CS 6822: Flows, Cuts, and Sparsifiers	**Lecture 2**
September 6, 2011	**Scribe: Bruno Abrahao**

In the last lecture, we showed that the random walk on the standard $2D$ lattice is recurrent. In that case, we got lucky and did not have to analyze the walk as we were able to explore a graph isomorphism. Above $2D$ the isomorphism does not hold and we need a different strategy to establish similar claims.

1 Proving transience for random walks on 3D lattices

In the last lecture we hinted at a non-rigorous way to show transience for random walks on 3D lattices by thinking of its behavior as three independent $1D$ random walks. Thus, we can write

$$E\left[\text{number of visits to } \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\right] = \Theta(\sum \frac{1}{s^{3/2}}) < \infty \tag{1}$$

Since this gives a finite number, this would explain transience. After the lecture a student offered a way to rigorously prove the above fact, which goes as follows.

Proof. Let us represent each step i of the walk on the $3D$ lattice as an ordered pair (d_i, s_i), where d_i is the i-th elements of a random **direction** sequence $Z, Z, X, Z, Y, X, ...$ which indicates a step either in the X, Y, or Z direction, and s_i the i-th element of a random **sign** sequence $-1, -1, +1, -1, ...$. Therefore, since each such ordered pair is equally likely, each of the six possible transitions are driven by (d_i, s_i) with d_i and s_i taken from the two independent sequences, one over a ternary and one over a binary alphabet.

At time t, let $i_X(t)$, $i_Y(t)$, and $i_Z(t)$ count the number of occurrences of X, Y, and Z symbols in the direction sequence. Notice that the X, Y, and Z subsequences of (d_i, s_i), defined as the sequences of signs s_i restricted to the pairs (d_i, s_i) where d_i is equal to one of the directions X, Y, or Z respectively, are mutually independent walks on $1D$. Thus, let P_t be the position of the walk at step t. We can use the counters to write

$$\Pr\left(P_t = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \middle| i_X(t), i_Y(t), i_Z(t)\right) = \begin{cases} \Theta((i_X(t) i_Y(t) i_Z(t))^{-1/2}) & \text{if } i_X(t), i_Y(t), i_Z(t) \text{ all even,} \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Thus the expect number of returns to $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, denoted by R_0 is

$$
\begin{aligned}
E[R_0] &= \Theta\left(\sum_{t=1}^{\infty}(i_X(t)i_Y(t)i_Z(t))^{-1/2} \times \Pr\left(i_X(t), i_Y(t), i_Z(t) \text{ are all even}\right)\right) &(3)\\
&\leq O\left(\sum_{t=1}^{\infty} E[(i_X(t)i_Y(t)i_Z(t))]^{-1/2}\right) &(4)
\end{aligned}
$$

Notice that there is a small but positive probability that one of the $i_X(t)$, $i_Y(t)$, or $i_Z(t)$ is equal to zero, in which case the corresponding term will contribute an infinite value to the expectation. In order to avoid this, let us assume without loss of generality that these terms are non-zero. For large values of t it is very likely that $i_X(t)$, $i_Y(t)$, and $i_Z(t)$ are all approximately equal to $t/3$. Therefore, ignoring the expectation operator, we can see that the last term reduces to $\sum_t t^{-3/2}$ which is finite. To make this precise, since this proof is tolerant of constant slack factors, we will instead say that for large t, $i_X(t)$, $i_Y(t)$, and $i_Z(t)$ are greater than or equal to $t/6$. Thus, we can break the above term up as

$$
\leq O\left(\sum_{t=1}^{\infty} \frac{t}{6}^{-3/2}\right) + O\left(\sum_{t=1}^{\infty} \Pr[(i_X(t) \wedge i_Y(t) \wedge i_Z(t) < \frac{t}{6})] \times t^{-1/2})\right) \qquad (5)
$$

where the first term covers all cases where t, $i_X(t)$, $i_Y(t)$, and $i_Z(t)$ are greater than or equal to $t/6$ and the second term is an error term. The error term can be shown to the exponentially small in t by Chernoff bound. ■

2 A more precise and more general proof of transience

The above proof is clumsy in the sense that it uses all the power of Chernoff bounds to prove something that is much weaker. In fact, we do not need an exponential tail for the error term.

