Recurrence vs Transience: An introduction to random walks

Pablo Lessa*

May 24, 2015

Preface

These notes are aimed at advanced undergraduate students of Mathematics. Their purpose is to provide a motivation for the study of random walks in a wide variety of contexts. I have chosen to focus on the problem of determining recurrence or transience of the simple random walk on infinite graphs, especially trees and Cayley graphs (associated to a symmetric finite generator of a discrete group).

None of the results here are new and even less of them are due to me. Except for the proofs the style is informal. I've used contractions and many times avoided the use of the editorial we. The bibliographical references are historically incomplete. I give some subjective opinions, unreliable anecdotes, and spotty historical treatment. This is not a survey nor a text for experts.

This work is dedicated to François Ledrappier.

Contents

*Partially supported by /CNPq/-Brazil project number 407129/2013-8.

1 Introduction

1.1 Pólya's Theorem

A simple random walk, or drunkard's walk, on a graph, is a random path obtained by starting at a vertex of the graph and choosing a random neighbor at each step. The theory of random walks on finite graphs is rich and interesting, having to do with diversions such as card games and magic tricks, and also being involved in the construction and analysis of algorithms such as the PageRank algorithm which at one point in the late 90s was an important part of the Google search engines.

In these notes we will be interested in walks on infinite graphs. I believe it's only a mild exaggeration to claim that the theory was born in 1920s with a single theorem due to George Pólya.

Theorem 1 (Pólya's Theorem)**.** *The simple random walk on the two dimensional grid is recurrent but on the three dimensional grid it is transient.*

Here the two dimensional grid $\mathbb{Z}^2$ is the graph whose vertices are pairs of integers and where undirected edges are added horizontally and vertically so that each vertex has 4 neighbors. The three dimensional grid $\mathbb{Z}^3$ is defined similarly, each vertex has 6 neighbors.

A simple random walk is called recurrent if almost surely (i.e. with probability 1) it visits every vertex of the graph infinitely many times. The reader can probably easily convince him or herself that such paths exist. The content of the first part of Pólya's theorem is that a path chosen at random on $\mathbb{Z}^2$ will almost surely have this property (in other words the drunkard always makes it home eventually).

Transience is the opposite of recurrence. It is equivalent to the property that given enough time the walk will eventually escape any finite set of vertices never to return. Hence it is somewhat counterintuitive that the simple random walk on $\mathbb{Z}^3$ is transient but its shadow or projection onto $\mathbb{Z}^2$ is recurrent.

1.2 The theory of random walks

Starting with Pólya's theorem one can say perhaps that the theory of random walks is concerned with formalizing and answering the following question: What is the relationship between the behavior of a random walk and the geometry of the underlying space?

Since it is possible for a drunkard to walk on almost any mathematical structure the theory has rich interactions with various parts of math. Meaningful and interesting answers to this question have been obtained in a wide variety of contexts ranging from random matrix products to Brownian motion on Riemannian manifolds.

We can illustrate this by looking at the most obvious generalization of the context of Pólya's theorem, which is simply to consider in place of $\mathbb{Z}^2$ and $\mathbb{Z}^3$ the simple random walk on the d-dimensional grid $\mathbb{Z}^d$ for $d = 1, 2, 3, \ldots$.

Pólya's theorem can be rephrased as follows: Suppose two random walkers (let's say lovers) begin at different vertices of $\mathbb{Z}^d$. If $d \leq 2$ then they will almost surely eventually meet. However, if $d \geq 3$ then there is a positive probability that they will never do so (how sad).

In spite of the fact that the two lovers might never meet on $\mathbb{Z}^3$, it is known that almost surely each lover will at some point be at the same vertex as the other lover was some time before (and hence be able to smell their perfume, there is hope!). Let's summarize the situation by saying that $\mathbb{Z}^3$ has the "Perfume property". A gem from the theory of random walks on $\mathbb{Z}^d$ (due to Erdös and Taylor) is the following.

Theorem 2 (The perfume theorem). *The d-dimensional grid has the perfume property if and only if $d \leq 4$.*

Brownian motion on $\mathbb{R}^d$ is a scaling limit of the simple random walk on $\mathbb{Z}^d$ (in a precise sense given by Donsker's theorem, which is a generalization of the central limit theorem). Hence one expects Brownian paths to have analogous behaviour to those of the simple random walk in the same dimension.

A theorem of Dvoretsky, Erdös, and Kakutani implies that two Brownian paths in $\mathbb{R}^4$ almost surely do not intersect (so the strict analog of the perfume

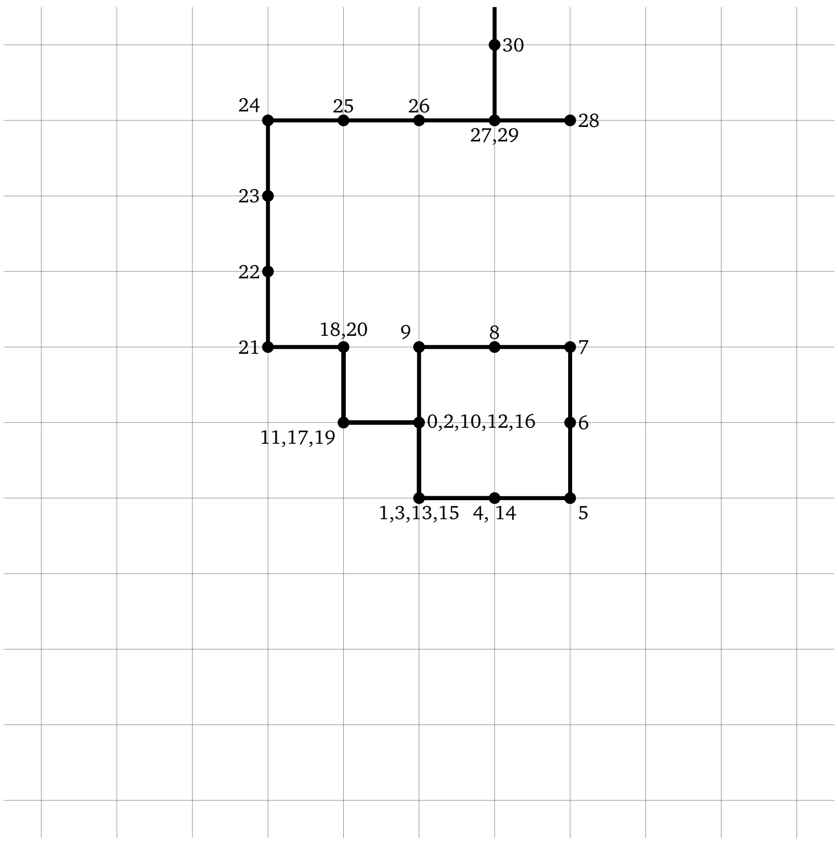

Figure 1: Bêbado
I generated this by flipping a brazilian “1 real” coin twice for each step. The result was almost to good of an illustration of Pólya’s theorem, returning to the origin 4 times before leaving the box from $(-5,-5)$ to $(5,5)$ after exactly 30 steps.

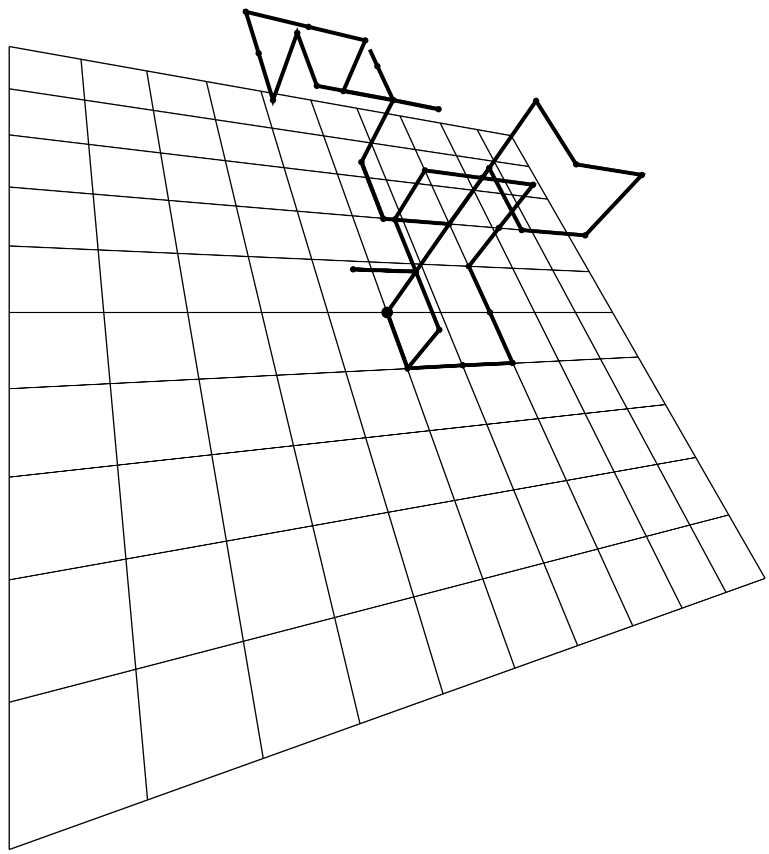

Figure 2: If you drink don't space walk.

More coin flipping. I kept the previous 2d walk and added a 3rd coordinate, flipping coins so that this third coordinate would change one third of the time on average (if you have to know, if both coins landed heads I would flip one again to see whether to increment or decrement the new coordinate, if both landed tails I would just ignore that toss, in any other case I would keep the same third coordinate and advance to the next step in the original 2d walk). The result is a 3d simple random walk. Clearly it returns to the origin a lot less than its 2d shadow.

property does not hold in $\mathbb{R}^4$). However, the two paths will pass arbitrarily close to one another so the perfume property does hold in a slightly weaker sense.

