Combinatorics, Probability and Computing (2003) 12, 345–358. © 2003 Cambridge University Press
DOI: 10.1017/S0963548302005588 Printed in the United Kingdom

Branching Random Walks on Quasi-Transitive Graphs

ALAN STACEY

Statistical Laboratory, Department of Pure Mathematics and Mathematical Statistics,
University of Cambridge, Centre for Mathematical Sciences,
Wilberforce Road, Cambridge CB3 OWB, England
(e-mail: A.M.Stacey@dpmms.cam.ac.uk)

Received 29 June 2002; revised 21 October 2002

The branching random walk on a *regular* graph turns out to be particularly easy to analyse using results for the corresponding simple random walk. In this way, one can show that there is an intermediate phase of weak survival if and only if the graph is nonamenable. No such simple analysis holds more generally, and it is known that the nonamenability equivalence does not extend to general connected graphs of bounded degree (although we observe that it does hold for such graphs if the branching random walk is modified in a certain natural way). The most important general class of (bounded degree, connected) graphs for which it is thought that the equivalence may hold is that of quasi-transitive graphs: we show that this is indeed the case.

1. Introduction

A good deal of work in recent years has studied stochastic processes, such as percolation and various interacting particle systems, on general (infinite, connected) graphs. Typically one kind of behaviour is observed on amenable graphs and another kind of behaviour on nonamenable graphs. Sometimes additional conditions are imposed, without which unexpected behaviour can occur: one nearly always assumes that graphs are of bounded degree; more strongly, one may insist that graphs are quasi-transitive, transitive, or transitive and unimodular; or restrict attention to special classes of graphs such as trees or Cayley graphs.

In this paper we consider the branching random walk, particularly on quasi-transitive graphs. Our aim throughout will be to give reasonably elementary and self-contained proofs, although in defining the processes it is necessary to assume a little familiarity with continuous-time Markov processes or interacting particle systems; see the books [8], [4] and [10]. We begin with the necessary definitions and background information.

Throughout this paper, any graph $G = (V, E)$ – consisting of a set of *vertices* or *sites* V and a set of (undirected) *edges* $E \subseteq V^{(2)}$ – will be assumed to be connected and of bounded

https://doi.org/10.1017/S0963548302005588 Published online by Cambridge University Press

degree; where, as usual, the *degree* $d(v)$ of a vertex v is the number of neighbours it has. We are generally interested in infinite graphs, although our theorems will be (trivially) true for finite graphs. A graph is said to be *transitive* if the automorphism group acts transitively on the vertex set. It is said to be *quasi-transitive* if this action has only finitely many orbits: informally speaking, there are only finitely many different types of vertex. Although the extension of results from transitive to quasi-transitive graphs can be very straightforward, there are cases where it not; as we shall see, the branching random walk is an example of the latter.

The *branching random walk* on $G = (V, E)$ is a continuous-time Markov process whose state space is a suitable subset (see remarks following) of $\mathbb{N}^V = \{0, 1, 2, \dots\}^V$. We shall denote the state at time t by η_t and the process evolves according to the following rules. For each $v \in V$,

$$\eta_t(v) \to \eta_t(v) - 1 \quad \text{at rate } \eta_t(v),$$
$$\eta_t(v) \to \eta_t(v) + 1 \quad \text{at rate } \lambda \sum_{u:u\sim v} \eta_t(u),$$

where $\lambda \geqslant 0$ is a fixed parameter and $u \sim v$ denotes that u is a neighbour of v. One typically interprets $\eta_t(v)$ as the number of particles at v at time t. Each particle dies at rate 1, independently of other particles, so the number of particles at a given site decreases at a rate equal to the number of particles at that site, $\eta_t(v)$; this explains the first transition rate given. The interpretation of the second rate is that each particle gives birth to new particles at each neighbouring site, at rate λ.

Some restriction on the starting state is necessary to prevent the process 'exploding', with infinitely many particles piling up at a single site. The issue of the existence and uniqueness of a process defined by these transition rates is discussed in detail in [14] and is not totally straightforward. For the purposes of this paper, it will be sufficient to consider the countable state-space process in which the total number of particles ($\sum_v \eta_t(v)$) is finite; it is not hard to see that on a graph of bounded degree, the branching random walk starting from such a state is a simple continuous-time Markov chain with only finitely many transitions taking place in any finite time period.

The branching random walk is of interest in its own right, and also because of comparisons with the contact process. The contact process has $\{0, 1\}^V$ as its state space, but otherwise is the same as the branching random walk, *i.e.*, it has the same transition rates, except that any transition which would cause more than one particle to be at a site is forbidden. The branching random walk behaves similarly to the contact process in many respects, and stochastically dominates it (see, for example, [10], for a definition of this stochastic monotonicity). However, it is typically easier to study since it has better independence properties: if η_t^α and η_t^β are branching random walks with starting states α and β respectively, then $\eta_t^\alpha + \eta_t^\beta$ is a branching random walk with starting state $\alpha + \beta$; no result of this kind holds for the contact process.