Next we present a proof that is stronger than the above as it precisely computes the probability of the number of returns to the origin of a random walk for any $D > 2$. It also illustrate the theme that, if the problem has enough symmetry, it becomes easy to do exact calculations.

Theorem 2.1 *Random walks are transient on any lattice of dimension $D > 2$*

Proof. Recall from lecture 1 that

$$
E\left[\text{number of visits to } \vec{0} \,\middle|\, s_0 = \vec{0}\right] = \frac{1}{1 - \Pr\left(\text{return to } \vec{0} \,\middle|\, s_0 = \vec{0}\right)} \qquad (6)
$$

Our strategy is to compute the left hand side exactly so that we can solve for the probability of returning to $\vec{0}$, given that we started there. Let

$$N(\overrightarrow{x}) = E\left[\text{number of visits to } \overrightarrow{0} \middle| s_0 = \overrightarrow{x}\right] \tag{7}$$

We will see that it is going to be easy to calculate $N(\overrightarrow{x})$ for all $\overrightarrow{x}$ but on a different graph. Let us define the torus graph T^n as the graph where vertices are d-tuples of integers modulo $2n$, i.e., $(\mathbb{Z}_{/(2n)})^d$. Further, let us identify two specific vertices, namely $s = \overrightarrow{0}$ and $t = (n, 0, 0, ..., 0)^T$. We are going to consider a random walk starting at s, walking until it hits t for the first time, at which step it ends. For this particular random walk, define

$$N^n(\overrightarrow{x}) = E\left[\text{number of visits to } s \middle| s_0 = \overrightarrow{x}\right] \tag{8}$$

We will show that the expected number of visits to $\overrightarrow{0}$ in an infinite random walk is the $\lim_{n\to\infty}$ of the expected number of times $\overrightarrow{0}$ is hit in a random walk which stops when t is hit for the first time in the finite torus.

Claim 2.2 $N(\overrightarrow{0}) = \lim_{n\to\infty} N^n(\overrightarrow{0})$

Proof. Our strategy is to show that for every r, $N(\overrightarrow{0}) > r \iff \lim_{n\to\infty} N^n(\overrightarrow{0}) > r$. Let us relate the two terms by writing

$$N(\overrightarrow{0}) = \lim_{q\to\infty} E\left[\text{number of visits to } \overrightarrow{0} \text{ in the first } q \text{ steps}\right] \tag{9}$$

and

$$N^n(\overrightarrow{0}) = \lim_{q\to\infty} E\left[\text{number of visits to } \overrightarrow{0} \text{ in the first } q \text{ steps or before hitting } t\right] \tag{10}$$

The key observation is that the Expressions 9 and 10 are equal if $q < n$, because in this case there is no chance of hitting the target in q steps. Moreover in $q < n$ steps it is impossible to wrap around the torus.

$$\begin{aligned}
N^n(\overrightarrow{0}) &= \lim_{n\to\infty}\lim_{q\to\infty} E\left[\text{number of visits to } \overrightarrow{0} \text{ in the first } q \text{ steps or before hitting } t\right] \\
&= \lim_{q\to\infty}\lim_{n\to\infty} E\left[\text{number of visits to } \overrightarrow{0} \text{ in the first } q \text{ steps or before hitting } t\right] \\
&= \lim_{q\to\infty} E\left[\text{number of visits to } \overrightarrow{0} \text{ in the first } q \text{ steps or before hitting } t\right] \\
&= N(\overrightarrow{0})
\end{aligned}$$

■

The reader should convince themselves that the interchange of limits in the second line is justified. Also, notice that the random walk in the third line is always in the grid, and never sees it is on a torus. Once we are on the torus, everything becomes not only symmetrical, but also finite. Thus, we will be able to analyze the walks using linear algebra.