Brownian motion makes sense on Riemannian manifolds (basically because a list of instructions of the type "forward 1 meter, turn right 90 degrees, forward 1 meter, etc" can be followed on any Riemannian surface, this idea is formalized by the so-called Eells-Ellsworthy-Malliavin construction of Brownian motion) so a natural object to study is Brownian motion on the homogeneous surface geometries (spherical, Euclidean, and hyperbolic). A beautiful result (due to Jean-Jaques Prat) in this context is the following:

Theorem 3 (Hyperbolic Brownian motion is transient)**.** *Brownian motion on the hyperbolic plane escapes to infinity at unit speed and has an asymptotic direction.*

The hyperbolic plane admits polar coordinates (r, θ) with respect to any chosen base point. Hence Brownian motion can be described as a random curve (r_t, θ_t) indexed on $t \geq 0$. Prat's result is that $r_t/t \to 1$ and the limit $\theta_\infty = \lim \theta_t$ exists almost surely (and in fact $e^{i\theta_\infty}$ is necessarily uniform on the unit circle by symmetry). This result is very interesting because it shows that the behaviour of Brownian motion can change drastically even if the dimension of the underlying space stays the same (i.e. curvature affects the behavior of Brownian motion).

Another type of random walk is obtained by taking two 2×2 invertible matrices A and B and letting $g_1, \ldots, g_n, \ldots$ be independent and equal to either A or B with probability 1/2 in each case. It was shown by Furstenberg and Kesten that the exponential growth of the norm of the product $A_n = g_1 \cdots g_n$ exists, i.e. $\chi = \lim \frac{1}{n} \log(|A_n|)$ (this number is called the first Lyapunov exponent of the sequence, incredibly this result is a trivial corollary of a general theorem of Kingman proved only a few years later using an almost completely disjoint set of ideas).

The sequence A_n can be seen as a random walk on the group of invertible matrices. There are three easy examples where $\chi = 0$. First, one can take both A and B to be rotation matrices (in this case one will even have recurrence in the following weak sense: A_n will pass arbitrarily close to the identity matrix). Second, one can take $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $B = A^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$. And third, one can take $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix}$. It is also clear that conjugation by a matrix C (i.e. changing A and B to $C^{-1}AC$ and $C^{-1}BC$ respectively) doesn't change the behavior of the walk. A beautiful result of Furstenberg (which has many generalizations) is the following.

Theorem 4 (Furstenberg's exponent theorem)**.** *If a sequence of independent random matrix products A_n as above has $\chi = 0$ then either both matrices A and B are conjugate to rotations with respect to the same conjugacy matrix, or they both fix a one-dimensional subspace of $\mathbb{R}^2$, or they both leave invariant a union of two one-dimensional subspaces.*

1.3 Recurrence vs Transience: What these notes are about

From the previous subsection the reader might imagine that the theory of random walks is already too vast to be covered in three lectures. Hence, we concentrate on a single question: Recurrence vs Transience. That is we strive to answer the following:

Question 1. *On which infinite graphs is the simple random walk recurrent?*

Even this is too general (though we will obtain a sharp criterion, the Flow Theorem, which is useful in many concrete instances). So we restrict even further to just two classes of graphs: Trees and Cayley graphs (these two families, plus the family of planar graphs which we will not discuss, have recieved special attention in the literature because special types of arguments are available for them, see the excellent recent book by Russell Lyons and Yuval Perez "Probability on trees and networks" for more information).

A tree is simply a connected graph which has no non-trivial closed paths (a trivial closed path is the concatenation of a path and its reversal). Here are some examples.

Example 1 (A regular tree)**.** *Consider the regular tree of degree three* $\mathbb{T}_3$ *(i.e. the only connected tree for which every vertex has exactly 3 neighbors). The random walk on* $\mathbb{T}_3$ *is clearly transient. Can you give a proof?*

Example 2 (The Canopy tree)**.** *The Canopy tree is an infinite tree seen from the canopy (It's branches all the way down!). It can be constructed as follows:*

1. *There is a "leaf vertex" for each natural number* $n = 1, 2, 3, \ldots$*.*
2. *The leaf vertices are split into consecutive pairs* $(1, 2), (3, 4), \ldots$ *and each pair is joined to a new "level two" vertex* $v_{1,2}, v_{3,4}, \ldots$*.*
3. *The level two vertices are split into consecutive pairs and joined to level three vertices and so on and so forth.*

Is the random walk on the Canopy tree recurrent?

Example 3 (Radially symmetric trees)**.** *Any sequence of natural numbers* $a_1, a_2, \ldots$ *defines a radially symmetric tree, which is simply a tree with a root vertex having* a_1 *children, each of which have* a_2 *children, each of which have* a_3 *children, etc. Two simple examples are obtained by taking* a_n *constant equal to* 1 *(in which case we get half of* $\mathbb{Z}$ *on which the simple walk is recurrent) or* 2 *(in which case we get an infinite binary tree where the simple random walk is transient). More interesting examples are obtained using sequences where all terms are either* 1 *or* 2 *but where both numbers appear infinitely many times. It turns out that such sequences can define both recurrent and transient trees (see Corollary 1).*

Example 4 (Self-similar trees)**.** *Take a finite tree with a distinguished root vertex. At each leaf attach another copy of the tree (the root vertex replacing the leaf). Repeat ad infinitum. That's a self-similar tree. A trivial example* *is obtained when the finite tree used has only one branch (the resulting tree is half of $\mathbb{Z}$ and therefore is recurrent). Are any other self-similar trees recurrent? A generalization of this construction (introduced by Woess and Nagnibeda) is to have n rooted finite trees whose leaves are labeled with the numbers 1 to n. Starting with one of them one attaches a copy of the k-th tree to each leaf labeled k, and so on ad infinitum.*

Given any tree one can always obtain another by subdividing a few edges. That is replacing an edge by a chain of a finite number of vertices (this concept of subdivision appears for example in the characterization of planar graphs, a graph is planar if and only if it doesn't contain a subdivision of the complete graph in 5 verteces or the 3 houses connected to electricity, water and telephone graph[1]). For example a radially symmetric tree defined by a sequence of ones and twos (both numbers appearing infinitely many times) is a subdivision of the infinite binary tree.

Question 2. *Can one make a transient tree recurrent by subdividing its edges?*

Besides trees we will be considering the class of Cayley graphs which is obtained by replacing addition on $\mathbb{Z}^d$ with a non-commutative operation. In general, given a finite symmetric generator F of a group G (i.e. $g \in F$ implies $g^{-1} \in F$, an example is $F = \{(\pm 1, 0), (0, \pm 1)\}$ and $G = \mathbb{Z}^2$) the Cayley graph associated to (G, F) has vertex set G and an undirected edge is added between x and y if and only if $x = yg$ for some $g \in F$ (notice that this relationship is symmetric).

Let's see some examples.

Example 5 (The free group in two generators)**.** *The free group $\mathbb{F}_2$ in two generators is the set of finite words in the letters N, W, E, W (north, south, east, and west) considered up to equivalence with respect to the relations $NS = SN = EW = WE = e$ (where e is the empty word). It's important to note that these are the only relations (for example $NE \neq EN$). The Cayley graph of $\mathbb{F}_2$ (we will always consider the generating set $\{N, W, E, W\}$) is a regular tree where each vertex has 4 neighbors.*

Example 6 (The modular group)**.** *The modular group Γ is the group of fractional linear transformations of the form*

$$g(z) = \frac{az + b}{cz + d}$$

where $a, b, c, d \in \mathbb{Z}$ and $ad - bc = 1$. We will always consider the generating set $F = \{z \mapsto z + 1, z \mapsto z - 1, z \mapsto -1/z\}$. Is the simple random walk on the corresponding Cayley graph recurrent or transient?

Example 7 (A wallpaper group)**.** *The wallpaper group $*632$ is a group of isometries of the Euclidean plane. To construct it consider a regular hexagon in the*

[1] A more standard name might be the $(3, 3)$ bipartite graph.

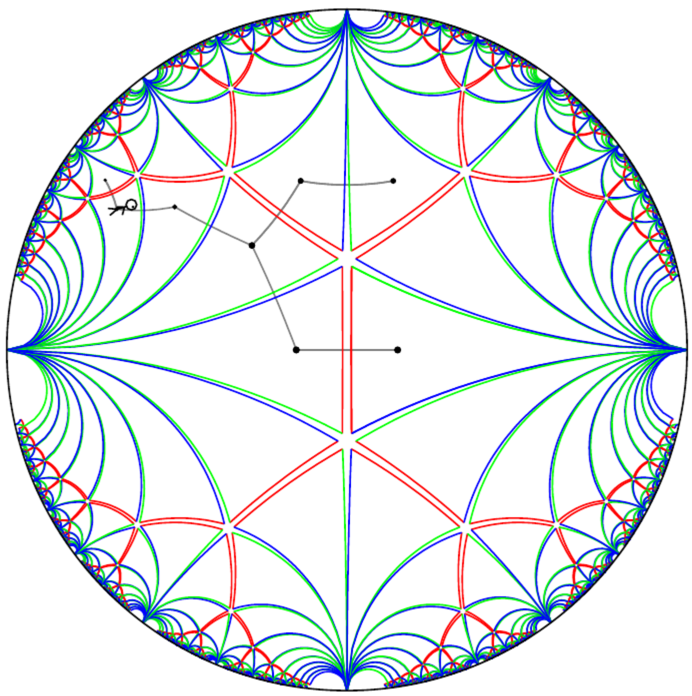

Figure 3: Random walk on the modular group
Starting at the triangle to the right of the center and choosing to go through either the red, green or blue side of the triangle one is currently at, one obtains a random walk on the modular group. Here the sequence red-blue-red-blue-green-red-blue-blue-red-red is illustrated.