We will be most interested in the branching random walk starting from a single particle at the origin (some distinguished vertex of our graph), which we denote by η_t^O. The key concepts, which we now introduce, describe the long-term behaviour of the process. We

https://doi.org/10.1017/S0963548302005588 Published online by Cambridge University Press

say that the branching random walk *survives globally* if

$$\mathbb{P}\big(\eta_t^O \equiv 0 \text{ for sufficiently large } t\big) < 1,$$

and that the branching random walk *survives locally* if

$$\mathbb{P}\big(\eta_t^O(O) = 0 \text{ for sufficiently large } t\big) < 1.$$

If the process survives globally but does not survive locally we say that it *survives weakly*. Essentially identical definitions are also made for the contact process. For each process we then define two critical values:

$$\lambda_1 = \inf\{\lambda \geqslant 0 : \eta_t^O \text{ survives globally}\},$$
$$\lambda_2 = \inf\{\lambda \geqslant 0 : \eta_t^O \text{ survives locally}\}.$$

It is well known that, for either process on an infinite connected graph of bounded degree, λ_1 and λ_2 take positive finite values (see, for example, the remarks preceding Proposition 1.1 of [14]). If $\lambda_1 < \lambda_2$ then there is a phase of weak survival. A fundamental problem is to classify those graphs for which this phase exists. For the contact process the answer is only known for a few specific graphs. For $\mathbb{Z}^d$ (any $d \geqslant 1$) it follows from results of Bezuidenhout and Grimmett [1] that $\lambda_1 = \lambda_2$. For a certain class of trees, including all homogeneous trees of degree $d \geqslant 3$, we have $\lambda_1 < \lambda_2$ (see [13], [9], [19]). In general, one expects that on 'large' or rapidly growing graphs we have $\lambda_1 < \lambda_2$, and that on 'small' graphs one has $\lambda_1 = \lambda_2$. In order to make this notion precise we need a further definition.

Given a graph $G = (V, E)$ we define the *Cheeger constant* of G by

$$\iota_E(G) = \inf_{\substack{S \subseteq V \\ 0 < |S| < \infty}} \frac{|\partial_E S|}{|S|}, \tag{1.1}$$

where $\partial_E S$ denotes the edge-boundary of S, *i.e.*, the set of those edges which have one endpoint lying in S and the other in $V \setminus S$. We say that a graph is *amenable* if and only if $\iota_E(G) = 0$. For a general background about processes and amenability, see, for example, [11] or [18].

It was conjectured in [13] that $\lambda_1 < \lambda_2$ for the contact process on any nonamenable tree of bounded degree (although the terminology used there is a little different) and conjectured in [12] that the branching random walk on a graph has $\lambda_1 < \lambda_2$ if and only if the contact process does. Both of these conjectures were shown to be false in [14]. Of particular relevance to this paper is the fact, established in [14], that there exist nonamenable trees of bounded degree where the branching random walk has no phase of weak survival, and amenable trees of bounded degree where it has such a phase. These trees arise naturally from branching processes, but they are not quasi-transitive.

Conjecture 5.1 of [14] states that, for both the branching random walk and the contact process on a quasi-transitive graph, there is a phase of weak survival if and only if the

https://doi.org/10.1017/S0963548302005588 Published online by Cambridge University Press

graph is nonamenable. In this paper we shall show that this conjecture holds for the branching random walk. This will be our principal result, Theorem 3.1.

2. Connections with spectral radii and simple random walks

In this section we shall see that the upper critical value for the branching random walk, λ_2, is closely related to the spectral radius of a certain operator. We shall also state a result about the spectral radius of the simple random walk and see what this implies about the branching random walk.

Let $G = (V, E)$ be a graph, assumed to be connected and of bounded degree, as stated at the start of Section 1. Let A be the adjacency matrix of G. Then, for any two vertices x and y, $A^n(x, y)$ (the (x, y)th entry of the matrix A^n) is equal to the number of walks of length n from x to y. Let O be some distinguished vertex of G, and let $M_n = A^n(O, O)$ be the number of walks of length n from O to itself. Then $M_{n+m} \geqslant M_n M_m$, so a standard supermultiplicativity argument shows that

$$\lim_{n\to\infty} M_{2n}^{1/2n} = M \tag{2.1}$$

exists and is equal to $\sup_n M_n^{1/n}$. Note that if G is bipartite then $M_n = 0$ for n odd; otherwise $\lim_{n\to\infty} M_n^{1/n}$ exists. Note also that, for any vertices x and y, if there is a walk of length d_x from O to x and one of length d_y from O to y, then we have

$$A^{n+2d_x+2d_y}(O, O) \geqslant A^{n+d_x+d_y}(x, y) \geqslant A^n(O, O),$$

for any n. Therefore

$$\lim_{n\to\infty} \left[A^n(x, y)\right]^{1/n} = M,$$

provided that, if the graph is bipartite, the limit is taken using only values of n of the appropriate parity.