If the reader will recall, in the last lecture we wrote the expected number of returns to the source starting from x as a function of the average over the neighbors y of x of the expected number of returns to the source starting at y:

$$N^n(\overrightarrow{x}) = \begin{cases} \frac{1}{2d}\sum_{\overrightarrow{y}:(\overrightarrow{x},\overrightarrow{y})\in E} N^n(\overrightarrow{y}) & \text{if } \overrightarrow{x} \neq s,t \\ \frac{1}{2d}\sum_{\overrightarrow{y}:(\overrightarrow{x},\overrightarrow{y})\in E} N^n(\overrightarrow{y}) + 1 & \text{if } \overrightarrow{x} = s \\ \frac{1}{2d}\sum_{\overrightarrow{y}:(\overrightarrow{x},\overrightarrow{y})\in E} N^n(\overrightarrow{y}) - 1 & \text{if } \overrightarrow{x} = t \end{cases} \tag{11}$$

What justifies the first line of Equation 11 is that if $\overrightarrow{x} \neq s,t$, then, when the random walk starts in $\overrightarrow{x}$, it will not be at s at time zero. Therefore, the expected number of visits to s is the same as taking a step in the random walk and then counting the number of visits to s took place after that initial step. For the second line, we count one visit to s, since we started there, and then once we take one step, we average the expected number over the walks starting at the neighbors. The third line is justified by the following equation:

$$\sum_{\overrightarrow{x}} N^n(\overrightarrow{x}) = \frac{1}{2d}\sum_{\overrightarrow{x}}\sum_{\overrightarrow{y}:(\overrightarrow{x},\overrightarrow{y})\in E} N^n(\overrightarrow{y}) \tag{12}$$

Equation 12 is valid because every $\overrightarrow{y}$ has exactly $2d$ neighbors in a d-dimensional torus. Therefore an $N^n(\overrightarrow{y})$ appears exactly $2d$ times in the double sum, which cancels out with the $\frac{1}{2d}$ term. Given that the first two lines are valid, the third must also be in order for this Equation 12 to hold.

An obvious fact is that $N^n(\overrightarrow{t}) = 0$. If the random walk starts at t, it stops and will not visit s. So, by the third line of Equation 11, we have

$$1 = \frac{1}{2d}\sum_{\overrightarrow{y}:(t,\overrightarrow{y})\in E} N^n(\overrightarrow{y}) \tag{13}$$

Now, t has exactly $2d$ neighbors and, since the graph is symmetrical under permuting the coordinates, which permutes the neighbors of t, it is the case that each term in the sum, on the right hand side of Equation 13 has to be individually equal to 1. Thus, starting a random walk on a d-dimensional torus at any neighbor of t, the walk is going to visit s once in expectation! A conceptual (rather than algebraic) justification of this fact is non-obvious.

In order to solve the system of linear equations given by Equation 11, we will define its terms using the Laplacian matrix.

Definition 2.3 *For any weighted graph $G = (V, E)$, the Laplacian matrix L_G is given by*

$$(L_G)_{uv} = \begin{cases} -W_{uv} & \textit{if } u \neq v \textit{ and } (u,v) \in E \\ \sum_{x:(x,u)\in E} W_{ux} & \textit{if } u = v \\ 0 & \textit{otherwise} \end{cases} \tag{14}$$

Definition 2.4 *For any weighted graph $G = (V, E)$, the Normalized Laplacian matrix is $\hat{L}_G = D^{-1}L_G$ where D is a diagonal matrix whose (u,u) entry is $(L_G)_{uu}$.*

For the torus T^n with $W_{uv} = 1$ if $(u, v) \in E$ and 0 otherwise, its Laplacian matrix will have $2d$ on the diagonal, and the other non-zero entries have value -1 in the pattern of the adjacent matrix of T^n. Its Normalized Laplacian matrix is simply $\hat{L}_{T^n} = \frac{1}{2d} L_{T^n}$.

Let us now think about the system of linear equations in Equation 11 using Laplacian matrix. In particular, we have that $\hat{L}_{T^n} N^n = \Delta_{st}$ where

$$\Delta_{st}(u) = \begin{cases} 1 & \text{if } u = s \\ -1 & \text{if } u = t \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

Our plan will be to work in a different basis where $\hat{L}_{T^n}$ is diagonal. The reason we will be able to do that is that the torus has enough symmetry that we will be able to guess its eigenvectors. To start with, let us reason about one of the eigenvectors of T^n.

