence also (2.17) holds.

Now we turn to the proof of (1.12) from (1.13)-(1.16). First note that (1.14) and (2.13) imply that (1.9) (or even (2.25)) holds (see a few lines after (2.7)). Thus it is again enough to prove (2.17). To this end, define

$$\tau = 2\left\lfloor \frac{n}{g(n/2)} \right\rfloor$$

and use Lemma 2 to obtain for fixed M, but large even n

$$\begin{aligned}
&\sum_{M\le r\le(1-\eta)n} f_r\frac{u_{n-r}}{u_n} \le \sum_{M\le r\le n-M}\frac{u_r u_{n-r}}{u_n}\\
&\le 2\sum_{M\le r\le n/2}\frac{u_r u_{n-r}}{u_n}\\
&\le 2\sum_{M\le r\le n/2, 2|r}\frac{u_r u_{n-r}}{u_n} + 2\sum_{M-1\le r\le (n-2)/2, 2|r}\frac{u_r u_{n-2-r}}{u_n} \text{ (by (2.10))}\\
&\le 2\sum_{r=M,2|r}^{\tau-1}\frac{u_r u_{n-r}}{u_n} + 2n\frac{u_\tau u_{n-\tau}}{u_n} + 2\sum_{r=M-1,2|r}^{\tau-1}\frac{u_r u_{n-2-r}}{u_n} + 2n\frac{u_\tau u_{n-2-\tau}}{u_n} \text{ (by (2.11))}\\
&\le 4\sum_{r=M-1}^{\infty} u_r\frac{u_{n-2-\tau}}{u_n} + 4n\frac{u_\tau u_{n-2-\tau}}{u_n} \text{ (by (2.9) and (2.10)).}
\end{aligned} \tag{2.27}$$

Next, (2.14) (with $2n$ replaced by $n-2-\tau$ and $2r$ by $\tau+2$) shows that for large even n,

$$\begin{aligned}
&\frac{u_{n-2-\tau}}{u_n}\\
&\le \frac{n}{(n-2-\tau)}\Big(\frac{n}{\tau+2}\Big)^{(\tau+2)/(n-2-\tau)}\exp\big[\frac{\tau+2}{n-2-\tau}g\big((n-2-\tau)/2\big)\big].
\end{aligned} \tag{2.28}$$

But the definition of τ and the monotonicity of $g(\cdot)$ show that $\tau = o(n)$ and

$$\frac{\tau+2}{n-2-\tau}g\big((n-2-\tau)/2\big) \le 2\frac{\tau}{n}g(n/2) \le 4$$

for large n. Therefore, the right hand side of (2.28) is bounded by some constant C_{16}. Substitution of this bound into (2.27) now shows that

$$\sum_{M\le r\le(1-\eta)n} f_r\frac{u_{n-r}}{u_n} \le 4C_{16}\sum_{r=M-1}^{\infty} u_r + 4C_{16}nu_\tau. \tag{2.29}$$

Theorem 1 already proves (1.12) if $\{S_n\}$ is recurrent, so we may assume that this random walk is transient. In this case we can make the first term in the right hand side of (2.29) small by choosing M large (see (9), Theorem XIII.3.2). Finally, by (1.16) and (1.15) also

$$nu_\tau \le nh(\tau/2) = nh\big(\lfloor\frac{n}{g(n/2)}\rfloor\big) \to 0.$$

This completes the proof of (2.17) for even n. The case of odd n, which arises only when $p=1$ can be reduced to the case of even n, because $u_{n-r}/u_n \sim u_{n+1-r}/u_{n+1}$ if $p=1$, by virtue of (2.16). ■

Proof of Theorem 3. The properties in (1.21) are well known (see (30), Theorem C and also Theorem 4.1 in (19)). We merely add a few comments concerning (1.21) which will be needed for the examples.