As we indicated when defining the branching random walk, one can explain the rate at which $\eta_t(v)$ increases by considering each particle to give birth to new particles at each neighbouring site at rate λ, with all particles and all neighbouring sites being treated independently in this regard. In this way, whenever $\eta_t(v)$ increases by 1 one can think of a new particle's being born at v; moreover, this particle has a parent, which is necessarily a particle at some neighbour of v. Now suppose that we label each new particle's by a pair (v, t), where v is the vertex at which the particle sits, and t is the time at which the particle is born. If the process starts from a single particle at the origin (labelled $(O, 0)$) then, for each particle (v, t), there is a chain of descendants of $(O, 0)$, $(v_0, t_0), \ldots, (v_n, t_n)$, with each v_i a neighbour of v_{i+1}, $t_0 < t_1 < \cdots < t_n$, $v_0 = O$, $t_0 = 0$, $v_n = v$ and $t_n = t$. Forgetting the times, to each new particle we have associated a finite walk starting at the origin, in this case $v_0 v_1 \cdots v_n$.

One can make the argument of the preceding paragraph rigorous via the graphical representation of the process. The details of this representation are a little involved, so we refer the interested reader to the first part of Section 3 of [14]; it is adapted from the well-known graphical representation of Harris for the contact process [7]; see also [4] or

https://doi.org/10.1017/S0963548302005588 Published online by Cambridge University Press

[10]. Nevertheless, it is hoped that the description of the previous paragraph is sufficient for an understanding of this paper.

The point of view just outlined also enables us to see the independence property of the branching random walk referred to earlier: if η_t^α and η_t^β are independent branching random walks with starting states α and β respectively, then $\eta_t^\alpha + \eta_t^\beta$ is a branching random walk with starting state $\alpha + \beta$. This also holds for a (countably) infinite decomposition, so the process from any starting state can be viewed as the superposition of numerous independent branching random walks, each starting from just one particle.

The point of view just outlined, and some standard calculations, allow one to relate M directly to λ_2, giving the following lemma. We give only a sketch of the proof and refer the reader to Lemma 3.1 of [14] for further details. The branching process comparisons referred to in the sketch are of a similar flavour to those described in detail in the proof of Proposition 3.2.

Lemma 2.1. *Let λ_2 be the upper critical value for the branching random walk on a graph G and let M be as in* (2.1). *Then*

$$\lambda_2 = 1/M.$$

Sketch of proof. Consider the branching random walk starting from a single particle at O. As just described, each new particle that appears at some vertex v, at any time, is associated to some walk from O to v.

For every particle that appears at a vertex u, and for each neighbour w of u, the particle at u gives birth to some random number of new particles at w before the particle at u itself dies. This random number has expectation λ. So it is easily seen, by induction on the length of walks, that expected number of particles associated to a given walk of length n (starting at the origin) is λ^n. The number of walks from the origin to itself, of length n, is approximately M^n, so the expected total number of particles that are born *at the origin* associated to all walks of length n is approximately $M^n\lambda^n$. It is not too hard to see from this, using some branching process comparisons, that with positive probability the origin is infected at arbitrarily large times if $\lambda M > 1$ and not if $\lambda M < 1$.

It is useful to characterize the quantity M, of (2.1), in terms of the norm and spectral radius of a certain operator. Given two functions on the vertices of our graph, $f, g : V \to \mathbb{R}$, define an inner product by

$$\langle f, g \rangle = \left(\sum_{v \in V} f(v)g(v) \right)$$

and a norm by

$$\|f\|_2 = \langle f, f \rangle^{1/2};$$

and let $l_2(V)$ consist of those functions with finite norm. Consider A as a linear

https://doi.org/10.1017/S0963548302005588 Published online by Cambridge University Press

endomorphism of $l_2(V)$ in the obvious way, so Af is defined by

$$Af(x) = \sum_{y:y\sim x} f(y).$$

It is clear that A is a bounded operator with $\|A\| \leqslant \Delta$, where Δ is the maximum degree of G. The following lemma is well known, at least in slightly variant forms. A proof is included for completeness, and is adapted from a proof of a similar result in Gerl [5].

Lemma 2.2. *Let A be the adjacency matrix of a graph G, regarded as a linear operator on $l_2(V)$, and let M be as in* (2.1). *Then*

$$\|A\| = \rho(A) = M,$$

where $\rho(A)$ is the usual spectral radius of A, $\lim_{n\to\infty}\|A^n\|^{1/n}$.

Proof. Observe that

$$\begin{aligned}\langle f, Ag\rangle &= 2\sum_{\text{edges } xy} f(x)g(y)\\ &= \langle Af, g\rangle,\end{aligned}$$

so A is self-adjoint. Therefore, we have that $\|A\| = \rho(A)$, by the following standard argument:

$$\begin{aligned}\|A\|^2 &= \sup_{\|f\|_2=1}\langle Af, Af\rangle\\ &= \sup_{\|f\|_2=1}\langle A^2f, f\rangle\\ &\leqslant \|A^2\|.\end{aligned}$$

Since, for any operator A, we have $\|A^2\| \leqslant \|A\|^2$, we see that $\|A\|^2 = \|A^2\|$ and, by induction, $\|A^{2^n}\| = \|A\|^{2^n}$ for any n; from which we conclude, using the definition of $\rho(A)$ given above, that $\rho(A) = \|A\|$.