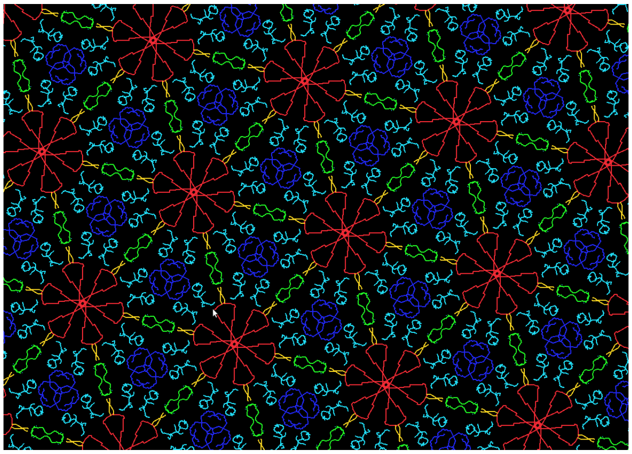

Figure 4: Wallpaper

I made this wallpaper with the program Morenaments by Martin von Gagern. It allows you to draw while applying a group of Euclidean transformations to anything you do. For this picture I chose the group $*632$.

plane and let F be the set of 12 axial symmetries with respect to the 6 sides, 3 diagonals (lines joining opposite vertices) and 3 lines joining the midpoints of opposite sides. The group $*632$ is generated by F and each element of it preserves a tiling of the plane by regular hexagons. The strange name is a case of Conway notation and refers to the fact that 6 axii of symmetry pass through the center of each hexagon in this tiling, 3 pass through each vertex, and 2 through the midpoint of each side (the lack of asterisk would indicate rotational instead of axial symmetry). Is the simple random walk on the Cayley graph of $*632$ recurrent?

Example 8 (The Heisenberg group)**.** *The Heisenberg group or Nilpotent group Nil is the group of 3×3 matrices of the form*

$$g = \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}$$

where $x, y, z \in \mathbb{Z}$. We consider the generating set F with 6 elements defined one of x, y or b being ± 1 and the other two being 0. The vertex set of the Cayley graph can be identified with $\mathbb{Z}^3$. Can you picture the edges? The late great Bill Thurston once gave a talk in Paris (it was the conference organized with the money that Perelman refused, but that's another story) where he said that Nil was a design for an efficient highway system . Several years later I understood what he meant (I think). How many different vertices can you get to from a given vertex using only n steps?

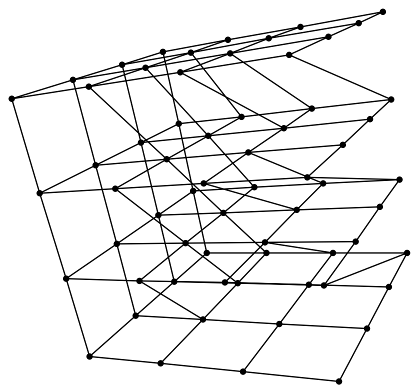

Figure 5: Heisenberg
A portion of the Cayley graph of the Heisenberg group.

Example 9 (The Lamplighter group)**.** *Imagine $\mathbb{Z}$ with a lamp at each vertex. Only a finite number of lamps are on and there's a single inhabitant of this world (the lamplighter) standing at a particular vertex. At any step he may either move to a neighboring vertex or change the state (on to off or off to on) of the lamp at his current vertex. This situation can be modeled as a group Lamplighter($\mathbb{Z}$). The vertex set is the set of pairs (x, A) where $x \in \mathbb{Z}$ and A is a finite subset of $\mathbb{Z}$ (the set of lamps which are on) the group operation is*

$$(x, A)(y, B) = (x + y, A \triangle (B + x)).$$

where the triangle denotes symmetric difference. The "elementary moves" correspond to multiplication by elements of the generating set $F = \{(\pm 1, \{\}), (0, \{0\})\}$. Is the random walk on Lamplighter($\mathbb{Z}$) recurrent?

Question 3. *Can it be the case that the Cayley graph associated to one generator for a group G is recurrent while for some other generator it's transient? i.e. Can we speak of recurrent vs transient groups or must one always include the generator?*

1.4 Notes for further reading

There are at least two very good books covering the subject matter of these notes and more: The book by Woess "Random walks on infinite graphs and groups" [Woe00] and "Probability on trees and infinite networks" by Russell Lyons and Yuval Peres [LP15]. The lecture notes by Benjamini [Ben13] can give the reader a good idea of where some of the more current research is heading.

The reader interested in the highly related area of random walks on finite graphs might begin by looking at the survey article by Laslo Lovasz [Lov96] (with the nice dedication to the random walks of Paul Erdös).

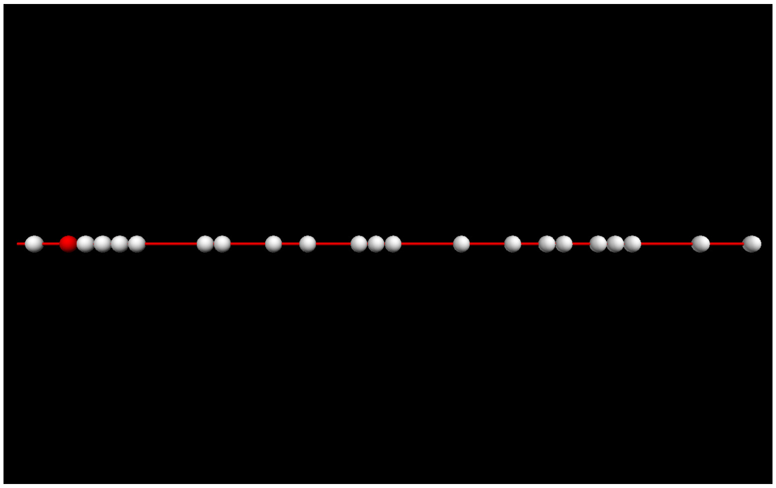

Figure 6: Drunken Lamplighter
I've simulated a random walk on the lamplighter group. After a thousand steps the lamplighter's position is shown as a red sphere and the white spheres represent lamps that are on. One knows that the lamplighter will visit every lamp infinitely many times, but will he ever turn them all off again?

For the very rich theory of the simple random walk on $\mathbb{Z}^d$ a good starting point is the classical book by Spitzer [Spi76] and the recent one by Lawler [Law13].

For properties of classical Brownian motion on $\mathbb{R}^d$ it's relatively painless and highly motivating to read the article by Einstein [Ein56] (this article actually motivated the experimental confirmation of the existence of atoms via observation of Brownian motion, for which Jean Perrin recieved the Nobel prize later on!). Mathematical treatment of the subject can be found in several places such as Mörters and Peres' [MP10].

Anyone interested in random matrix products should start by looking at the original article by Furstenberg [Fur63], as well as the excellent book by Bougerol and Lacroix [BL85].

Assuming basic knowledge of stochastic calculus and Riemannian geometry I can recommend Hsu's book [Hsu02] for Brownian motion on Riemannian manifolds, or Stroock's [Str00] for those preferring to go through more complicated calculations in order to avoid the use of stochastic calculus . Purely analytical treatment of the heat kernel (including upper and lower bounds) is well given in Grigoryan's [Gri09]. The necessary background material in Riemannian geometry is not very advanced into the subject and is well treated in several places (e.g. Petersen's book [Pet06]), similarly, for the Stochastic calculus background there are several good references such as Le Gall's [LG13].

2 The entry fee: A crash course in probability theory

Let me quote Kingman who said it better than I can.

> The writer on probability always has a dilemma, since the mathematical basis of the theory is the arcane subject of abstract measure theory, which is unattractive to many potential readers...

That's the short version. A longer version is that there was a long conceptual struggle between discrete probability (probability = favorable cases over total cases; the theory of dice games such as Backgammon) and geometric probability (probability = area of the good region divided by total area; think of Buffon's needle or Google it). After the creation of measure theory at the turn of the 20th century people started to discover that geometric probability based on measure theory was powerful enough to formulate all concepts usually studied in probability. This point of view was clearly stated for the first time in Kolmogorov's seminal 1933 article (the highlights being, measures on spaces of infinite sequences and spaces of continuous functions, and conditional expectation formalized as a Radon-Nikodym derivative). Since then probabilistic concepts are formalized mathematically as statements about measurable functions whose domains are probability spaces. Some people were (and some still are) reluctant about the predominant role played by measure theory in probability theory (after all, why should a random variable be a measurable function?) but this formalization has been tremendously successful and has not only permitted the understanding of somewhat paradoxical concepts such as Brownian motion and Poisson point processes but has also allowed for deep interactions between probability theory and other areas of mathematics such as potential theory and dynamical systems.

We can get away with a minimal amount of measure theory here thanks to the fact that both time and space are discrete in our case of interest. But some is necessary (without definitions there can be no proofs). So here it goes.

2.1 Probability spaces, random elements

A probability space is a triplet consisting of a set Ω (the points of which are sometimes called "elementary outcomes"), a family $\mathcal{F}$ of subsets of Ω (the elements of which are called "events") which is closed under complementation and countable union (i.e. $\mathcal{F}$ is a σ-algebra), and a function $\mathbb{P} : \mathcal{F} \to [0, 1]$ (the probability) satisfying $\mathbb{P}[\Omega] = 1$ and more importantly

$$\mathbb{P}\left[\bigcup_n A_n\right] = \sum_n \mathbb{P}[A_n]$$

for any countable (or finite) sequence $A_1, \ldots, A_n, \ldots$ of pairwise disjoint elements of $\mathcal{F}$.

