Our objective of study today is the random walk algorithm for deciding if two vertices in an undirected graph are connected by a path. Last time we gave a heuristic argument that if G is a graph on n vertices and there is a path from s to t, the random walk algorithm will detect that t is reachable from s in at most $2n^3$ steps with probability $1/2$.

Today we give a rigorous analysis of the random walk algorithm (which will give a weaker but still polynomial bound on the length of the random walk). We then look at ways to remove the randomness from this algorithm and see how this can be achieved using some additional tools that can be obtained via the methods we'll see in the analysis. In the process we will see how to construct *expander graphs*, an important combinatorial tool that is useful in many scenarios in computer science.

As last time, we will assume that the graph G is connected, is d-regular, and has a loop around every vertex.

1 Adjacency matrix and eigenvalues

We now turn to some algebraic tools that will allow us to carry out a rigorous analysis of the random walk process. At this point it will be helpful to slightly change our perspective and look at the following question: If the particle starts at s, how long will it take the particle to reach not a specific s', but a random vertex in the graph?

To answer this question, it will be helpful to represent the random walk by a sequence of probability distributions $\mathbf{p}^0, \mathbf{p}^1, \ldots$ on the vertices of G, with the following interpretation: At each step t, $\mathbf{p}^t(u)$ is the probability of the particle ending up at vertex u after t steps of the walk. Initially, we have $\mathbf{p}^0$ assign probability 1 to vertex s, and probability 0 to all the other vertices. The distribution $\mathbf{p}^{t+1}$ can be calculated from $\mathbf{p}^t$ via the formula

$$\mathbf{p}^{t+1}(u) = \sum_{v:(v,u) \text{ is an edge}} \frac{1}{d} \cdot \mathbf{p}^t(v). \tag{1}$$

We are now interested in the following question: When t gets large, how close does the distribution $\mathbf{p}^t$ get to the uniform distribution $\mathbf{u}$ on the set of vertices? To answer this question, we need some way of measuring how "close" two distributions are. In our setting the most convenient measure is the ℓ_2 *norm*. The ℓ_2 norm of a vector $\mathbf{v}$ is the quantity

$$\|\mathbf{v}\| = \Big(\sum_i \mathbf{v}_i^2\Big)^{1/2}$$

and the ℓ_2 distance between two vectors $\mathbf{v}$ and $\mathbf{v}'$ is the ℓ_2 norm of $\mathbf{v} - \mathbf{v}'$. We will think of probability distributions as vectors in $\mathbb{R}^n$ (with one entry for each vertex in the graph), and we will say that two distributions $\mathbf{p}$ and $\mathbf{p}'$ are ϵ-close (in ℓ_2 distance) if $\|\mathbf{p} - \mathbf{p}'\| \leq \epsilon$.

The (normalized) *adjacency matrix* of G is an $n \times n$ matrix A defined as follows:

$$A_{u,v} = \frac{\text{number of edges between } u \text{ and } v \text{ in } G}{d}$$

This matrix is symmetric and the entries in each row add up to one. Using A, we can write equation 1 in matrix form as $\mathbf{p}^{t+1} = \mathbf{p}^t A$ (it is customary to represent $\mathbf{p}^t$ as row vectors) and so we immediately obtain that $\mathbf{p}^t = \mathbf{p}_0 A^t$.

It turns out that the eigenvalues and eigenvectors of A play a significant role in determining the behavior of random walks on G. Recall that an eigenvalue-eigenvector pair is a complex number λ and a vector $\mathbf{v}$ such that $\mathbf{v}A = \lambda\mathbf{v}$. It is a basic theorem in linear algebra that symmetric matrices have an orthonormal basis of eigenvectors with real eigenvalues. Let's denote these pairs by $(\lambda_1, \mathbf{v}_1), \ldots, (\lambda_n, \mathbf{v}_n)$ where $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n$. (Some of the λ_i may be negative.)