As usual, if $\{S_n\}$ is a random walk on a Cayley graph $(G, \mathcal{S})$ then we take e to be the identity element of G. If $\{S_n\}$ is simple random walk on some other vertex transitive graph $\mathcal{G}$ then e

is any fixed vertex of $\mathcal{G}$. f_n and u_n are as in (1.7). In all cases $u_{n+m} \geq u_n u_m$. As we already pointed out in (2.7) and the following lines this implies that $\lim_{n\to\infty}[u_{np}]^{1/(np)}$ exists and is at least equal to $\sqrt{u_2} > 0$. Moreover, $u_m = 0$ if $p \nmid m$. Thus $U(z) = \sum_{n\geq 0} u_{np} z^{np}$ has the radius of convergence $\lim_{n\to\infty}[u_{np}]^{-1/(np)} \in [1, \infty)$ (recall $|u_n| \leq 1$). From the theory of recurrent events (see (9), Section XIII.3) it then follows that

$$U(z) = \frac{1}{[1 - F(z)]} \text{ for } |z| < 1,$$

where $F(z) = \sum_{n=1}^{\infty} f_n z^n$, as defined in the statement of the theorem. But the right hand side here is analytic on the open disc $\{z : |z| < \rho\}$ with $\rho = \min\{\rho_1, \rho_2\}$, where ρ_1 equals the first singularity of $F(\cdot)$ on the positive real axis, and $\rho_2 = \sup\{x > 0 : F(x) < 1\}$. Thus, also the power series $U(z) = \sum_{n=0}^{\infty} u_n z^n$ converges at least for $|z| < \rho$. On the other hand it cannot be that the powerseries for U converges on all of the disc $\{|z| < \rho + \varepsilon\}$ for some $\varepsilon > 0$, because $1/[1 - F(z)]$ cannot be analytic in such a disc. In fact, since F is a powerseries with non-negative coefficients its smallest singularity must be on the positive real axis. Thus the radius of convergence of U equals ρ. The fact that $\rho \leq \rho_2$ (and Fatou's lemma) shows that $F(\rho) \leq 1$, as claimed in (1.21).

To prove (1.24) from the conditions (1.22) and (1.23) we shall show that for all $0 < \eta < 1$

$$\lim_{n\to\infty, p|n} \frac{1}{n} E\{R_{(1-\eta)n} | \mathcal{E}_n\} = (1 - \eta)[1 - F(\rho)], \tag{2.30}$$

and then show that a good approximation to $(1/n)R_{(1-\eta)n}$ on $\mathcal{E}_n$ has a conditional variance, given $\mathcal{E}_n$, which tends to 0 as $n \to \infty$ (see (2.39)). Since the approximation can be made as precise as desired, and since $|R_n - R_{(1-\eta)n}| \leq \eta n$ this will give us (1.24). (Here and in the sequel we drop the largest integer symbol in $\lfloor (1 - \eta)n \rfloor$ for brevity; the proof of (2.19) demonstrates that this has no serious consequences.)

The proof of (2.30) is straightforward. We define

$$\begin{aligned} &J(v, k, r) \\ &\qquad = I\big[S_k = v, \text{ first time after } k \text{ at which } S_\cdot \text{ returns to } v \text{ is } k + r\big]. \end{aligned} \tag{2.31}$$

Again, by a last exit decomposition we have

$$R_{(1-\eta)n} = \sum_{0\leq k<(1-\eta)n} \sum_{v\in\mathcal{G}} \big[I[S_k = v] - \sum_{1\leq r<(1-\eta)n-k} J(v, k, r)\big] \tag{2.32}$$

For brevity we write W for the triple sum

$$\sum_{0\leq k<(1-\eta)n} \sum_{v\in\mathcal{G}} \sum_{1\leq r<(1-\eta)n-k} J(v, k, r).$$

With $I_n = I[\mathcal{E}_n]$ as before, we have as in (2.22)

$$E\{W I_n\} = \sum_{0\leq k<(1-\eta)n} \sum_{1\leq r<(1-\eta)n-k} f_r u_{n-r},$$

and hence

$$E\frac{1}{n}\{R_{(1-\eta)n}|\mathcal{E}_n\} = (1-\eta) - \frac{1}{n}\sum_{0\le k<(1-\eta)n}\sum_{\substack{1\le r<(1-\eta)n-k\\ p|r}} f_r\frac{u_{n-r}}{u_n}$$
$$= (1-\eta) - \frac{1}{n}\sum_{0\le k<(1-\eta)n}\sum_{\substack{1\le r<(1-\eta)n-k\\ p|r}} f_r\rho^r\frac{u_{n-r}\rho^{n-r}}{u_n\rho^n}. \tag{2.33}$$