A random element is a function x from a probability space Ω to some complete separable metric space X (i.e. a Polish space) with the property that the preimage of any open set belongs to the σ-algebra $\mathcal{F}$ (i.e. x is measurable). Usually if the function takes values in $\mathbb{R}$ it's called a random variable, and as Kingman put it, a random elephant is a measurable function into a suitable space of elephants.

The point is that given a suitable (which means Borel, i.e. belonging to the smallest σ-algebra generated by the open sets) subset A of the Polish space X defined by some property P one can give meaning to "the probability that the random element x satisfies property P" simply by assigning it the number $\mathbb{P}\left[x^{-1}(A)\right]$ which sometimes will be denoted by $\mathbb{P}\left[x \text{ satisfies property } P\right]$ or $\mathbb{P}\left[x \in A\right]$.

Some people don't like the fact that $\mathbb{P}$ is assumed to be countably additive (instead of just finitely additive). But this is crucial for the theory and is necessary in order to make sense out of things such as $\mathbb{P}\left[\lim x_n \text{ exists }\right]$ where x_n is a sequence of random variables (results about the asymptotic behaviour of sequences of random elements abound in probability theory, just think of or look up the Law of Large Numbers, Birkhoff's ergodic theorem, Doob's martingale convergence theorem, and of course Pólya's theorem).

2.2 Distributions

The distribution of a random element x is the Borel (meaning defined on the σ-algebra of Borel sets) probability measure μ defined by $\mu(A) = \mathbb{P}\left[x \in A\right]$. Similarly the joint distribution of a pair of random elements x and y of spaces X and Y is a probability on the product space $X \times Y$, and the joint distribution of a sequence of random variables is a probability on a sequence space (just group all the elements together as a single random element and consider its distribution).

Two events A and B of a probability space Ω are said to be independent if $\mathbb{P}\left[A \cap B\right] = \mathbb{P}\left[A\right]\mathbb{P}\left[B\right]$. Similarly, two random elements x and y are said to be indpendent if $\mathbb{P}\left[x \in A, y \in B\right] = \mathbb{P}\left[x \in A\right]\mathbb{P}\left[y \in B\right]$ for all Borel sets A and B in the corresponding ranges of x and y. In other words they are independent if their joint distribution is the product of their individual distributions. This definitions generalizes to sequences. The elements of a (finite or countable) sequence are said to be independent if their joint distribution is the product of their individual distributions.

A somewhat counterintuitive fact is that independence of the pairs $(x, y), (x, z)$ and (y, z) does not imply indepedence of the sequence x, y, z. An example is obtained by letting x and y be independent with $\mathbb{P}\left[x = \pm 1\right] = \mathbb{P}\left[y = \pm 1\right] = 1/2$ and $z = xy$ (the product of x and y). A random element is independent from itself if and only if it is almost surely constant (this strange legalistic loophole is actually used in the proof of some results, such as Kolmogorov's zero-one law or the ergodicity of the Gauss continued fraction map; one shows that an event has probability either zero or one by showing that it is independent from itself).

The philosophy in probability theory is that the hypothesis and statements of the theorems should depend only on (joint) distributions of the variables involved (which are usually assumed to satisfy some weakened form of independence) and not on the underlying probability space (of course there are some exceptions, notably Skorohod's representation theorem on weak convergence where the results is that a probability space with a certain sequence of variables defined on it exists). The point of using a general space Ω instead of some fixed Polish space X with a Borel probability μ is that, in Kolmogorov's framework, one may consider simultaneously random objects on several different spaces and combine them to form new ones. In fact one could base the whole theory (pretty much) on $\Omega = [0,1]$ endowed with Lebesgue measure on the Borel sets and a few things would be easier (notably conditional probabilities), but not many people like this approach nowadays (the few who do study "standard probability spaces").

2.3 Markov chains

Consider a countable set X. The space of sequences $X^{\mathbb{N}} = \{\omega = (\omega_1, \omega_2, \ldots) : \omega_i \in X\}$ of elements of X is a complete separable metric space when endowed with the distance

$$d(\omega, \omega') = \sum_{\{n:\omega_n \neq \omega'_n\}} 2^{-n}$$

and the topology is that of coordinate-wise convergence.

For each finite sequence $x_1, \ldots, x_n \in X$ we denote the subset of $X^{\mathbb{N}}$ consisting of infinite sequences begining with $x_1, \ldots, x_n$ by $[x_1, \ldots, x_n]$.

A probability on X can be identified with a function $p : X \to [0,1]$ such that $\sum\limits_x p(x) = 1$. By a transition matrix we mean a function $q : X \times X \to [0,1]$ such that $\sum\limits_{y \in X} q(x,y) = 1$ for all x.

The point of the following theorem is that, interpreting $p(x)$ to be the probability that a random walk will start at x and $q(x,y)$ as the probability that the next step will be at y given that it is currently at x, the pair p, q defines a unique probability on $X^{\mathbb{N}}$.

Theorem 5 (Kolmogorov extension theorem)**.** *For each probability p on X and transition matrix q there exists a unique probability $\mu_{p,q}$ on $X^{\mathbb{N}}$ such that*

$$\mu_{p,q}([x_1, \ldots, x_n]) = p(x_1)q(x_1, x_2) \cdots q(x_{n-1}, x_n).$$

A Markov chain with initial distribution p and transition matrix q is a sequence of random elements $x_1, \ldots, x_n, \ldots$ (defined on some probability space) whose joint distribution is $\mu_{p,q}$.

Markov chains are supposed to model "memoryless processes". That is to say that what's going to happen next only depends on where we are now (plus randomness) and not on the entire history of what happened before. To formalize this we need the notion of conditional probability with respect to an event which in our case can be simply defined as

$$\mathbb{P}[B|A] = \mathbb{P}[B] / \mathbb{P}[A]$$

where $\mathbb{P}[A]$ is assumed to be positive (making sense out of conditional probabilities when the events with respect to which one wishes to condition have probability zero was one of the problems which pushed towards the development of the Kolmogorov framework... but we wont need that).

We can now formalize the "memoryless" property of Markov chains (called the Markov property) in its simplest form.

Theorem 6 (Weak Markov property)**.** *Let $x_0, x_1, \ldots$ be a Markov chain on a countable set X with initial distribution p and transition matrix q. Then for each fixed n the sequence $y_0 = x_n, y_1 = x_{n+1}, \ldots$ is also a Markov chain with transition matrix q (the initial distribution is simply the distribution of x_n). Furthermore the sequence $\mathbf{y} = (y_0, y_1, \ldots)$ is conditionally independent from $x_0, \ldots, x_{n-1}$ given x_n by which we mean*

$$\mathbb{P}[y \in A | x_0 = a_0, \ldots, x_n = a_n] = \mathbb{P}[y \in A | x_n = a_n]$$

for all $a_0, \ldots, a_n \in X$.

Proof. When the set A yields an event of $\{y_1 = b_1, \ldots, y_m = b_m\}$ the proof is by direct calculation. The general result follows by approximation. One uses the fact from measure theory (which we will not prove) that given a probability measure μ on $X^{\mathbb{N}}$ any Borel set A can be approximated by a countable union of sets defined by fixing the value of a finite number of coordinates (approximation means that the probability symmetric difference can be made arbitrarily small). This level of generality is needed in applications (e.g. a case we will use is when A is the event that $\mathbf{y}$ eventually hits a certain point $x \in X$). □

A typical example of a Markov chain is the simple random walk on $\mathbb{Z}$ (i.e. a Markov chain on $\mathbb{Z}$ whose transition matrix satisfies $q(n, n \pm 1) = 1/2$ for all n). For example, let p_k be the probability that a simple random walk $x_0, x_1, \ldots$ on $\mathbb{Z}$ starting at k eventually hits 0 (i.e. $p_k = \mathbb{P}[x_n = 0 \text{ for some } n]$). Then from the weak Markov property (with $n = 1$) one may deduce $p_k = \frac{1}{2}p_{k-1} + \frac{1}{2}p_{k+1}$ for all $k \neq 0$ (do this now!). Since $p_0 = 1$ it follows from this that in fact $p_k = 1$ for all k. Hence, the simple random walk on $\mathbb{Z}$ is recurrent!

A typical example of a random sequence which is not a Markov chain is obtained by sampling without replacement. For example, suppose you are turning over the cards of a deck (after shuffling) one by one obtaining a sequence $x_1, \ldots, x_{52}$ (if you want the sequence to be infinite, rinse and repeat). The more cards of the sequence one knows the more one can bound the conditional probability of the last card being the ace of spades for example (this fact is used by people who count cards in casinos like in the 2008 movie "21"). Hence this sequence is not reasonably modeled by a Markov chain.

2.4 Expectation

Given a real random variable x defined on some probability space its expectation is defined in the following three cases:

1. If x assumes only finitely many values $x_1, \ldots, x_n$ then $\mathbb{E}[x] = \sum_{k=1}^{n} \mathbb{P}[x = x_k] x_k$.

2. If x assumes only non-negative values then $\mathbb{E}[x] = \sup\{\mathbb{E}[y]\}$ where the supremum is taken over all random variables with $0 \le y \le x$ which assume only finitely many values. In this case one may have $\mathbb{E}[x] = +\infty$.

3. If $\mathbb{E}[|x|] < +\infty$ then one defines $\mathbb{E}[x] = \mathbb{E}[x^+] - \mathbb{E}[x^-]$ where $x^+ = \max(x, 0)$ and $x^- = \max(-x, 0)$.

Variables with $\mathbb{E}[|x|] = +\infty$ are said to be non-integrable. The contrary to this is to be integrable or to have finite expectation.