What is the meaning of this? Initially the position of our particle is determined by the distribution $\mathbf{p}^0$. Since the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ form an orthonormal basis we can decompose $\mathbf{p}^0$ in the form

$$\mathbf{p}^0 = \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n$$

where $\alpha_i = \langle \mathbf{p}^0, \mathbf{v}_i \rangle$ and $\alpha_1^2 + \cdots + \alpha_n^2 = 1$.

After one step of the random walk, the distribution becomes

$$\mathbf{p}^1 = \mathbf{p}^0 A = \alpha_1\mathbf{v}_1 A + \cdots + \alpha_n\mathbf{v}_n A = \alpha_1\lambda_1\mathbf{v}_1 + \cdots + \alpha_n\lambda_n\mathbf{v}_n$$

and after t steps

$$\mathbf{p}^t = \mathbf{p}^0 A^t = \alpha_1\lambda_1^t\mathbf{v}_1 + \cdots + \alpha_n\lambda_n^t\mathbf{v}_n. \tag{2}$$

Let's think of what happens when t becomes large. We will assume the values α_i are nonzero since the initial position of the particle can be arbitrary.[1] Eventually the right hand side of the expression will be dominated by the term in which λ_i has largest absolute value; this is either $|\lambda_1|$ or $|\lambda_n|$. This absolute value cannot exceed 1, because $\mathbf{p}^t$ would then become very large, but its norm is bounded since it is a probability distribution. Similarly, the absolute value cannot be less than 1 because then $\mathbf{p}^t$ would become very small when t gets large.

Therefore, it must be the case that $\lambda_1 = 1$, and

$$\max\{|\lambda_i| : 2 \leq i \leq n\} = \max(\lambda_2, -\lambda_n) \leq 1.$$

The quantity on the left side is denoted by $\lambda = \lambda(G)$ and plays a very important role because of the following. First, note that $\mathbf{u}A = \lambda_1\mathbf{u}$, so the eigenvector $\mathbf{v}_1$ associated to $\lambda_1 = 1$ equals $\sqrt{n} \cdot \mathbf{u}$. Now from (2) we have that

$$\|\mathbf{p}^t - \alpha_1\mathbf{v}_1\|^2 = \alpha_2^2\lambda_2^{2t} + \cdots + \alpha_n^2\lambda_n^{2t} \leq \lambda^{2t}.$$

The left hand side has a natural interpretation. Recall that $\alpha_1 = \langle \mathbf{p}^0, \mathbf{v}_1 \rangle = 1/\sqrt{n}$, so $\alpha_1\mathbf{v}_1$ equals the uniform distribution $\mathbf{u}$. Thus λ^t measures how close $\mathbf{p}^t$ gets to the uniform distribution after t steps of the walk: $\|\mathbf{p}^t - \mathbf{u}\| \leq \lambda^t$. Another way of saying this is that λ determines the *rate* at which $\mathbf{p}^t$ converges to the uniform distribution: The smaller λ is, the faster we will get to a uniformly random vertex.

[1] This is not quite right: The correct way to say it is that for every index i there exists an initial position for the particle that makes $\alpha_i \neq 0$.

2 Bounding the eigenvalue gap

We now introduce some tools that will allow us to upper bound the value $\lambda = \max(\lambda_2, -\lambda_n)$. Let us begin with λ_n.

Claim 1. *Suppose G is a d-regular graph with a loop around every vertex. Then $-\lambda_n \leq 1 - 2/d$.*

Proof. If G is d-regular with a loop around every vertex, we can write the adjacency matrix A of G as

$$A = \frac{d-1}{d} A' + \frac{1}{d} I$$

where A' is the adjacency matrix of the graph obtained by removing the loops from G, and I is the identity matrix. Notice that the eigenvalues of A and A' are then related by the formula

$$\lambda_i = \lambda'_i \cdot \frac{d-1}{d} + \frac{1}{d}$$

for $1 \leq i \leq n$, so in particular, $\lambda_n \geq (-1) \cdot (d-1)/d + 1/d \geq -1 + 2/d$. □

2.1 Rayleigh quotients

To bound λ_2, it will be useful to describe this eigenvalue in an alternative way. To see how this can be done, we diagonalize the matrix A as $S^{\mathrm{T}} \Lambda S$, where S is an orthonormal matrix whose rows are the eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$, and Λ is a diagonal matrix consisting of the entries $\lambda_1, \lambda_2, \ldots, \lambda_n$ in that order. (Recall that when S is orthonormal, $S^{\mathrm{T}} = S^{-1}$.) Then the eigenvectors of A and those of Λ are related by an orthonormal change of basis.