Now, by (1.22),(1.23) and the fact that $F(\rho)<\infty$ we have, for any fixed $\eta'>0$, uniformly in $k\in[0,(1-\eta-\eta')n)$ that

$$\lim_{n\to\infty,p|n}\sum_{\substack{1\le r<(1-\eta)n-k\\ p|r}} f_r\rho^r\frac{u_{n-r}\rho^{n-r}}{u_n\rho^n} = \sum_{r\ge1,p|r} f_r\rho^r = F(\rho). \tag{2.34}$$

Moreover, for each fixed $\eta>0$, and $k\le(1-\eta)n$,

$$\sum_{\substack{1\le r<(1-\eta)n-k\\ p|r}} f_r\rho^r\frac{u_{n-r}\rho^{n-r}}{u_n\rho^n}$$

is bounded in n. (2.30) is immediate from these observations and (2.33).

We also obtain that for any given $\delta\in(0,1)$ we can choose $M=M(\delta)$, independent of n, such that for $p|n$

$$\frac{1}{n}E\Big\{\sum_{0\le k<(1-\eta)n}\sum_{v\in\mathcal{G}}\sum_{M<r\le(1-\eta)n-k} J(v,k,r)\big|\mathcal{E}_n\Big\}\le\delta.$$

This leads us to write

$$W_M = \sum_{0\le k<(1-\eta)n}\sum_{v\in\mathcal{G}}\sum_{1\le r\le M\wedge\big((1-\eta)n-k\big)} J(v,k,r).$$

With this notation we find

$$\lim_{n\to\infty,p|n}\frac{1}{n}E\{W_M|\mathcal{E}_n\} = (1-\eta)\sum_{r=1}^{M} f_r\rho^r \tag{2.35}$$

and for $p|n$

$$\frac{1}{n}E\big\{\big|R_{(1-\eta)n}-(1-\eta)n+W_M\big|\big|\mathcal{E}_n\big\}\le\delta. \tag{2.36}$$

Next we estimate $E\{W_M^2|\mathcal{E}_n\}$ for a fixed M. From the definition of W_M it follows that

$$E\{W_M^2 I_n\}$$
$$= \sum_{0\le k<(1-\eta)n}\sum_{v\in\mathcal{G}}\sum_{1\le r\le M\wedge\big((1-\eta)n-k\big)}\sum_{0\le\ell<(1-\eta)n}\sum_{w\in\mathcal{G}}\sum_{1\le s\le M\wedge\big((1-\eta)n-k\big)} E\{J(v,k,r)J(w,\ell,s)I_n\} \tag{2.37}$$

Now note that for fixed $\ell, \sum_{w\in\mathcal{G}}\sum_{1\le s\le M} J(w,\ell,s)$ is a sum of indicator functions of disjoint events and is therefore bounded by 1. Thus the terms in the multiple sum in (2.37) with $|k-\ell|\le M$ contribute at most

$$\sum_{0\le k<(1-\eta)n}\sum_{v\in\mathcal{G}}\sum_{1\le r\le M}\sum_{\ell:|k-\ell|\le M} E\{J(v,k,r)I_n\}$$
$$\le \sum_{0\le k<(1-\eta)n}\sum_{\ell:|k-\ell|\le M} P\{\mathcal{E}_n\}\le n(2M+1)u_n.$$