We say a sequence of random variables x_n converges almost surely to a random variable x if $\mathbb{P}[\lim x_n = x] = 1$. The bread and butter theorems regarding expectation are the following:

Theorem 7 (Monotone convergence)**.** *Let x_n be a monotone increasing (in the weak sense) sequence of non-negative random variables which converges almost surely to a random variable x. Then $\mathbb{E}[x] = \lim \mathbb{E}[x_n]$ (even in the case where the limit is $+\infty$).*

Theorem 8 (Dominated convergence)**.** *Let x_n be a sequence of random variables converging almost surely to a limit x, and suppose there exists an integrable random variable y such that $|x_n| \le y$ almost surely for all n. Then $\mathbb{E}[x] = \lim \mathbb{E}[x_n]$ (here the limit must be finite and less than $\mathbb{E}[y]$ in absolute value) .*

In probability theory the dominated convergence theorem can be used when one knows that all variables of the sequence x_n are uniformly bounded by some constant C. This is because the probability of the entire space is finite (counterexamples exist on infinite measure spaces such as $\mathbb{R}$ endowed with the Lebesgue measure).

As a student I was sometimes bothered by having to think about non-integrable functions to prove results about integrable ones (e.g. most theorems in probability, such as the Law of Large Numbers or the ergodic theorem, require some sort of integrability hypothesis). Wouldn't life be easier if we just did the theory for bounded functions for example? The thing is there are very natural examples of non-bounded and even non-integrable random variables. For example, consider a simple random walk $x_0, x_1, \ldots$ on $\mathbb{Z}$ starting at 1. Let τ be the first $n \ge 1$ such that $x_\tau = 0$. Notice that τ is random but we have shown using the weak Markov property that it is almost surely finite. However, it's possible to show that $\mathbb{E}[\tau] = +\infty$! This sounds paradoxical: τ is finite but its expected value is infinite. But, in fact, I'm pretty sure that with some thought a determined and tenacious reader can probably calculate $\mathbb{P}[\tau = n] = \mathbb{P}[x_1 \neq 0, \ldots, x_{n-1} \neq 0, x_n = 0]$ for each n. The statement is that $\sum \mathbb{P}[\tau = n] = 1$ but $\sum n\mathbb{P}[\tau = n] = +\infty$. That doesn't sound so crazy does it?

In fact there's an indirect proof that $\mathbb{E}[\tau] = +\infty$ using the Markov property. Let $e_1 = \mathbb{E}[\tau]$, e_2 be the expected value of the first hitting time of 0 for the random walk starting at 2, etc (so e_n is the expected hitting time of 0 for a simple random walk starting at n). Using the weak Markov property one can get

$$e_n = 1 + \frac{e_{n-1} + e_{n+1}}{2}$$

for all $n \geq 1$ (where $e_0 = 0$) which gives us $e_{n+1} - e_n = (e_n - e_{n-1}) - 2$ so any solution to this recursion is eventually negative. The expected values of the hittings times we're considering, if finite, would all be positive, so they're not finite.

2.5 The strong Markov property

A stopping time for a Markov chain $x_0, x_1, \ldots$ is a random non-negative integer τ (we allow $\tau = +\infty$ with positive probability) such that for all $n = 0, 1, \ldots$ the set $\{\tau = n\}$ can be written as a countable union of sets of the form $\{x_0 = a_0, \ldots, x_n = a_n\}$. Informally you can figure out if $\tau = n$ by looking at $x_0, \ldots, x_n$, you don't need to look at "the future".

The concept was originally created to formalize betting strategies (the decision to stop betting shouldn't depend on the result of future bets). And is central to Doob's Martingale convergence theorem which roughly states that all betting strategies are doomed to failure if the game is rigged in favor of the casino.

Stopping times are also natural and important for Markov chains. An example, which we have already encountered is the first hitting time of a subset $A \subset X$ which is defined by

$$\tau = \min\{n \geq 0 : x_n \in A\}.$$

A non-example is the last exit time $\tau = \max\{n : x_n \in A\}$. It's not a stopping time because one cannot determine if the chain is going to return or not to A after time n only by looking at $x_0, \ldots, x_n$.

On the set on which a stopping time is τ finite for a chain $x_0, x_1, \ldots$ one can define x_τ as the random variable which equals x_k exactly when $\tau = k$ for all $k = 0, 1, \ldots$.

If τ is a stopping time so are $\tau + 1, \tau + 2, \ldots$. But the same is not true in general for $\tau - 1$ (even if one knows that $\tau > 1$).

An event A is said to occur "before the stopping time τ" if for each finite n one can write $A \cap \{\tau = n\}$ as a countable union of events of the form $\{x_0 = a_0, \ldots, x_n = a_n\}$. A typical example is the event that the chain hits a point x before time τ (it sounds tautological but that's just because the name for this type of events is well chosen). The point is on each set $\{\tau = n\}$ you can tell if A happened by looking at $x_0, \ldots, x_n$.

The following is a generalization of the Weak Markov Property (Theorem 6) to stopping times. To state it, given a stopping time τ for a Markov chain $x_0, \ldots, x_n, \ldots$ defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we will need to consider the modified probability defined by $\mathbb{P}_\tau = \mathbb{P}1_{\tau < +\infty} / \mathbb{P}[\tau < +\infty]$ (i.e. the measure of an event A is $\mathbb{P}[A \cap \tau < +\infty] / \mathbb{P}[\tau < +\infty]$). Here it goes.

Theorem 9 (Strong Markov property)**.** *Let $x_0, x_1, \ldots$ be a Markov chain with transition matrix q and τ be a finite stopping. Then with respect to the probability $\mathbb{P}_\tau$ the sequence $\mathbf{y} = (y_0, y_1, \ldots)$ where $y_n = x_{\tau+n}$ is a Markov chain with transition matrix q. Furthermore $\mathbf{y}$ is independent from event prior to τ conditionally on on x_τ by which we mean that*

$$\mathbb{P}_\tau \left[\mathbf{y} \in B | A, x_\tau = x\right] = \mathbb{P}_\tau \left[\mathbf{y} \in B | x_\tau = x\right]$$

for all A occurring prior to τ, all $x \in X$, and all Borel $B \subset X^{\mathbb{N}}$.

The reader should try to work out the proof in enough detail to understand where the fact that τ is a stopping time is used.

To see why the generality of the strong Markov property is useful let's consider the simple random walk on $\mathbb{Z}$ starting at 0. We have shown that the probability of returning to 0 is 1. The strong Markov property allows us to conclude the intuitive fact that almost surely the random walk will hit 0 infinitely many times. That is one can go from

$$\mathbb{P}\left[x_n = 0 \text{ for some } n > 0\right] = 1$$

to

$$\mathbb{P}\left[x_n = 0 \text{ for infinitely many } n\right] = 1$$

.

To see this just consider $\tau = \min\{n > 0 : x_n = 0\}$ the first return time to 0. We have already seen that the x_n returns to 0 almost surely. The strong Markov property allows one to calculate the probability that x_n will visit 0 at least twice as $\mathbb{P}\left[\tau < +\infty\right] \mathbb{P}_\tau \left[x_{\tau+n} = 0 \text{ for some } n > 0\right]$ and tells us that the second factor is exactly the probability that x_n returns to zero at least once (which is 1). Hence the probability of returning twice is also equal to 1 and so on and so forth. Finally, after we've shown that event $\{x_n = 0 \text{ at least } k \text{ times}\}$ has full probability their intersection (being countable) also does.

This application may not seem very impressive. It gets more interesting when the probability of return is less than 1. We will use this in the next section.

2.6 Simple random walks on graphs

We will now define precisely, for the first time, our object of study for these notes.

Given a graph X (we will always assume graphs are connected, undirected, and locally finite) a simple random walk on X is a Markov chain $x_0, \ldots, x_n, \ldots$ whose transition probabilities are defined by

$$q(x, y) = \frac{\text{number of edges joining } x \text{ to } y}{\text{total number of edges with } x \text{ as an endpoint}}.$$

The distribution of the initial vertex x_0 is unspecified. The most common choice is just a deterministic vertex x and in this case we speak of a simple random starting at x.

Definition 1 (Recurrence and Transience)**.** *A simple random walk* $\{x_n : n \geq 0\}$ *on a graph* X *is said to be recurrent if almost surely it visits every vertex in* X *infinitely many times i.e.*

$$\mathbb{P}\left[x_n = x \textit{ for infinitely many } n\right] = 1.$$

If this is not the case then we say the walk is transient.

Since there are countably many vertices in X recurrence is equivalent to the walk visiting each fixed vertex infinitely many times almost surely (that is $\mathbb{P}\left[x_n = x \text{ for infinitely many } n\right] = 1$). In fact we will now show something stronger and simultaneously answer the question of whether or not recurrence depends on the initial distribution of the random walk (it doesn't).

Lemma 1 (Recurrence vs Transience Dichotomy)**.** *Let* x *be a vertex of a graph* X *and* p *be the probability that a simple random walk starting at* x *will return to* x *at some positive time. Then* $p = 1$ *if and only if all simple random walks on* X *are recurrent. Furthermore for any simple random walk* $x_0, x_1, \ldots$ *the expected number of visits to* x *is given by*

$$\mathbb{E}\left[|\{n > 0 : x_n = x\}|\right] = \sum_{n=0}^{+\infty} \mathbb{P}\left[x_n = x\right] = \mathbb{P}\left[x_n = x \textit{ for some } n\right]/(1-p).$$

Proof. Define $f : X \to \mathbb{R}$ by setting $f(y)$ to be the probability that the simple random walk starting at y will visit x (in particular $f(x) = 1$). By the weak Markov property one has

$$f(y) = \sum_{z \in X} q(y,z) f(z)$$

for all $y \neq x$ (functions satisfying this equation are said to be harmonic) and

$$p = \sum_{z \in X} q(x,z) f(z).$$

If $p = 1$ then using the above equation one concludes that $f(y) = 1$ at all neighbors of x. Iterating the same argument one sees that $f(y) = 1$ at all vertices at distance 2 from x and etc. This shows that the probability that a simple random walk starting at any vertex will eventually visit x is 1, and hence all simple random walks on X visit the vertex x almost surely at least once.