Let's now look at the eigenvalues of Λ, which are its diagonal entries. One way to describe the first eigenvalue λ_1 is to look at all possible vectors $\mathbf{v}$ of norm one, and take the one that maximizes the expression $\mathbf{v}\Lambda\mathbf{v}^{\mathrm{T}}$. Clearly this quantity is maximized by the vector $\mathbf{v} = \mathbf{e}_1 = (1, 0, \ldots, 0)$, which yields the value λ_1. But now notice that

$$\lambda_1 = \max_{\|\mathbf{v}\|=1} \mathbf{v}\Lambda\mathbf{v}^{\mathrm{T}} = \max_{\|\mathbf{v}\|=1} \mathbf{v}(SAS^{\mathrm{T}})\mathbf{v}^{\mathrm{T}} = \max_{\|\mathbf{v}\|=1} (\mathbf{v}S)A(\mathbf{v}S)^{\mathrm{T}} = \max_{\|\mathbf{v}\|=1} \mathbf{v}A\mathbf{v}^{\mathrm{T}}$$

because as $\mathbf{v}$ cycles over all vectors of norm one, so does $\mathbf{v}S$. Now notice that

$$\mathbf{v}A\mathbf{v}^{\mathrm{T}} = \sum_{u,w=1}^{n} A_{uw}\mathbf{v}(u)\mathbf{v}(w) = \sum_{(u,\,w) \text{ is an edge}} \frac{1}{d}\mathbf{v}(u)\mathbf{v}(w) = n\,\mathrm{E}_{(u,w)}[\mathbf{v}(u)\mathbf{v}(w)]$$

where $\mathrm{E}_{(u,w)}[\cdot]$ denotes expectation taken over a random *directed* edge (u, w) in G. This gives the following formula for λ_1:

$$\lambda_1 = n \cdot \max_{\|\mathbf{v}\|=1} \mathrm{E}_{(u,w)}[\mathbf{v}(u)\mathbf{v}(w)].$$

What about λ_2? Again, we look at Λ, but now instead of maximizing over all vectors, we only maximize over those that are orthogonal to the first vector $\mathbf{e}_1$:

$$\lambda_2 = \max_{\|\mathbf{v}\|=1, \mathbf{v}\perp\mathbf{e}_1} \mathbf{v}\Lambda\mathbf{v}^{\mathrm{T}} = \max_{\|\mathbf{v}\|=1, \mathbf{v}\perp\mathbf{e}_1} (\mathbf{v}S)A(\mathbf{v}S)^{\mathrm{T}}.$$

Now notice that as $\mathbf{v}$ cycles over all vectors of norm 1 that are perpendicular to $\mathbf{e}_1$, $\mathbf{v}S$ will cycle over all vectors of norm 1 that are perpendicular to $\mathbf{e}_1 S = \mathbf{v}_1$, which is parallel to $\mathbf{u}$. So we obtain the following expression for λ_2:

$$\lambda_2 = \max_{\|\mathbf{v}\|=1, \mathbf{v}\perp\mathbf{u}} \mathbf{v}A\mathbf{v}^{\mathrm{T}} = n \max_{\|\mathbf{v}\|=1, \mathbf{v}\perp\mathbf{u}} \mathrm{E}_{(u,w)}[\mathbf{v}(u)\mathbf{v}(w)].$$