These terms are therefore $o(n^2u_n)$. Similarly to (2.22) the remaining terms contribute

$$\begin{aligned}
&2\sum_{0\le k<(1-\eta)n}\sum_{v\in\mathcal{G}}\sum_{1\le r\le M\wedge\big((1-\eta)n-k\big)}\sum_{k+M<\ell<(1-\eta)n}\sum_{w\in\mathcal{G}}\sum_{1\le s\le M\wedge\big((1-\eta)n-\ell\big)}\\
&\qquad\qquad E\{J(v,k,r)J(w,\ell,s)I_n\}\\
&=2\sum_{0\le k<(1-\eta)n}\sum_{v\in\mathcal{G}}\sum_{1\le r\le M\wedge\big((1-\eta)n-k\big)}\sum_{k+M<\ell<(1-\eta)n}\sum_{w\in\mathcal{G}}\sum_{1\le s\le M\wedge\big((1-\eta)n-\ell\big)}\\
&\qquad P\{S_k=v\}f_rP\{S_\ell=w|S_{k+r}=v\}f_sP\{S_n=e|S_{\ell+s}=w\}\\
&=2\sum_{0\le k<(1-\eta)n}\sum_{1\le r\le M\wedge\big((1-\eta)n-k\big)}\sum_{k+M<\ell<(1-\eta)n}\sum_{1\le s\le M\wedge\big((1-\eta)n-\ell\big)} f_rf_su_{n-r-s}.
\end{aligned}\tag{2.38}$$

After division by n^2u_n we find, just as in (2.33), (2.34) that

$$\limsup_{n\to\infty,p|n}\frac{1}{n^2}E\{W_M^2|\mathcal{E}_n\}\le(1-\eta)^2\Big[\sum_{1\le r\le M} f_r\rho^r\Big]^2.$$

Together with (2.35) this shows that

$$\lim_{n\to\infty,p|n}\frac{1}{n^2}\operatorname{Var}\{W_M|\mathcal{E}_n\}=0,\tag{2.39}$$

and therefore, $(1/n)W_M$, conditioned on $\mathcal{E}_n$, tends to $(1-\eta)\sum_{r=1}^M f_r\rho^r$ in probability as $n\to\infty$. By first taking M large and then η small in (2.36) we obtain the desired (1.24). ■

Example. Simple random walk on a regular tree. Here we shall explicitly calculate the values of F and $F(\rho)$ which were stated in (1.26). We take an arbitrary vertex of the tree $\mathcal{G}_b$ for e. This will remain fixed throughout the calculation. Also $\{S_n\}$ will be simple random walk on this tree. Unless otherwise stated $S_0=e$. This random walk has period $p=2$. For any vertex v, $d(v)$ denotes the number of edges in the simple path on $\mathcal{G}$ from e to v; this is also called the *height* of v. We set

$$T_n:=d(S_n).\tag{2.40}$$

Again, unless stated otherwise $T_0=d(S_0)=0$.

It is well known and easy to prove that with $\{S_n\}$ simple random walk on $\mathcal{G}_b$, $\{T_n\}$ is a nearest neighbor random walk on $\mathbb{Z}_+=\{0,1,2,\dots\}$ with transition probabilities

$$P(x,y)=\begin{cases}\frac{b}{b+1} & \text{if } x\ge1, y=x+1\\ \frac{1}{b+1} & \text{if } x\ge1, y=x-1\\ 1 & \text{if } x=0, y=1.\end{cases}\tag{2.41}$$

$P(x,y)=0$ if $|x-y|>1$ or $y<0$. From this observation we have

$$\begin{aligned} f_r &= P\{\text{first return by } T_\cdot \text{ to the origin is at time } r\}, \\ u_r &= P\{T_r = 0\}. \end{aligned} \tag{2.42}$$

$f_r = u_r = 0$ if r is odd, while explicit formulae for f_{2k} and for the generating functions of the f_r and u_r are known. Indeed, set

$$\lambda = \frac{b}{(b+1)^2}.$$

Then, by the arguments in (9) for the equations XIII.4.6-XIII.4.8 (see also (9), Section XI.3)

$$F(z) := \sum_{r=1}^{\infty} f_r z^r = \frac{b+1}{2b} - \frac{b+1}{2b}\big[1-4\lambda z^2\big]^{1/2},\ |z| \le 1. \tag{2.43}$$

Also,

$$F = F(1) = \frac{b+1}{2b} - \frac{b+1}{2b}[1-4\lambda]^{1/2} = \frac{b+1}{2b} - \frac{b+1}{2b}\cdot\frac{b-1}{b+1} = \frac{1}{b},$$

as claimed in (1.26).