To see that $\|A\|$ is equal to the quantity M, we take any $\epsilon > 0$, and we may then find a function f of finite support with $\|f\|_2 = 1$ and $\|Af\| \geqslant \|A\| - \epsilon$. We see that

$$\begin{aligned}\|Af\|^2 &= \langle Af, Af\rangle\\ &= \langle A^2f, f\rangle\\ &\leqslant \|A^2f\|.\end{aligned} \tag{2.2}$$

We then have

$$\begin{aligned}\|Af\|^4 &= \left(\|Af\|^2\right)^2\\ &\leqslant \|A^2f\|^2, \qquad \text{by (2.2)},\\ &= \langle A^4f, f\rangle, \qquad \text{by self-adjointness},\\ &\leqslant \|A^4f\|.\end{aligned}$$

https://doi.org/10.1017/S0963548302005588 Published online by Cambridge University Press

Similarly, for any n, we have that $\langle A^{2^n} f, f\rangle \geqslant \|Af\|^{2^n}$. Hence

$$\begin{aligned}\|A\| - \epsilon &\leqslant \langle A^{2^n} f, f\rangle^{1/2^n} \\ &= \left(\sum_{x \in \mathrm{supp}(f)} f(x) \sum_{y \in \mathrm{supp}\,(f)} A^{2^n}(x,y) f(y) \right)^{1/2^n} \\ &\to M \text{ as } n \to \infty.\end{aligned}$$

Since this is true for any $\epsilon > 0$, we see that $\|A\| \leqslant M$. However, since $A^n(O,O) \leqslant \|A^n\|$, we certainly have $M \leqslant \rho(A)$, so $M = \rho(A) = \|A\|$. □

We now state a result of Gerl [5] concerning the simple random walk on a graph G. Recall that the simple random walk on a connected locally finite graph is a discrete-time Markov chain whose state space consists of the vertices of graph. At each step the state jumps to one of the neighbours of the current state, chosen uniformly at random. The spectral radius of such a walk is defined to be $\sigma(P) = \lim_{n\to\infty}[P^{2n}(O,O)]^{1/2n}$, where P is the transition matrix of the walk and O is some distinguished vertex. This is also equal to the usual spectral radius, much as above, when P is regarded as an operator. (A slightly different norm has to be used, to make P self-adjoint. See [5] for further details. The study of Markov chains as linear operators on normed spaces is highly developed: see, for example, [16].) In [5], the graphs (of bounded degree) for which the spectral radius is less than 1 are characterized. We state the theorem only for the simple random walk, although it was originally stated for more general (strongly reversible) walks.

Theorem 2.3. (Gerl) *The simple random walk on a connected graph of bounded degree has spectral radius less than* 1 *if and only if the graph is nonamenable.*

This theorem, and the preceding observations, immediately enable us to determine when $\lambda_1 < \lambda_2$ for graphs in one particular class, namely the class of regular graphs. Recall that a graph is said to be *regular of degree* d if every vertex has degree d.

Theorem 2.4. *Suppose that G is a regular graph of degree d. Then the branching random walk on G satisfies $\lambda_1 < \lambda_2$ if and only if G is nonamenable.*

Proof. Since every vertex of G has d neighbours, every particle in the branching random walk, regardless of its location, gives birth to new particles at rate λd. It also dies at rate 1. Therefore the total number of particles behaves as a simple continuous-time branching process. This is easy to analyse, either via an embedded discrete-time branching process, or using the jump chain which is just a simple random walk on $\mathbb{N}$ with jump probabilities $1/(\lambda d + 1)$ and $\lambda d/(\lambda d + 1)$. We see that, starting from a single particle, there is a positive probability that the total number of particles never reaches 0 if and only if $\lambda d > 1$.

https://doi.org/10.1017/S0963548302005588 Published online by Cambridge University Press

Therefore $\lambda_1 = 1/d$. Since we have $\lambda_2 = 1/\rho(A)$ by Lemmas 2.1 and 2.2, we see that

$$\begin{aligned} \lambda_1 < \lambda_2 &\Leftrightarrow \rho(A) < d \\ &\Leftrightarrow \rho\left(\frac{1}{d}A\right) < 1. \end{aligned} \tag{2.3}$$

However, $(1/d)A$ is just the transition matrix for the simple random walk on G so, by Theorem 2.3, condition (2.3) holds if and only if G is nonamenable. □

The proof just given clearly only applies to regular graphs. We know from the examples mentioned in the Introduction that the result does not extend, even as a one-way implication, to all connected graphs of bounded degree. However, it is interesting to note in passing that, if we modify the definition of the branching random walk in a certain simple way, then the result does extend, and the proof just given applies essentially unchanged to all such graphs.

Define the *modified branching random walk* to have the same state space as the branching random walk and the following transition rates:

$$\begin{aligned} &\eta_t(v) \to \eta_t(v) - 1 \quad \text{at rate } \eta_t(v), \\ &\eta_t(v) \to \eta_t(v) + 1 \quad \text{at rate } \lambda \sum_{u:u\sim v} \frac{1}{d(u)} \eta_t(u), \end{aligned}$$

where $d(u)$ is the degree of a vertex u. One can interpret this as each particle's giving birth to new particles at rate λ, with each new particle going to one of the neighbouring vertices chosen uniformly at random. Therefore, the total number of particles behaves as a simple continuous-time branching process, even if the graph is not regular, and we have $\lambda_1 = 1$. The analogue of Theorem 2.4 is the following, much more general, result.