Since our goal is to bound the value λ_2 away from 1, it will be convenient to look at the expression $1 - \lambda_2$. Using the above formula and simplifying a bit, we obtain

$$1 - \lambda_2 = \frac{n}{2} \min_{\|\mathbf{v}\|=1, \mathbf{v}\perp\mathbf{u}} \mathrm{E}_{(u,w)}[(\mathbf{v}(u) - \mathbf{v}(w))^2] \tag{3}$$

$$= \frac{1}{2} \min_{\mathbf{v}\perp\mathbf{u}} \frac{\mathrm{E}_{(u,w)}[(\mathbf{v}(u) - \mathbf{v}(w))^2]}{\mathrm{E}_u[\mathbf{v}(u)^2]}. \tag{4}$$

2.2 Bounds on the second eigenvalue

Looking at (3) we immediately see that $1 - \lambda_2 \geq 0$. Recall that our goal is to bound this value from below, so as a first step let's ask if $1 - \lambda_2$ can ever equal zero. Notice that if $1 - \lambda_2 = 0$, then there must exist a $\mathbf{v} \perp \mathbf{u}$ such that $\mathbf{v}(u) = \mathbf{v}(w)$ for every edge (u, w). Since G is connected, this implies that $\mathbf{v}$ is a constant vector, namely it is a multiple of $\mathbf{u}$. But since $\mathbf{v} \perp \mathbf{u}$, it follows that $\mathbf{v} = 0$, which is not allowed. So if G is connected, it must be that $\lambda_2 < 1$.

Our objective will now be to quantify this reasoning. In some sense, we will try to say that the better connected G is, the farther λ_2 will be from 1. To get a better understanding of $1 - \lambda_2$, it will be easier to start with an *upper bound* instead of a lower bound. To do this we define the *edge expansion* $h(G)$ of a graph G by the formula

$$h(G) = \min_{S:\, |S| \leq n/2} \Pr_{(u,w)}[w \notin S \mid u \in S]$$

where S ranges over all subsets of vertices of G of size at most $n/2$. This value tells us how likely we are to get out of a set S after one step of the random walk, if we start at a random vertex of S.

Theorem 2. $1 - \lambda_2 \leq 2h(G)$.

Proof. Let S be any set of vertices of size at most $n/2$ and set

$$\mathbf{v}(u) = \begin{cases} |\overline{S}|/n, & \text{if } u \in S \\ -|S|/n, & \text{if } u \notin S. \end{cases}$$

Notice that $\mathbf{v} \perp \mathbf{u}$, and that $\mathbf{v}(u) - \mathbf{v}(w)$ is 1 exactly when (u, w) exactly one of u and w is in S and the other is in $\overline{S}$, and 0 otherwise. In the first case we will say (u, w) *crosses* $(S, \overline{S})$. Plugging into (4) we obtain

$$1 - \lambda_2 \leq \frac{1}{2} \frac{\Pr_{(u,w)}[(u, w) \text{ crosses } (S, \overline{S})]}{\mathrm{E}_u[\mathbf{v}(u)^2]}.$$

Now notice that $\Pr_{(u,w)}[(u, w) \text{ crosses } (S, \overline{S})] = 2 \Pr[w \notin S \mid u \in S](|S|/n)$ and

$$\mathrm{E}_u[\mathbf{v}(u)]^2 = \frac{1}{n}\Big[|S| \cdot \Big(\frac{|\overline{S}|}{n}\Big)^2 + |\overline{S}| \cdot \Big(\frac{|S|}{n}\Big)^2\Big] = \frac{|S||\overline{S}|}{n^2} \geq |S|/2n. \qquad \square$$

This theorem tells us that if $1-\lambda_2$ is large, then the edge expansion is large. We will now see that the opposite relation holds as well: If the edge expansion is large, then so is $1-\lambda_2$.

Theorem 3 (Cheeger's inequality)**.** $1-\lambda_2 \geq h(G)^2/2$.

We now prove this theorem. Fix a $\mathbf{v}$ that minimizes (3). We may assume at most half (but at least one) of the entries of $\mathbf{v}$ are nonnegative, otherwise we can use $-\mathbf{v}$. Without loss of generality, let us assume that $\mathbf{v}(1) \geq \mathbf{v}(2) \geq \cdots \geq \mathbf{v}(n)$. The proof will show that if $\Pr_{(u,w)}[w \geq i+1 \mid u \in\leq i]$ is large for all i, then λ_2 must be large.