As pointed out in the beginning of the proof of Theorem 3, $\rho = \min\{\rho_1, \rho_2\}$. In the present example ρ_1 is the first place where $1-4\lambda x^2$ becomes 0, i.e., $\rho_1 = 1/(2\sqrt{\lambda})$. $\rho_2 > \rho_1$ because $F(\rho_1) = F\big(1/(2\sqrt{\lambda}\big) = (b+1)/(2b)$, which is still less than 1. Hence, $\rho = 1/(2\sqrt{\lambda}) = (b+1)/(2\sqrt{b})$ and $F(\rho) = (b+1)/(2b)$. This proves (1.26). ■

To conclude we prove the equivalence of (1.11) and polynomial growth of the group on which the random walk takes place.

Lemma 4. *Let $(G, \mathcal{S})$ be a Cayley graph and let $\mathcal{V}(n)$ be as in* (1.18) *Then, for a random walk on $(G,\mathcal{S})$ which satisfies* (1.5) *and* (1.6), (1.11) *is equivalent to*

$$\begin{aligned} &\textit{there exist some constants } C_i < \infty \textit{ such that} \\ &\mathcal{V}(n) \le C_{19} n^{C_{20}} \textit{ for } n \ge 1. \end{aligned} \tag{2.44}$$

Proof. Some version of this result was known to N. Varopoulos. We learned the following argument from Laurent Saloff-Coste. Polynomial growth of G as in (2.44) implies (1.11) by means of Theorem 5.1 in (14), as we already observed for (1.17). For the converse, assume (1.11) holds, and let the constant $C_{21} > 0$ be such that $u_{4n}/u_{2n} \ge C_{21}$ for all $n \ge 1$. It is shown in (14), equation (10) that

$$u_{2n+m} \le \frac{2}{\mathcal{V}\big(r(n,m)\big)}, n, m \ge 1 \tag{2.45}$$

where

$$r(n,m) = \sqrt{m}\frac{u_{2n+m}}{u_{2n}}.$$

If we take $m = 2n$ we get from (2.45)

$$u_{4n} \le \frac{2}{\mathcal{V}\big(C_{21}\sqrt{2n}\big)}.$$

But we already saw in (2.26) that (1.11) implies $u_{2n} \ge n^{-C_{22}}$ for some constant C_{22} and large n. Thus

$$\mathcal{V}\big(C_{21}\sqrt{2n}\big) \le 2[2n]^{C_{22}} \text{ for large } n,$$

so that $\mathcal{V}(n)$ cannot grow faster than a power of n. ■

References

[1] A. Avez, Limite de quotients pour les marches aléatoires sur les groupes, C. R. Acad. Sc. Paris, Sér. A **276** (1973) 317-320. MR0315750

[2] Ph. Bougerol and T. Jeulin, Brownian bridge on hyperbolic spaces and on homogeneous trees, Probab. Theory Relat. Fields **115** (1999) 95-120. MR1715541

[3] D. I. Cartwright, Some examples of random walks on free products of discrete groups, Ann. Mat. Pura Appl. **151** (1988) 1-15. MR0964500

[4] Y. S. Chow and H. Teicher, Probability Theory, 3$^{\text{nd}}$ ed., Springer-Verlag, 1997. MR1476912

[5] F. M. Dekking, On transience and recurrence of generalized random walks, Z. Wahrsch. verw. Gebiete **61** (1982) 459-465. MR0682573

[6] M. D. Donsker and S. R. S. Varadhan, On the number of distinct sites visited by a random walk, Comm. Pure Appl. Math. **32** (1979) 721-747. MR0539157

[7] A. Dvoretzky and P. Erdős, Some problems on random walk in space, Proc. Second Berkeley Symposium on Mathematical Statistics and Probability (J. Neyman ed.), pp. 353-368, Univ. of California Press, 1951 MR0047272