Theorem 2.5. *Let G be a connected graph of bounded degree. Then the modified branching random walk on G has $\lambda_1 < \lambda_2$ if and only if the graph is nonamenable.*

Proof. Let P be the transition matrix for the simple random walk on G. Then the expected number of particles born at the origin corresponding to walks of length n is exactly $\lambda^n P^n(O, O)$ so, exactly as in the proof of Lemma 2.1, we have $\lambda_2 = 1/\sigma(P)$, where $\sigma(P)$ is the spectral radius of the walk as defined above. Since $\lambda_1 = 1$, we see that $\lambda_1 < \lambda_2$ if and only if $\sigma(P) < 1$ which, by Theorem 2.3, occurs if and only if G is nonamenable. □

What we have called the modified branching random walk has been used as the definition of the branching random walk by some authors, such as Schinazi [17]. In [17], λ_2 is related to a variant of the spectral radius, so Theorem 2.5 in essence just combines results of Gerl [5] and Schinazi [17].

It has been generally observed (*e.g.*, [15]) that it is not always obvious how best to generalize a process defined on lattices or general transitive graphs to quasi-transitive or more general classes of graphs. The fact that Theorem 2.5 holds for the modified branching random walk suggests that this modified definition is in some ways a better

https://doi.org/10.1017/S0963548302005588 Published online by Cambridge University Press

one. Note that the two versions of the branching random walk are equivalent, via a reparametrization, for regular graphs. However, our original branching random walk has some significant advantages: it is, for example, closest in definition to the contact process on the same graph, and stochastically dominates it. Of course, the definition of the contact process on general graphs could be modified to bring it into line with the modified branching random walk, but then it would lose the important desirable property of self-duality (see [8] for a definition and an explanation of why the contact process has this). This claim is very easily seen by considering almost any small example. Take, for instance, the modified contact process on the graph P_2, with vertex set $\{0,1,2\}$ in which there are two edges, $\{0,1\}$ and $\{1,2\}$. Identify the state of the process at any time with the subset of vertices that are in state 1 and let ξ_t^X denote the modified contact process with starting state X and parameter $\lambda = 1$. Let $B = \{0,2\}$ and let $C = \{1\}$. Then we have for small t that

$$\mathbb{P}\big(\xi^B \cap C \neq \emptyset\big) = 2t + o(t), \qquad \mathbb{P}\big(\xi^C \cap B \neq \emptyset\big) = t + o(t),$$

so self-duality fails.

In the last part of this section, we prepare the ground for our main theorem of Section 3 by establishing a fundamental property of nonamenable graphs. This property is the key part of Gerl [5], and also appears elsewhere in variant forms. We include it here, in a straightforward way, for completeness. Given a graph $G = (V,E)$ of bounded degree and a function $f \in l_2(V)$, define

$$\|f\|_{D(2)} = \left(\sum_{\{x,y\} \in E} (f(x) - f(y))^2 \right)^{1/2}.$$

We shall prove the following theorem. In the remainder of this section, including the statement of this theorem, graphs are no longer assumed to be connected. We shall later need to apply Theorem 2.6 to a graph, G^2, which may have two connected components.

Theorem 2.6. *Let $G = (V,E)$ be a nonamenable graph of bounded degree. Then there exists $c > 0$ such that, for any function $f \in l_2(V)$,*

$$\|f\|_{D(2)} \geqslant c \, \|f\|_2. \tag{2.4}$$

Note that the definition of nonamenability is exactly the statement (2.4) in the case when f is an indicator function. In order to prove Theorem 2.6 we shall first prove the equivalent statement for l_1 norms. Define

$$\|f\|_1 = \sum_{v \in V} |f(v)|, \qquad \|f\|_{D(1)} = \sum_{\{x,y\} \in E} |f(x) - f(y)|.$$

Proposition 2.7. *Let $G = (V,E)$ be a nonamenable graph. Then, for any $g \in l_1(V)$,*

$$\|g\|_{D(1)} \geqslant \iota_E \|g\|_1, \tag{2.5}$$

where ι_E is the Cheeger constant defined by (1.1).

https://doi.org/10.1017/S0963548302005588 Published online by Cambridge University Press

Proof. By taking limits, it is sufficient to prove that (2.5) holds for functions g of finite support. Given such a g, let the set of values (including 0) taken by g be $\{m_1, \ldots, m_n\}$ with $m_1 < m_2 < \cdots < m_n$. We proceed by induction on n. Clearly the statement holds when $n = 1$ (when $f \equiv 0$). Now suppose $n > 1$ and the statement holds for smaller values of n. By replacing g by $-g$ if necessary, we may assume $m_n > 0$. Let $W = \{v \in V : g(v) = m_n\}$ and define a function g^- by

$$g^-(v) = \begin{cases} m_{n-1} & \text{if } v \in W, \\ g(v) & \text{otherwise.} \end{cases}$$

Note that $m_{n-1} \geqslant 0$, since one of the values taken by g is 0. Now, it is clear that

$$\|g\|_1 - \|g^-\|_1 = (m_n - m_{n-1})|W|, \qquad \|g\|_{D(1)} - \|g^-\|_{D(1)} = (m_n - m_{n-1})|\partial_e W|. \tag{2.6}$$