For somewhat obscure technical reasons, the negative entries of $\mathbf{v}$ are difficult to deal with, so we start by removing them.

Step 1: Remove the negative entries of v. Let $\mathbf{v}^+$ be the vector $\mathbf{v}^+(u) = \max(\mathbf{v}(u), 0)$. The first t entries of there vector are nonzero, where $1 \leq t \leq n/2$. We begin by showing that

$$1-\lambda_2 \geq \frac{1}{2} \cdot \frac{\mathrm{E}_{(u,w)}[(\mathbf{v}^+(u)-\mathbf{v}^+(w))^2]}{\mathrm{E}_u[\mathbf{v}^+(u)^2]}. \tag{5}$$

To prove this we write

$$\mathrm{E}_{(u,w)}[(\mathbf{v}^+(u)-\mathbf{v}^+(w))^2] = 2\,\mathrm{E}_u[\mathbf{v}^+(u)^2] - 2\,\mathrm{E}_{(u,w)}[\mathbf{v}^+(u)\mathbf{v}^+(w)].$$

By the definition of λ_2, for every u we have $\mathrm{E}_{(u,w)}[\mathbf{v}(w) \mid u] = \lambda_2 \mathbf{v}(u)$. Since $\mathbf{v}^+(w) \geq \mathbf{v}^+(u)$, it follows that $\mathrm{E}_{(u,w)}[\mathbf{v}^+(w) \mid u] \geq \lambda_2 \mathbf{v}(u)$, and since all $\mathbf{v}^+(w)$ are nonnegative, we get $\mathrm{E}_{(u,w)}[\mathbf{v}^+(w) \mid u] \geq \lambda_2 \mathbf{v}^+(u)$. Averaging over u, it follows that

$$\mathrm{E}_{(u,w)}[\mathbf{v}^+(u)\mathbf{v}^+(w)] \geq \lambda_2\,\mathrm{E}_u[\mathbf{v}^+(u)^2]$$

so $\mathrm{E}_{(u,w)}[(\mathbf{v}^+(u)-\mathbf{v}^+(w))^2] \leq (2-2\lambda_2)\,\mathrm{E}_u[\mathbf{v}^+(u)^2]$, proving (5).

Step 2: Cauchy-Schwarz. We now apply the Cauchy-Schwarz inequality to lower bound the numerator of (5):

$$\mathrm{E}_{(u,w)}[(\mathbf{v}^+(u)-\mathbf{v}^+(w))^2]\,\mathrm{E}_{(u,w)}[(\mathbf{v}^+(u)+\mathbf{v}^+(w))^2] \geq \mathrm{E}_{(u,w)}\left[|\mathbf{v}^+(u)^2-\mathbf{v}^+(w)^2|\right].$$

For the second term we have the upper bound:

$$\mathrm{E}_{(u,w)}[(\mathbf{v}^+(u)+\mathbf{v}^+(w))^2] \leq \mathrm{E}_{(u,w)}[2(\mathbf{v}^+(u)^2+\mathbf{v}^+(w)^2)] = 2\,\mathrm{E}_u[\mathbf{v}^+(u)^2],$$

so substituting in (5) we obtain:

$$\sqrt{1-\lambda_2} \geq \frac{1}{2} \cdot \frac{\mathrm{E}_{(u,w)}\left[|\mathbf{v}^+(u)^2-\mathbf{v}^+(w)^2|\right]}{\mathrm{E}_u[\mathbf{v}^+(u)^2]}. \tag{6}$$