[8] A. Erschler, Isoperimetry for wreath products of Markov chains and multiplicity of selfintersections of raqndom walks, Probab. Theory Relat. Fields, **136** (2006) 560-586. MR2257136

[9] W. Feller, An Introduction to Probability Theory and its Applications, vol I, 3$^{\text{rd}}$ ed., John Wiley & Sons, 1968. MR0228020

[10] P. Gerl, Über die Anzahl der Darstellungen von Worten, Monatsh. Math. **75** (1971) 205-214. MR0295923

[11] P. Gerl, Diskrete, mittelbare Gruppen, Monatsh. Math. **77** (1973) 307-318. MR0341550

[12] R. I. Grigorchuk, Degrees of growth of finitely generated groups and the theory of invariant means, Math. USSR Izvestiya **25** (1985) 259-300. MR0764305

[13] Y. Hamana, An almost sure invariance principle for the range of random walks, Stochastic Process. Appl. **78** (1998) 131-143, MR1657371

[14] W. Hebisch and L. Saloff-Coste, Gaussian estimates for Markov chains and random walks on groups, Ann. Probab. **21** (1993) 673-709. MR1217561

[15] N. C. Jain and W. E. Pruitt, The range of random walk, Proc. Sixth Berkeley Symposium on Mathematical Statistics and Probability (L. M. Le Cam, J. Neyman and E. L. Scott eds.) Vol 3, pp. 31-50, Univ. of California Press, 1972. MR0410936

[16] V. A. Kaimanovich and A. M. Vershik, Random walks on discrete groups: boundary and entropy, Ann. Probab. **11** (1983) 457-490. MR0704539

[17] H. Kesten, Full Banach mean values on countable groups, Math. Scan. **7** (1959) 146-156. MR0112053

[18] H. Kesten, Subdiffusive behavior of random walk on a random cluster, Ann. Inst. Henri Poincaré, Probabilités et Statistiques **22** (1986) 425-487. MR0871905

[19] S. P. Lalley, Finite range random walk on free groups and homogeneous trees, Ann. Probab. **21** (1993) 2087-2130. MR1245302

[20] B. Ya. Levit and S. A. Molchanov, Invariant chains on a free group with a finite number of generators, Vestnik Moscow Univ. **26** (1971) 80-88. MR0298721

[21] R. Lyons and Y. Peres, Probability on Trees and Networks. To appear (2006); see http://mypage.iu.edu/ rdlyons/prbtree/prbtree.html.

[22] S. K. Nechaev, A. Yu. Grosberg and A. M. Vershik, Random walks on braid groups: Brownian bridges, complexity and statistics, J. Physics A Math. Gen. **29** (1996) 2411-2434. MR1396938

[23] A. L. T. Paterson, Amenability, Math Surveys and Monographs **29**, Amer. Math. Soc. 1988. MR0961261

[24] C. Pittet and L. Saloff-Coste, On random walks on wreath products, Ann. Probab. **30** (2002) 948-977. MR1905862

[25] W. Rudin, Functional Analysis, 2^{nd} ed., Mc-Graw Hill, 1991. MR1157815

[26] S. Sawyer, Isotropic random walks in a tree, Z. Wahrsch. verw. Gebiete **42** (1978) 279-292. MR0491493

[27] A. Sinclair, Algorithms for Random Generation and Counting: a Markov Chain Approach, Birkhauser Verlag, Boston, 1993. MR1201590

[28] F. Spitzer, Principles of Random Walk, 2^{nd} ed., Springer-Verlag, 1976. MR0388547

[29] F. Spitzer in Discussion of Subadditive ergodic theory by J. F. C. Kingman, Ann Probab. **1** (1973) 904-905. MR0356192

[30] D. Vere-Jones, Ergodic properties of nonnegative matrices I, Pacific J. Math. **22** (1967) 361-386. MR0214145

[31] W. Woess, Random Walks on Infinite Graphs and Groups. Cambridge Univ. Press, 2000. MR1743100

[32] R. Young, Averaged Dehn functions for nilpotent groups, arXiv math. GR/0510665.