Define $h = (2n)^{d/2} = \sqrt{|V(T^n)|}$ (where $V(G)$ denotes the set of nodes of G). Further define the primitive root of unit $w = e^{\pi/n}$ so that $w^{2n} = 1$. We are now ready to define the *Fourier basis* which is given by the vectors ξ_x whose y component is

$$\xi_x(y) = h^{-1} w^{<x,y>} \tag{16}$$

where $< x, y >$ is the inner product[1]. Now, x and y are vertices of the torus and, therefore, integers between 1 and $2n$. Thus, their dot product is well defined as an integer module $2n$ but it is not a well defined integer. A possible problem is that we are raising w to a power which is not an actual number but an equivalence class of an integer modulo $2n$. Nevertheless, this does not lead to problems as $w^{2n} = 1$, and, thus, whichever representative of the equivalence class we choose, we get the same number. Here are some useful facts about these vectors.

Property 2.5 $< \xi_x, \xi_x >= 1$.

$$\begin{aligned} < \xi_x, \xi_x > &= \sum_y \overline{\xi_x(y)} \xi_x(y) && (17) \\ &= h^{-2} \sum_y (\overline{w} w)^{<x,y>} && (18) \\ &= h^{-2} \sum_y 1 = 1 && (19) \end{aligned}$$

Property 2.6 *The vectors in the Fourier basis are mutually orthogonal:* $< \xi_x, \xi_y >= 0$ *if* $x \neq y$.

[1] Notice that we are indicating the particular component of vectors within parens, instead of using the traditional subscript notation. This is because we want to save the subscript notation for indexing the eigenvalues.

$$h^{-2}\sum_z \overline{w}^{<x,z>} w^{<y,z>} = h^{-2}\sum_z w^{<y-x,z>} \tag{20}$$

$$= h^{-2}\prod_{i=1}^{d}\sum_{z_i=1}^{2n-1} w^{(y_i-x_i)z_i} \tag{21}$$

As long as at least one of the $y_i = x_i$ are non-zero, we can use the formula for the geometric progression:

$$\sum_{z_i=1}^{2n-1} w^{(y_i-x_i)z_i} = \frac{w^{2n(y_i-x_i)}-1}{w^{y_i-x_i}-1} = 0 \tag{22}$$

Given the above properties, we have shown that ξ_x is an orthonormal basis. Every other vector η satisfies

$$\eta = \sum_x < \eta, \xi_x > \xi_x \tag{23}$$

The reason why we have picked this basis is that these vectors are the eigenvectors of the Normalized Laplacian. We now proceed to show that this is the case and compute the eigenvalues. What is $\hat{L}_{T^n}\xi_x$?

$$(\hat{L}_{T^n}\xi_x)(y) = \sum_z (\hat{L}_{T^n})_{yz}\xi_x(z) \tag{24}$$

$$= \sum_{z:(y,z)\in E} \frac{\xi_x(y)-\xi_x(z)}{2d} \tag{25}$$

$$= \frac{1}{2d}\sum_{i=1}^{d}[1-w^{x_i}+1-w^{-x_i}]\xi_x(y) \tag{26}$$

$$= \frac{1}{d}\sum_{i=1}^{d}\left[1-\cos(\frac{\pi x_i}{n})\right]\xi_x(y) \tag{27}$$

The second line of Equation 24 is justified as the Nomalized Laplacian has a $\frac{-1}{2d}$ in the yz component if y is adjacent to z. Thus, it has $\frac{-1}{2d}\xi_x(z)$ for every neighbor of z, it has a 1 in the yy component, so we have $\xi_x(y)$ appearing $2d$ times, which is why we are dividing by $2d$. In the third line, we grouped the neighbors in pairs, the positive neighbor in the i direction and the negative neighbor in the i direction. Thus, if we move from y to its neighbor in the positive i direction, then $\xi_x(z) = \xi_x(y)w^{x_i}$, and if we move in the negative direction, we multiply it by w^{-x_i}. The fourth line uses the fact that $\frac{w^{x_i}+w^{-x_i}}{2} = \cos(\frac{2\pi x_i}{2n})$. Notice that the only part that depends on y in the resulting term is $\xi_x(y)$. The other part of the term is the eigenvalue, which we will refer heretofore as λ_x.