Step 3: Change the order of summation. Order the vertices by decreasing values of $\mathbf{v}^+$, breaking ties arbitrarily. Direct the edges of the graph in a way compatible with this ordering and let $(u < w)$ denote the corresponding distribution on directed edges. Then

$$\begin{aligned}\mathrm{E}_{(u,w)}\left[|\mathbf{v}^+(u)^2 - \mathbf{v}^+(w)^2|\right] &= \mathrm{E}_{(u<w)}[\mathbf{v}^+(u)^2 - \mathbf{v}^+(w)^2] \\ &= \mathrm{E}_{(u<w)}\left[\sum_{i=w}^{u-1} \mathbf{v}^+(i)^2 - \mathbf{v}^+(i+1)^2\right] \\ &= \sum_{i=1}^{t} f_i \cdot \left(\mathbf{v}^+(i)^2 - \mathbf{v}^+(i+1)^2\right)\end{aligned}$$

Where $f_i = \Pr_{(u<w)}[u \le i \text{ and } w \ge i+1]$. Therefore $f_i = (i/n)\Pr[w \ge i+1 \mid u \le i]$. Finally,

$$\begin{aligned}\sum_{i=1}^{t} f_i \cdot \left(\mathbf{v}^+(i)^2 - \mathbf{v}^+(i+1)^2\right) &= \sum_{i=1}^{t} \Pr[w \ge i+1 \mid u \le i] \cdot i\left(\mathbf{v}^+(i)^2 - \mathbf{v}^+(i+1)^2\right)/n \\ &\ge \min_{i\in\{1,\ldots,t\}} \Pr[w \ge i+1 \mid u \le i] \cdot \sum_{i=1}^{t} i\left(\mathbf{v}^+(i)^2 - \mathbf{v}^+(i+1)^2\right)/n\end{aligned}$$

Now notice that the summation equals $(\mathbf{v}^+(1)^2+\mathbf{v}^+(2)^2+\cdots+\mathbf{v}^+(t)^2-t\mathbf{v}^+(t+1)^2)/n = \mathrm{E}_u[\mathbf{v}^+(u)]^2$, so putting everything in (6) we get:

$$\sqrt{1-\lambda_2} \ge \frac{1}{2} \cdot \min_{i\in\{1,\ldots,t\}} \Pr[w \ge i+1 \mid u \le i] \ge \frac{h(G)}{2}.$$

3 Analysis of the random walk algorithm

Using Theorem 3 it is now easy to analyze the random walk algorithm. Since G is connected, for every set S of size at most $n/2$ there must be at least one edge going out of S, so $h(G) \ge 1/dn$, and we get that $\lambda_2 \le 1-1/2(dn)^2$. Since every vertex has a loop, $\lambda_n \ge -1+2/d \ge -1+1/2(dn)^2$. Therefore after t steps of the random walk, we have that $\|\mathbf{p}^t - \mathbf{u}\| \le (1-1/2(dn)^2)^t$. For $t = \log(4n^2)/2(dn)^2$ we get that $\|\mathbf{p}^t - \mathbf{u}\| \le 1/4n^2$, so for every vertex v, $(\mathbf{p}^t(v) - \mathbf{u}(v))^2 \le 1/4n^2$, and $\mathbf{p}^t(v) \ge 1/2n$.

To recapitulate, after walking for t steps, for every vertex v, we have probability at least $1/2n$ of reaching it at time t. To reach the vertex s', consider what happens when we perform $4n$ such consecutive t step walks. Since in each walk we reach s' independently with probability at least $1/2n$, the probability that we fail to reach s' after all these walks is at most $(1-1/2n)^{4n} \le 1/2$. Therefore, after a total of $O(d^2n^3 \log n)$ steps, we have reached s' with probability at least $1/2$.

4 Highly connected graphs

In the beginning we asked the following question: Given a particle that sits at vertex s in the graph, how many steps of a random walk does it take for the particle to reach a random vertex in

the graph? We showed that for a d-regular, n vertex graph, we can always reach a random vertex after polynomially many (in n and d) steps.

However, in certain cases the number of steps can be quite large – as large as dn^2. Let us now reverse the question and ask the following: How can we design a d-regular graph G (think of d as very small compared to n) such that starting from *any* vertex s, we can reach a random vertex as soon as possible?