Given any simple random walk x_n and letting τ be the first hitting time of x, one knows that τ is almost surely finite. Also, by the strong Markov property, one has that $x_\tau, x_{\tau+1}, \ldots$ is a simple random walk starting at x. Hence τ_2 the second hitting time of x is also almost surely finite, and so on so forth. This shows that $p = 1$ implies that every simple random walk hits x infinitely many times almost surely. Also, given any vertex y the probability q that a simple random walk starting at x will return to x before hitting y is strictly less than 1. The strong Markov property implies that the probability of returning to x for the n-th time without hitting y is q^n, since this goes to 0 one obtains that any simple random walk will almost surely hit y eventually, and therefore infinitely many times. In short all simple random walks on X are recurrent.

We will now establish the formula for the expected number of visits.

Notice that the number of visits to x of a simple random walk x_n is simply

$$\sum_{n=0}^{+\infty} 1_{\{x_n=x\}}$$

where 1_A is the indicator of an event A (i.e. the function taking the value 1 exactly on A and 0 everywhere else in the domain probability space Ω). Hence the first equality in the formula for the expected number of visits follows simply by the monotone convergence theorem.

The equality of this to the third term in the case $p = 1$ is trivial (all terms being infinite). Hence we may assume from now on that $p < 1$.

In order to establish the second equality we use the sequence of stopping times $\tau, \tau_1, \tau_2, \ldots$ defined above. The probability that the number of visits to x is exactly n is shown by the strong Markov property to be $\mathbb{P}[\tau < +\infty]\, p^{n-1}(1-p)$. Using the fact that $\sum np^{n-1} = 1/(1-p)^2$ one obtains the result. □

2.7 A counting proof of Pólya's theorem

For such a nice result Pólya's theorem admits a decidedly boring proof. Let $p_n = \mathbb{P}[x_n = 0]$ for each n where x_n is a simple random walk starting at 0 on $\mathbb{Z}^d$ ($d = 1$ or 2). In view of the previous subsection it suffices to estimate p_n well enough to show that $\sum p_n = +\infty$ when $d = 2$ and $\sum p_n < +\infty$ when $d = 3$.

Noticing that the total number of paths of length n in $\mathbb{Z}^d$ starting at 0 is $(2d)^n$ one obtains

$$p_n = (2d)^{-n}|\text{closed paths of length of length } n \text{ starting at } 0|$$

so that all one really needs to estimate is the number of closed paths of length n starting and ending at 0 (in particular notice that $p_n = 0$ if n is odd, which is quite reasonable).

For the case $d = 2$ the number of closed paths of length $2n$ can be seen to be $\binom{2n}{n}^2$ by bijection as follows: Consider the $2n$ increments of a closed path. Associate to each increment one of the four letters $aAbB$ according to the following table

$(1,0)$	a
$(-1,0)$	B
$(0,1)$	A
$(0,-1)$	b

.

Notice that the total number of 'a's (whether upper or lowercased) is n and that the total number of uppercase letters is also n. Reciprocally if one is given two subsets A, U of $\{1, \ldots, 2n\}$ with n elements each one can obtain a closed path by interpreting the set A as the set of increments labeled with 'a' or 'A' and the set U as those labelled with an upper case letter.

Hence, one obtains for $d = 2$ that

$$p_{2n} = 4^{-2n} \binom{2n}{n}^2.$$

Stirling's formula $n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$ implies that

$$p_{2n} \sim \frac{1}{\pi n},$$

which clearly implies that the series $\sum p_n$ diverges. This shows that the simple random walk on $\mathbb{Z}^2$ is recurrent.

For $d = 3$ the formula gets a bit tougher to analyze but we can grind it out.

First notice that if $\sum p_{6n}$ converges then so does $\sum p_n$ (because the number of closed paths of length $2n$ is an increasing function of n). Hence we will bound only the number of closed paths of length $6n$.

The number of closed paths of length $6n$ is (splitting the $6n$ increments into 6 groups according to their value and noticing that there must be the same number of $(1, 0, 0)$ as $(-1, 0, 0)$,etc)

$$\sum_{a+b+c=3n} \frac{(6n)!}{(a!)^2(b!)^2(c!)^2} \le \frac{(6n)!}{(n!)^3} \sum_{a+b+c=3n} \frac{1}{a!b!c!} \le \frac{(6n)!}{(3n)!(n!)^3} \sum_{a+b+c=3n} \frac{(3n)!}{a!b!c!}$$
$$= \frac{(6n)!}{(3n)!(n!)^3} 3^{3n} \sim 6^{6n} \frac{1}{\sqrt{4\pi^3 n^3}},$$

where the first inequality is because $a!b!c! \ge (n!)^3$ when $a + b + c = 3n$ and we use that 3^{3n} is the sum of $(3n)!/(a!b!c!)$ over $a + b + c = 3n$.

Hence we have shown that for all $\epsilon > 0$ one has $p_{6n} \le (1 + \epsilon)/\sqrt{4\pi^3 n^3}$ for n large enough. This implies $\sum p_n < +\infty$ and hence the random walk on $\mathbb{Z}^3$ is transient.

This proof is unsatisfactory (to me at least) for several reasons. First, you don't learn a lot about why the result is true from this proof. Second, it's very inflexible, remove or add a single edge from the $\mathbb{Z}^d$ graph and the calculations become a lot more complicated, not to mention if you remove or add edges at random or just consider other types of graphs. The ideas we will discuss in the following section will allow for a more satisfactory proof, and in particular will allow one to answer questions such as the following:

Question 4. *Can a graph with a transient random walk become recurrent after one adds edges to it?*

At one point while learning some of these results I imagined the following: Take $\mathbb{Z}^3$ and add an edge connecting each vertex (a, b, c) directly to 0. The simple random walk on the resulting graph is clearly recurrent since at each step it has probability $1/7$ of going directly to 0. The problem with this example is that it's not a graph in our sense and hence does not have a simple random walk (how many neighbors does 0 have? how would one define the transition probabilities for 0?). In fact, we will show that the answer to the above question is negative.

2.8 The classification of recurrent radially symmetric trees

David Blackwell is a one of my favorite mathematicians because of his great life story combined with some really nice mathematical results. He was a black man, born in Centralia, Illinois in 1919 (a town which today has some 13 thousand inhabitants, and which is mostly notable for a mining accident which killed around 100 people in the late 1940s). He went to the university at Urbana-Champagne starting at the age of 16 and ended up doing his Phd. under the tutelage of Joseph Doob (who was one of the pioneering adopters of Kolmogorov's formalism for probability). He did a postdoc at the Institute of Advanced Study (where he recalled in an interview having fruitful interactions for example with Von Neumann). Apparently during that time he wasn't allowed to attend lectures at Princeton University because of his color. Later on he was also considered favorably for a position at Berkeley which it seems also didn't work out because of racism (though Berkeley did hire him some 20 years later). He was the first black member of the National Academy of Sciences.

Anyway, I'd like to use this section to talk about a very elegant result from a paper he wrote in 1955. The result will allows us to immediately classify the radially symmetric trees whose simple random walk is recurrent. The idea is to study the Markov chain on the non-negative integers whose transition matrix satisfies $q(0,0) = 1$, $q(n, n+1) = p_n$ and $q_n = 1 - p_n = q(n, n-1)$ for all $n \geq 1$. The question of interest is: What is the probability that such a Markov chain started at a vertex n will eventually hit zero? Let's look at two examples.

When $p_n = 1/2$ for all n one gets the simple random walk "stopped (or absorbed or killed, terminologies vary according to the different author's moods apparently) at 0". We have already shown that the probability of hitting zero for the walk started at any n is equal to 1. We did this by letting $f(n)$ be the corresponding hitting probability and noting that (by the weak Markov property) $f(0) = 1$ and f satisfies

$$f(n) = \frac{1}{2}f(n-1) + \frac{1}{2}f(n+1)$$

for all $n \geq 1$. Since there are no non-constant bounded solutions to the above equation we get that $f(n) = 1$ for all n.

When $p_n = 2/3$ (a chain which goes right with probability 2/3 and left with probability 1/3) we can set $f(n)$ once again to be the probability of hitting 0 if the chain starts at n and notice that $f(0) = 1$ and f satisfies

$$f(n) = \frac{1}{3}f(n-1) + \frac{2}{3}f(n+1)$$

for all $n \geq 1$. It turns out that there are non-constant bounded solutions to this equation, in fact $g(n) = 2^{-n}$ is one such solution and all other solutions are of the form $\alpha + \beta g$ for some constants α and β. Does this mean that there's a positive probability of never hitting 0? The answer is yes and that's the content of Blackwell's theorem (this is a special case of a general result, the original paper is very readable and highly recommended).

Theorem 10 (Blackwell's theorem). *Consider the transition matrix on $\mathbb{Z}^+$ defined by $q(0,0) = 1$, $q(n,n+1) = p_n$ and $q_n = 1 - p_n = q(n,n-1)$ for all $n \geq 1$. Then any Markov chain with this transition matrix will eventually hit* 0 *almost surely if and only if the equation*

$$f(n) = q_n f(n-1) + p_n f(n+1)$$

has no non-constant bounded solution.