Since g^- takes only $n-1$ values, equation (2.5) holds, by the induction hypothesis, with g replaced by g^-. Combining this fact with (2.6) and the fact that $|\partial_E W| \geqslant \iota_E |W|$ gives (2.5) as required. □

Proof of Theorem 2.6. Let G be as in the statement of the theorem, with isoperimetric constant i_E and maximum degree Δ. Take $f \in l_2(V)$ and apply Proposition 2.7 with $g = f^2$. Then we have

$$\begin{aligned} \|f\|_2^2 = \|f^2\|_1 &\leqslant \iota_E \|f^2\|_{D(1)} \\ &= \iota_E \sum_{xy \in E} |f^2(x) - f^2(y)| \\ &= \iota_E \sum_{xy \in E} |(f(x) - f(y))(f(x) + f(y))| \\ &\leqslant \iota_E \left(\sum_{xy \in E} (f(x) - f(y))^2 \right)^{1/2} \left(\sum_{xy \in E} (f(x) + f(y))^2 \right)^{1/2} \\ &\leqslant 2\Delta \iota_E \|f\|_{D(2)} \|f\|_2, \end{aligned}$$

the penultimate inequality using Cauchy–Schwartz and the last inequality using the fact that $(f(x) + f(y))^2 \leqslant 2(f^2(x) + f^2(y))$. Dividing both sides of the above by $\|f\|_2$, we obtain

$$\|f\|_2 \leqslant 2\Delta \iota_E \|f\|_{D(2)},$$

giving (2.4) with $c = 2\Delta \iota_E$. □

We close this section with a remark on the proof of Theorem 2.6. Looked at slightly differently, this theorem asserts that, if there is an f with $\|f\|_{D(2)}/\|f\|_2$ small, then there is a set W with $|\partial_E W|/|W|$ small; yet the proof does not really give us a simple recipe for finding such a W. The Cauchy–Schwartz inequality muddies the waters somewhat, and a proof without it might be rather desirable. On the other hand, the converse is trivial: given a W with $|\partial_E W|/|W|$ small, just set $f = \mathbf{1}_W$.

https://doi.org/10.1017/S0963548302005588 Published online by Cambridge University Press

3. Proof of the main theorem

In this section we shall prove the following theorem, which extends Theorem 2.4 to quasi-transitive graphs. The branching random walk referred to throughout this section is the unmodified one defined in Section 1.

Theorem 3.1. *Let G be a quasi-transitive connected locally finite graph. Then the branching random walk on G has two phase transitions, i.e., $\lambda_1 < \lambda_2$, if and only if G is nonamenable.*

Our first goal is to identify the value of λ_1 for quasi-transitive graphs. For such a graph, G, the automorphism group acts on the vertex set with finitely many orbits; let these orbits be denoted $\mathcal{O}_1, \dots, \mathcal{O}_n$. We shall say that a vertex in orbit $\mathcal{O}_i$ has *type* i. For $1 \leqslant i, j \leqslant n$, let $\tilde{A}(i, j)$ be the number of neighbours of type j that each vertex of type i has. We shall refer to $\tilde{A}$ as the *quotient adjacency matrix*. $\tilde{A}^n(i, j)$ counts the number of walks of length n starting from some fixed vertex of type i and finishing at any vertex of type j.

By the Perron–Frobenius theorem (see, for example, [3]), $\tilde{A}$ has a largest positive eigenvalue, μ, and associated positive eigenvector $(\gamma_1, \dots, \gamma_n)$; furthermore, for any i, j, we have

$$\limsup_{n\to\infty} \left(\tilde{A}^n(i, j)\right)^{1/n} = \mu. \tag{3.1}$$

It is not too hard to see that λ_1 is the reciprocal of μ, as we now show.

Proposition 3.2. *Let λ_1 be the lower critical value for the branching random walk on a quasi-transitive graph G. Let $\tilde{A}$ be the quotient adjacency matrix of G and let μ be its largest eigenvalue. Then*

$$\lambda_1 = 1/\mu.$$

Proof. Consider the branching random walk on G with parameter λ. Suppose firstly that $\lambda < 1/\mu$, and choose $\epsilon > 0$ so that $\lambda(\mu + \epsilon) < 1$. Then (3.1) implies that, for sufficiently large n, the total number of walks of length n starting from the origin is at most $(\mu + \epsilon)^n$, and so the expected number of new particles that appear, resulting from a single initial particle at the origin and associated with walks of length n, is at most $\lambda^n(\mu + \epsilon)^n$, much as in the proof of Lemma 2.1. Summing over n, the expected total number of new particles to appear is finite, implying that the process almost surely does not survive.

Now suppose that $\lambda > 1/\mu$, and again consider the branching random walk starting from a single particle at the origin. Suppose, without loss of generality, that the origin is of type 1. Let X_n be the number of particles arising at all vertices of type 1 whose associated walk (described at start of Section 2) has length n. As in the proof of Lemma 2.1, the expected number of particles associated to each such walk is λ^n, so $\mathbb{E}X_n = \lambda^n \tilde{A}^n(1, 1)$. Then (3.1) implies that, for some n, we have $\mathbb{E}X_n > 1$.