It seems a good idea to make as many vertices of G reachable using short walks out of s. This suggests that G should look like a tree rooted at s.

If we start as s, very quickly (after about $\log_d n$ steps) we will find the particle near the leaves of the tree. However, the particle is unlikely to stick at any particular leaf because there is only one path leading to it. A random walk on the tree favors the interior vertices, so the vertex at which the particle ends up won't look random.

In some sense, this is a bit unfair because the leaves have degree one, and the graph is not d-regular. We can "connect up" the leaves in some way so as to make the graph be d-regular. Once we do this, it seems plausible that after enough steps the vertex where the particle sits will indeed be uniform (and this is in fact the case), but also that a random vertex is reachable from s rather quickly (because in a tree, paths starting from s "expand out" very quickly).

However, in the end there is nothing special about s, and what we want in some way is that if we choose any vertex as the root, from the perspective of that vertex the graph looks a lot like a tree.

To be a bit more quantitative, if we start at s, even in the ideal case of a tree, we need just $\Omega(\log n)$ steps out of s to "cover" all the possible vertices in G. So we cannot hope to end up at a random vertex of G before we have completed at least $\Omega(\log n)$ steps. Can we do so in $O(\log n)$ steps no matter at which vertex s we started?

Recall that after t steps of the walk, our distance to the uniform distribution is upper bounded by the value λ^t, where $\lambda = \max(\lambda_2, -\lambda_n)$. This suggests that we want to design a graph whose value λ is as small as possible.

5 Expander graphs

The above discussion indicates that λ can in fact never get too small. To get a lower bound on λ, notice that after t steps of the walk, the potential number of vertices that could have been reached from s never exceeds d^{t+1}; there are at most this many vertices at distance $\leq t$ from s. So even when $t = \log_d n - 2$, less than half of the vertices of the random walk are reachable. Therefore the distribution $\mathbf{p}^t$ must assign probability zero to the other half vertices, and

$$\lambda^t \geq \|\mathbf{p}^t - \mathbf{u}\| \geq \left(n/2 \cdot (0 - 1/n)^2\right)^{1/2} = 1/\sqrt{2n}.$$

It follows that $\lambda = \Omega(1/\sqrt{d})$. A more precise analysis shows that for every graph, $\lambda \geq 2/\sqrt{d} - o_n(1)$, where $o_n(1)$ is quantity that converges to zero as n gets large. However, there exist graphs such

that $\lambda = 2/\sqrt{d}$ for infinitely many values of n. Such graphs are called *Ramanujan graphs*.[2]

For our purposes, it will be enough to consider graphs for which as n grows, λ stays bounded away from one. If this is the case, then after only $t = \Theta(\log n)$ steps of the random walk, we have that

$$\|\mathbf{p}^t - \mathbf{u}\| \leq \lambda^{\Theta(\log n)} = n^{-\Theta(1)} \tag{7}$$

so $\mathbf{p}^t$ gets very close to the uniform distribution, and in fact *all vertices* of G are reached with probability $\Theta(1/n)$.

Definition 4. A family of graphs $\{G_n\}$, where G_n has n vertices and is d-regular, is called an *expander family* if there is a constant $\epsilon > 0$ such that $\lambda(G_n) \leq 1 - \epsilon$ for every sufficiently large n.

For example, the following family of 4-regular graphs is an expander family: For every prime number p, think of the vertices of G_p as elements of the finite field $\mathbb{F}_p$. For every $x \in \mathbb{F}_p$, the edges going out of x are $(x, x-1)$, (x, x), $(x, x+1)$, (x, x^{-1}) where all operations are over $\mathbb{F}_p$ (and we set $0^{-1} = 0$).

Another example comes from random graphs: Suppose n is even and we construct a graph by taking a union of d random matchings on its vertices. It can be shown that with high probability the resulting family of graphs is expanding (and in fact has λ close to $2/\sqrt{d}$).

[2]Ramanujan graphs are known to exist for every n such that $n+1$ is a power of a prime larger than two.