Proof. One direction is easy. Setting $f(n)$ to be the probability of hitting 0 for the chain starting at n one has that f is a solution and $f(0) = 1$. Therefore if the probability doesn't equal 1 one has a non-constant bounded solution.

The other direction is a direct corollary of Doob's Martingale convergence theorem (which regrettably we will not do justice to in these notes). The proof goes as follows. Suppose f is a non-constant bounded solution. Then the Martingale convergence theorem yields that the random limit $L = \lim\limits_{n \to +\infty} f(x_n)$ exists almost surely and is non-constant. If the probability of hitting 0 were 1 then one would have $L = f(0)$ almost surely contradicting the fact that L is non-constant. □

Blackwell's result allows the complete classification of radially symmetric trees whose random walk is recurrent.

Corollary 1 (Classification of recurrent radially symmetric trees). *Let T be a radially symmetric tree where the n-th generation vertices have a_n children. Then the simple random walk on T (with any starting distribution) is recurrent if and only if*

$$\sum_{n=1}^{+\infty} \frac{1}{a_1 a_2 \cdots a_n} = +\infty.$$

Proof. The distance to the root of a random walk on T is simply a Markov chain on the non-negative integers with transition probability $q(0,1) = 1$, $q(n,n-1) = 1/(1+a_n) = q_n$, and $q(n,n+1) = a_n/(1+a_n) = p_n$. Clearly such a chain will eventually almost surely hit 0 if and only if the corresponding chain with modified transition matrix $q(0,0) = 1$ does.

Hence by Blackwell's result the probability of hitting 0 is 1 for all simple random walks on T if and only if there are no non-constant bounded solutions to the equation

$$f(n) = q_n f(n-1) + p_n f(n+1)$$

on the non-negative integers.

It's a linear recurrence of order two so one can show the space of solutions is two dimensional. Hence all solutions are of the form $\alpha + \beta g$ where g is the solution with $g(0) = 0$ and $g(1) = 1$. Manipulating the equation one gets the equation $g(n+1) - g(n) = (g(n) - g(n-1))/a_n$ for the increments of g. From this it follows that

$$g(n) = 1 + \frac{1}{a_1} + \frac{1}{a_1 a_2} + \cdots + \frac{1}{a_1 a_2 \cdots a_{n-1}}$$

for all n. Hence the random walk is recurrent if and only if the $g(n) \to +\infty$ when $n \to +\infty$, as claimed. □

2.9 Notes for further reading

For results from basic measure theory and introductory measure-theoretic probability I can recommend Ash's book [Ash00]. Varadhan's books [Var01] and [Var07] are also excellent and cover basic Markov chains (including the Strong Markov property which is [Var01, Lemma 4.10]) and more.

I quoted Kingman's book [Kin93] on Poisson point processes a couple of times in the section and, even though it is in principle unrelated to the material actually covered in these notes, any time spent reading it will not be wasted.

The history of measure-theoretic probability is very interesting in its own right. The reader might want to take a look at the articles by Doob (e.g. [Doo96] or [Doo89]) about resistance to Kolmogorov's formalization. The original article by Kolmogorov [Kol50] is still readable and one might also have a look at the discussion given in [SV06].

The counting proof of Polya's theorem is essentially borrowed from [Woe00].

Blackwell's original article is [Bla55] and the reader might want to look at Kaimanovich's [Kai92] as well.

3 The Flow Theorem

Think of the edges of a graph as tubes of exactly the same size which are completely filled with water. A flow on the graph is a way of specifying a direction and speed for the water in each tube in such a way that the amount of water which enters each vertex per unit time equals the amount exiting it. We allow a finite number of vertices where there is a net flux of water outward or inwards (such vertices are called sources and sinks respectively). A flow with a single source and no sink (such a thing can only exist on an infinite graph) is called a source flow. The energy of a flow is the total kinetic energy of the water. A flow on an infinite graph may have either finite or infinite energy. Here's a beautiful theorem due to Terry Lyons (famous in households world wide as the creator of the theory of Rough paths, and deservedly so).

Theorem 11 (The flow theorem)**.** *The simple random walk on a graph is transient if and only if the graph admits a finite energy source flow.*

3.1 The finite flow theorem

The flow theorem can be reduced to a statement about finite graphs but first we need some notation.

So far, whenever X was a graph, we've been abusing notation by using X to denote the set of vertices as well (i.e. $x \in X$ means x is a vertex of X). Let's complement this notation by setting $E(X)$ to be the set of edges of X. We consider all edges to be directed so that each edge $e \in E(X)$ has a starting vertex e_- and an end e_+. The fact that our graphs are undirected means there is a bijective involution $e \mapsto e^{-1}$ sending each edge to another one with the start and end vertex exchanged (loops can be their own inverses).

A field on X is just a function φ from $E(X)$ to $\mathbb{R}$ satisfying $\varphi(e^{-1}) = -\varphi(e)$ for all edges $e \in E(X)$. Any function $f : X \to \mathbb{R}$ has an asociated gradient field ∇f defined by $\nabla f(e) = f(e_+) - f(e_-)$. The energy of a field $\varphi : E(X) \to \mathbb{R}$ is defined by $\mathcal{E}(\varphi) = \sum \frac{1}{2}\varphi(e)^2$. The divergence of a field $\varphi : E(X) \to \mathbb{R}$ is a function $\text{div}(\varphi) : X \to \mathbb{R}$ defined by $\text{div}(\varphi)(x) = \sum\limits_{e_-=x} \varphi(e)$.

Obviously all the definitions were chosen by analogy with vector calculus in $\mathbb{R}^n$. Here's the analogous result to integration by parts.

Lemma 2 (Summation by parts)**.** *Let f be a function and φ a field on a finite graph X. Then one has*

$$\sum_{e \in E(X)} \nabla f(e)\varphi(e) = -2 \sum_{x \in X} f(x)\,div(\varphi)(x).$$

A flow between two vertices a and b of a graph X is a field with divergence 0 at all vertices except a and b, with positive divergence at a and negative divergence at b. The point of the following theorem is that if X is finite there is a unique (up to constant multiples) energy-minimizing flow from a to b and it's directly related to the behaviour of the simple random walk on the graph X.

Theorem 12 (Finite flow theorem)**.** *Let a and b be vertices of a finite graph X and define $f : X \to \mathbb{R}$ by*

$$f(x) = \mathbb{P}\left[x_n \text{ hits } b \text{ before } a\right]$$

where x_n is a simple random walk starting at x. Then the following properties hold:

1. *The gradient of f is a flow from a to b and satisfies $div(\nabla f)(a) = \mathcal{E}(\nabla f) = \mathbb{P}\left[x_n = b \text{ before returning to } a\right]$ where x_n is a simple random walk starting at a.*

2. *Any flow φ from a to b satisfies $\mathcal{E}(\varphi) \geq \frac{\mathcal{E}(\nabla f)}{div(\nabla(f))(a)^2} div(\varphi)(a)^2$ with equality if and only if $\varphi = \lambda \nabla f$ for some constant λ.*

Proof. Notice that $f(a) = 0, f(b) = 1$ and $f(x) \in (0, 1)$ for all other vertices x. Hence the divergence at a is positive and the divergence at b is negative. The fact that the divergence of the gradient is zero at the rest of the vertices follows from the weak Markov property (f is harmonic except at a and b). Hence f is a flow from a to b as claimed. The formula for divergence at a follows directly from the definition, and for the energy one uses summation by parts.

The main point of the theorem is that ∇f is the unique energy minimizing flow (for fixed divergence at a). To see this consider any flow φ from a to b with the same divergence as ∇f at a.

We will first show that unless φ is the gradient of a function it cannot minimize energy. For this purpose assume that $e_1, \ldots, e_n$ is a closed path in the graph (i.e. $e_{1+} = e_{2-}, \ldots, e_{n+} = e_{1-}$) such that $\sum \varphi(e_k) \neq 0$. Let ψ be the flow such that $\psi(e_k) = 1$ for all e_k and $\psi(e) = 0$ unless $e = e_k$ or $e = e_k^{-1}$ for some k (so ψ is the unit flow around the closed path). One may calculate to obtain

$$\partial_t \mathcal{E}\left(\varphi + t\psi\right)_{|t=0} \neq 0$$

so one obtains for some small t (either positive or negative) a flow with less energy than φ and the same divergence at a. This shows that any energy minimizing flow with the same divergence as ∇f at a must be the gradient of a function.

Hence it suffices to show that if $g : X \to \mathbb{R}$ is a function which is harmonic except at a and b and such that $\operatorname{div}(\nabla g)(a) = \operatorname{div}(\nabla f)(a)$ then $\nabla g = \nabla f$. For this purpose notice that because ∇g is a flow one has $\operatorname{div}(\nabla g)(b) = -\operatorname{div}(\nabla g)(a) = -\operatorname{div}(\nabla f)(a) = \operatorname{div}(\nabla f)(b)$. Hence $f - g$ is harmonic on the entire graph and therefore constant. □

Exercise 1. *Show that all harmonic functions on a finite graph are constant.*

3.2 A proof of the flow theorem

Let X be an infinite graph now and x_n be a simple random walk starting at a. For each n let p_n be the probability that $d(x_n, a) = n$ before x_n returns to a and notice that

$$\lim_{n \to +\infty} p_n = p = \mathbb{P}\left[x_n \text{ never returns to } a\right].$$

For each n consider the finite graph X_n obtained from X by replacing all vertices with $d(x, a) \geq n$ by a single vertex b_n (edges joining a vertex at distance $n-1$ to one at distance n now end in this new vertex, edges joining two vertices at distance larger than n from a in X are erased).