Now we have a complete description of the Laplacian in this eigenbasis. It is a diagonal matrix whose entries are given by the above expression. This makes it easy to solve the linear system, because all we need is to invert each of the diagonal entries, and that gives us the inverse. Let us go back to compute the right hand side of the linear system, namely Δ_{st}, in terms of the Fourier basis.

$$\Delta_{st} = \sum_x < \Delta_{st}, \xi_x > \xi_x \tag{28}$$

$$\begin{aligned} < \Delta_{st}, \xi_x > &= \sum_y \overline{\Delta_{st}(y)} \xi_x(y) & (29) \\ &= \xi_x(s) - \xi_x(t) & (30) \\ &= \left[\frac{1}{n} - \frac{1}{n}(-1)^{x_1} \right] & (31) \\ &= h^{-1}[1 - (-1)^{x_1}] & (32) \\ &= \begin{cases} 2h^{-1} & \text{for odd } x_1 \\ 0 & \text{for even } x_1 \end{cases} & (33) \end{aligned}$$

In the second line, s is the all zero vector, so $\xi_x(s) = 1/h$ and $\xi_x(t) = \frac{1}{h} w^{<x,t>} = \frac{1}{h} w^{n x_1}$, because the only non-zero component of t is the first. Thus, we get

$$\begin{aligned} < N^n, \xi_x > &= \lambda_x^{-1} < \Delta_{st}, \xi_x > & (34) \\ &= \begin{cases} 2\lambda_x^{-1} h^{-1} & \text{for odd } x_1 \\ 0 & \text{for even } x_1 \end{cases} & (35) \end{aligned}$$

In order to get back to the non-Fourier basis

$$\begin{aligned} N^n &= \sum < N^n, \xi_x > \xi_x & (36) \\ N^n(s) &= \sum < N^n, \xi_x > \xi_x(s) & (37) \\ &= \frac{1}{h} \sum_x \begin{cases} 2\lambda_x^{-1} h^{-1} & \text{for odd } x_1 \\ 0 & \text{for even } x_1 \end{cases} & (38) \\ &= \frac{2}{h^2} \sum_{x (x_i \text{ odd})} \lambda_x^{-1} & (39) \end{aligned}$$

If we draw the entire torus, and put an X in the points that contribute to the sum and an O in those which do not, we get a a grid that uniformly covers the torus, skipping all the even numbered coordinates in the x_1 direction and including all the odd numbered ones. The total number of summands is half of the number of vertices of the torus, $h^2/2$. Thus, the term $\frac{2}{h^2}$ has the effect of averaging the terms of the sum, and, therefore, this above term is averaging λ_x^{-1} at the points labeled X in our construction. This means that this is a Riemman sum approximating the value of that function over the entire torus.

As $n \to \infty$, the sum converges to

$$\int_0^1 ... \int_0^1 \left[\frac{1}{d} \sum_{i=1}^{d} (1 - \cos(2\pi x_i)) \right]^{-1} dx_1 dx_2 ... dx_d \tag{40}$$

Now we have an exact formula for $N^n(s)$ for $\lim_{n\to\infty}$. If random walks are recurrent in $2D$ and transient in $3D$, it must be that this integral is divergent when one integrates over 2 variables, but finite when one integrates over 3 or more.

The integrand is bounded, except near 0, because $1 - \cos(2\pi x)$ is either 0 or a positive number. It is 0 when x is either 0 or 1 (on a torus $0 = 1$). Therefore, the sum is strictly positive, except at 0 where we have a division by 0, which blows the integrand up to infinity. Near $x = 0$, the term $1 - \cos(2\pi x) = \Theta(x^2)$, by its Taylor expansion. Thus, to determine whether the integral converges or diverges is equivalent to look at the convergence or divergence of $\int ||x||_2^{-2} dx_1 ... dx_d$.

If we write it in polar coordinates, in a neighborhood of 0 we have

$$\begin{aligned} \int ||x||_2^{-2} dx_1 ... dx_d &= \int_0^{\epsilon} r^{-2} (r^{d-1} Vol(S^{d-1}) dr d\theta &(41)\\ &= c \int_0^{\epsilon} r^{d-3} dr &(42)\\ &= \begin{cases} < \infty & \text{for } d \geq 3 \\ \infty & \text{for } d = 1, 2 \end{cases} &(43) \end{aligned}$$

where $Vol(S^{d-1})$ is the volume of the $(d-1)$-dimensional sphere, which is a constant c (used in the second line) that does not depend on r. ■