At the start of Section 2, it was explained how each particle can be regarded as having a chain of ancestors leading back to some particle which was present at time 0. Now, in this way, each of the X_n particles associated to walks of length n has a certain number

https://doi.org/10.1017/S0963548302005588 Published online by Cambridge University Press

of descendants associated with walks of length $2n$, and this number of descendants has the same distribution as X_n. So we see that, conditional on $X_n = x$, X_{2n} is at least the sum of x independent random variables, each having the same distribution as X_n. Iterating this observation, we can see that $(X_{kn})_{k\geqslant 0}$ dominates a branching process with offspring distribution equal to the distribution of X_n. This branching process has a positive probability of survival, so there is a positive chance that there are infinitely many particles born at the vertices of type 1; this implies via standard and straightforward arguments (see, for example, [14, Lemma 3.1]) that the branching random walk survives. □

We are now in a position to prove our principal theorem.

Proof of Theorem 3.1. Let $G = (V, E)$ be a quasi-transitive locally finite (and hence bounded degree) nonamenable graph. The critical values λ_1 and λ_2 are characterized by Lemma 2.1 and Proposition 3.2, and in order to show that $\lambda_1 < \lambda_2$ we must show that $M < \mu$, or equivalently $\|A\| < \mu$, where A is as in Lemma 2.2.

For each $x \in V$, let $i(x)$ denote its type so that $x \in \mathcal{O}_{i(x)}$. The fact that $(\gamma_1, \ldots, \gamma_n)$ is an eigenvalue of the quotient adjacency matrix with eigenvalue μ means that, for each $x \in V$,

$$\sum_{y:y\sim x} \gamma_{i(y)} = \mu\gamma_{i(x)}. \tag{3.2}$$

Now let $f \in l_2(V)$. Using (3.2) twice to substitute for μ, we obtain

$$\begin{aligned}
\mu^2\|f\|_2^2 &= \mu^2 \sum_{y\in V} (f(y))^2 \\
&\overset{(3.2)}{=} \mu \sum_{y\in V} \left(\sum_{x:x\sim y} \frac{\gamma_{i(x)}}{\gamma_{i(y)}} \right) (f(y))^2 \\
&= \mu \sum_{x\in V} \gamma_{i(x)} \sum_{y:y\sim x} \frac{1}{\gamma_{i(y)}} (f(y))^2 \\
&\overset{(3.2)}{=} \sum_{x\in V} \left(\sum_{z:z\sim x} \gamma_{i(z)} \right) \left(\sum_{y:y\sim x} \frac{(f(y))^2}{\gamma_{i(y)}} \right).
\end{aligned}$$

So we see that

$$\begin{aligned}
\mu^2\|f\|_2^2 - \|Af\|_2^2 &= \sum_{x\in V} \left[\left(\sum_{z:z\sim x} \gamma_{i(z)} \right) \left(\sum_{y:y\sim x} \frac{(f(y))^2}{\gamma_{i(y)}} \right) - \left(\sum_{y:y\sim x} f(y) \right)^2 \right] \\
&= \sum_{x\in V} \sum_{y:y\sim x} \sum_{z:z\sim x} \left[\frac{\gamma_{i(z)}}{\gamma_{i(y)}} (f(y))^2 - f(y)f(z) \right] \\
&= \sum_{x\in V} \sum_{\substack{\{y_1,y_2\}: \\ y_1,y_2\sim x}} \left[\frac{\gamma_{i(y_2)}}{\gamma_{i(y_1)}} (f(y_1))^2 + \frac{\gamma_{i(y_1)}}{\gamma_{i(y_2)}} (f(y_2))^2 - 2f(y_1)f(y_2) \right]
\end{aligned}$$

https://doi.org/10.1017/S0963548302005588 Published online by Cambridge University Press

$$= \sum_{x \in V} \sum_{\substack{\{y_1, y_2\}: \\ y_1, y_2 \sim x}} \gamma_{i(y_1)} \gamma_{i(y_2)} \left[\frac{f(y_1)}{\gamma_{i(y_1)}} - \frac{f(y_2)}{\gamma_{i(y_2)}} \right]^2$$

$$\geqslant \left(\min_i \gamma_i \right)^2 \sum_{xy \in E(G^2)} \left(\frac{f(x)}{\gamma_{i(x)}} - \frac{f(y)}{\gamma_{i(y)}} \right)^2, \tag{3.3}$$

where, in the last line, G^2 is the graph with the same vertex set as G and whose edge set, $E(G^2)$, consists of those pairs of distinct vertices, $\{x, y\}$, with the property that there is a path of length exactly 2 from x to y in G. It is easy to see that G^2 is also nonamenable and of bounded degree. Therefore, applying Theorem 2.6 to this graph and the function g given by $g(x) = f(x)/\gamma_{i(x)}$, we see that there is a constant $C > 0$, independent of f, so that the right-hand side of (3.3) exceeds $C\|g\|_2^2$; that is

$$\mu^2 \|f\|_2^2 - \|Af\|_2^2 \geqslant C\|g\|_2^2.$$

However, $\|g\|_2^2 \geqslant \|f\|_2^2/D$, where D is another constant, independent of f, namely $D = (\max_i \gamma_i)^2$. So we obtain, for some constant $k = k(G)$,

$$\mu^2 \|f\|_2^2 - \|Af\|_2^2 \geqslant k\|f\|_2^2,$$

which implies that $\|A\| \leqslant \sqrt{\mu^2 - k} < \mu$, thereby showing that $\lambda_1 < \lambda_2$.

The converse, namely that if G is nonamenable then $\lambda_1 = \lambda_2$, is much easier. In that case, for any $\epsilon > 0$, we can find a finite set of vertices A with $|\partial_E A|/|A| < \epsilon$. For such an A, define f by

$$f(x) = \begin{cases} \gamma_{i(x)} & \text{if } x \in A, \\ 0 & \text{otherwise.} \end{cases}$$

Now let $\mathrm{int}(A)$ consist of those vertices in A which do not have a neighbour outside A. Note that $|\mathrm{int}(A)|/|A| \geqslant 1 - \epsilon$. Equation (3.2) implies that if $x \in A$ then $Af(x) = \mu f(x)$. Therefore

$$\|Af\|_2^2 \geqslant \mu^2 \left(\|f\|_2^2 - \epsilon |A| \left(\max_i \gamma_i \right)^2 \right),$$

from which we deduce

$$\frac{\|Af\|_2^2}{\|f\|_2^2} \geqslant \mu^2 - \epsilon \left(\frac{\max_i \gamma_i}{\min_i \gamma_i} \right)^2. \tag{3.4}$$

Since the right-hand side of (3.4) can be made arbitrarily close to μ^2 by an appropriate choice of ϵ, we see that $\|A\| = \mu$, and hence $\lambda_1 = \lambda_2$. □

As remarked in the Introduction, the equivalent of Theorem 3.1 for the contact process remains a significant open problem. Neither direction is known, except for very special classes of graphs either explicitly defined (*e.g.*, $\mathbb{Z}^d$) or defined in terms of a certain concrete construction (the class of trees in [19]). Equivalent problems for other stochastic processes also remain open. For example, in percolation, it is widely believed (see, *e.g.*, [11]) that on transitive locally finite connected graphs, one has $p_c < p_u$ if and only if the graph

https://doi.org/10.1017/S0963548302005588 Published online by Cambridge University Press

is nonamenable; where p_c is the threshold for the existence of an infinite component and p_u is the threshold for the uniqueness of such components. A proof of Burton and Keane [2] (see also [6]) shows that amenability implies $p_c = p_u$, but the converse remains a celebrated and tantalizing open problem.

References

[1] Bezuidenhout, C. and Grimmett, G. (1990) The critical contact process dies out. *Ann. Probab.* **18** 1462–1482.
[2] Burton, R. M. and Keane, M. (1989) Density and uniqueness in percolation. *Comm. Math. Phys.* **121** 501–505.
[3] Cox, D. R. and Miller, H. D. (1965) *The Theory of Stochastic Processes*, Chapman and Hall, London.
[4] Durrett, R. (1988) *Lecture Notes on Particle Systems and Percolation*, Wadsworth, Belmont, CA.
[5] Gerl, P. (1988) Random walk on graphs with a strong isoperimetric property. *J. Theoret. Probab.* **1** 171–187.
[6] Grimmett, G. R. (1999) *Percolation*, 2nd edn, Springer, New York–Berlin.
[7] Harris, T. E. (1978) Additive set-valued Markov processes and graphical methods. *Ann. Probab.* **6** 355–378.
[8] Liggett, T. M. (1985) *Interacting Particle Systems*, Springer, New York–Berlin.
[9] Liggett, T. M. (1996) Multiple transition points for the contact process on the binary tree. *Ann. Probab.* **24** 1675–1710.
[10] Liggett, T. M. (1999) *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*, Springer, New York–Berlin.
[11] Lyons, R. (2000) Phase transitions on nonamenable graphs. *J. Math. Phys.* **41** 1099–1126.
[12] Madras, N. and Schinazi, R. (1992) Branching random walks on trees. *Stoch. Process. Appl.* **42** 255–267.
[13] Pemantle, R. (1992) The contact process on trees. *Ann. Probab.* **20** 2089–2116.
[14] Pemantle, R. and Stacey, A. M. (2001) The branching random walk and contact process on Galton–Watson and non-homogeneous trees. *Ann. Probab.* **29** 1563–1590.
[15] Peres, Y. Personal communication.
[16] Saloff-Coste, L. (1997) *Lectures on Finite Markov Chains*, Ecole d'Eté de Probabilités de Saint-Fleur XXVI-1996, *Lectures on Probability Theory and Statistics*, Springer, Berlin, pp. 301–413.
[17] Schinazi, R. (1993) On multiple phase transitions for branching Markov chains. *J. Statist. Phys.* **71** 521–525.
[18] Schonmann, R. H. (2001) Multiplicity of phase transitions and mean-field criticality on highly non-amenable graphs. *Comm. Math. Phys.* **219** 271–322.
[19] Stacey, A. M. (1996) The existence of an intermediate phase for the contact process on trees. *Ann. Probab.* **24** 1711–1726.

https://doi.org/10.1017/S0963548302005588 Published online by Cambridge University Press