The finite flow theorem implies that any flow from a to b_n in X_n with divergence 1 at a has energy greater than or equal to $1/p_n$. Notice that if X is recurrent $p_n \to 0$ and this shows there is no finite energy source flow (with soure a) on X.

On the other hand if X is transient then $p > 0$, so by the finite flow theorem there are fields φ_n with divergence 1 at a, divergence 0 at vertices with $d(x, a) \neq n$ and with energy less than $1/p$. For each edge $e \in E(X)$ the sequence $\varphi_n(e)$ is bounded, hence, using a diagonal argument, there is a subsequence φ_{n_k} which converges to a finite energy flow with divergence 1 at a and 0 at all other vertices (i.e. a finite energy source flow). This completes the proof (in fact one could show that the sequence φ_n converges directly if one looks at how it is defined in the finite flow theorem).

3.3 Adding edges to a transient graph cannot make it recurrent

A first, very important, and somewhat surprising application of the flow theorem is the following.

Corollary 2. *If a graph X has a transient subgraph then X is transient.*

3.4 Recurrence of groups is independent of finite symmetric generating set

Suppose G is a group which admits both F and F' as finite symmetric generating sets. The Cayley graphs associated to F and F' are different and in general one will not be a subgraph of the other (unless $F \subset F'$ or something like that, for example the Cayley graph of $\mathbb{Z}^2$ with respect to $\{(\pm 1, 0), (0, \pm 1), (\pm 1, \pm 1)\}$ is also recurrent by the previous corollary). However the flow theorem implies the following:

Corollary 3. *Recurrence or transience of a Cayley graph depends only on the group and not on the finite symmetric generator one chooses.*

Proof. Suppose that the graph with respect to one generator admits a finite energy source flow φ. Notice that φ is the gradient of a function $f : G \to \mathbb{R}$ with $f = 0$ at the source. The gradient of the same function with respect to the other graph structure also has finite energy (prove this!).

This does not conclude the proof because the gradient of f on the second graph is not necessarily a flow. We will sketch how to solve this issue.

Modifying f slightly one can show that for any n one can find a function taking the value 0 at the identity element of the group and 1 outside of the the ball of radius n and such that the energy of the flow is bounded by a constant which does not depend on n. Mimicking the argument used to prove the finite flow one can in fact find a a harmonic function for each n satisfying the same restrictions and whose gradient has energy bounded by the same constant. Taking a limit (as in the proof of the flow theorem) one obtains a function whose gradient is a finite energy source flow on the second Cayley graph. □

This is a special case of a very important result which we will not prove but which can also be proved via the flow theorem. Two metric spaces (X, d) and (X', d') are said to be quasi-isometric if they are large-scale bi-lipshitz i.e. there exists a function $f : X \to X'$ (not necesarilly continuous) and a constant C such that

$$d(x, y)/C - C \leq d(f(x), f(y)) \leq Cd(x, y) + C,$$

and $f(X)$ is C-dense in X'.

The definition is important because it allows one to relate discrete and continuous objects. For example $\mathbb{Z}^2$ and $\mathbb{R}^2$ are quasi-isometric (projection to the nearest integer point gives a quasi-isometry in one direction, and simple inclusion gives one in the other).

Exercise 2. *Prove that Cayley graphs of the same group with respect to different generating sets are quasi-isometric.*

An important fact is that recurrence is invariant under quasi-isometry.

Corollary 4 (Kanai's lemma)**.** *If two graphs of bounded degree X and X' are quasi-isometric then either they are both recurrent or they are both transient.*

The bounded degree hypothesis can probably be replaced by something sharper (I really don't know how much is known in that direction). The problem is that adding multiple edges between the same two vertices doesn't change the distance on a graph but it may change the behavior of the random walk.

Exercise 3. *Prove that the random walk on the graph with vertex set $\{0, 1, 2, \ldots\}$ where n is joined to $n+1$ by 2^n edges is transient.*

3.5 Boundary and Core of a tree

Let T be a tree and fix a root vertex $a \in T$. We define the boundary ∂T of T as the set of all infinite sequences of edges $(e_0, e_1, \ldots) \in E(T)^{\mathbb{N}}$ such that $e_{n+} = e_{(n+1)-}$ and $d(a, e_{n-}) = n$ for all n. We turn the boundary into a metric space by using the following distance

$$d((e_0, e_1, \ldots), (e'_0, e'_1, \ldots)) = e^{-\min\{n : e_n \neq e'_n\}}.$$

Exercise 4. *Prove that ∂T is a compact metric space.*

The core $\mathring{T}$ of T is the subtree consisting in all edges and which appear in some path of ∂T and the vertices they go through.

Exercise 5. *Prove that given the metric space ∂T it is possible to reconstruct $\mathring{T}$. That is, two rooted trees have isometric boundaries if and only if their cores are isomorphic as graphs.*

The flow theorem implies the following.

Corollary 5. *The transience or recurrence of a tree depends only on its boundary (or equivalently on its core). That is, if two trees have isometric boundaries then they are either both recurrent or both transient.*

3.6 Logarithmic capacity and recurrence

Short digresion: There's a beautiful criterion due to Kakutani which answers the question of whether or not a Brownian motion in $\mathbb{R}^2$ will hit a compact set K with positive probability or not. Somewhat surprisingly it is not necesary for K to have positive measure for this to be the case. The sharp criterion is given by whether or not K can hold a distribution of electric charge in such a way that the potential energy (created by the repulsion between charges of the same sign) is finite. In this formulation charges are considered to be infinitely divisible so that if one has a unit of charge at a single point then the potential energy is infinite (because one can consider that one has two half charges at the same spot, and they will repulse each other infinitely). Sets with no capacity to hold charges (such as a single point, or a countable set) will not be hit by Brownian motions, but sets with positive capacity will. End digresion.

Definition 2. *Let (X, d) be a compact metric space. Then (X, d) is is said to be Polar (or have zero capacity) if and only if for every probability measure μ on X one has*

$$\int -\log(d(x,y))\mathrm{d}\mu(x)\mathrm{d}\mu(y) = \infty.$$

Otherwise the space is said to have positive capacity. In general the capacity is defined as (I hope I got all the signs right, another solution would be to let the potential energy above be negative and use a supremum in the formula below)

$$\mathcal{C}(X,d) = \exp\left(-\inf\left\{\int -\log(d(x,y))\mathrm{d}\mu(x)\mathrm{d}\mu(y)\right\}\right)$$

where the infimum is over all probability measures on X.

There is a correspondence between finite energy source flows on a tree and finite measures on the boundary. Furthermore the energy of a source flow has a simple relationship to the above type of integral for the corresponding measure.

Theorem 13. *Let T be a rooted tree with root a. Given a unit source flow φ with source a there is a unique probability measure μ on ∂T such that for each edge $e \in E(T)$ the measure of the set of paths in ∂T which contain e is $\varphi(e)$. This correspondence between unit flows and probability measures is bijective. Furthermore, one has*

$$\mathcal{E}(\varphi) = \int -\log(d(x,y))\mathrm{d}\mu_\varphi(x)\mathrm{d}\mu_\varphi(y).$$

Proof. For each $e \in E(T)$ let $[e]$ denote the subset of ∂T consisting of paths containing the edge e. Notice that $\partial T \setminus [e] = \bigcup_{i=1}^{n}[e_i]$ where the e_i are the remaining edges at the same level (i.e. joining vertices at the same distance from a) as e. Also the sets $[e]$ are compact so that if we write $[e]$ as a countable union of disjoint sets $[e_i]$ then in fact the union is finite. Since this implies that $[e] \mapsto \varphi(e)$ is countably additive on the algebra of sets generated by the $[e]$ one obtains by Caratheodory's extension theorem that it extends to a unique probability measure.

The inverse direction is elementary. Given a measure μ on ∂T one defines $\varphi(e) = \mu([e])$ and verifies that it is a unit flow simply by additivity of μ.

The important part of the statement is the relationship between the energy of the flow φ and the type of integral used to define the capacity of ∂T.

To prove this we use a well known little trick in probability. It's simply the observation that if X is a random variable taking value in the non-negative integers then

$$\mathbb{E}[X] = \sum_n \mathbb{P}[X \geq n].$$

The random variable we will consider is $X = -\log(d(x,y))$ where x and y are independent elements of ∂T with distribution μ. Notice that $\mathbb{E}[X]$ is exactly the integral we're interested in. Our result follows from the observation that

$$\mathbb{P}[X \geq n] = \sum_{\{e \in E(X): d(e_-,a)=n-1, d(e_+,a)=n\}} \varphi(e)^2$$

which is simply the sum of probabilities of all the different ways two infinite paths can coincide up to distance n from a. □

The folowing corollary is immediate (notice that team Lyons has the ball and Terry has passed it over to Russell).

Corollary 6 (Russell Lyons (1992))**.** *The simple random walk on a tree T is transient if and only if ∂T has positive capacity. It is recurrent if and only if ∂T is polar.*

3.7 Average meeting height and recurrence

One can also state the criterion for recurrence of trees is in more combinatorial terms using the fact that $\int -\log(d(x,y))\mathrm{d}\mu(x)\mathrm{d}\mu(y)$ is the expected value of the distance from a of the first vertex where the paths x and y separate. To formalize this, given two vertices $c, d \in T$ define their meeting height $m(c,d)$ to be the distance from a at which the paths joining c to a and d to a first meet. The average meeting height of a set $b_1, \ldots, b_n$ of vertices is defined as

$$\frac{1}{\binom{n}{2}} \sum_{i<j} m(b_i, b_j).$$

Corollary 7 (Benjamini and Peres (1992))**.** *The simple random walk on a tree T is transient if and only if there is a finite constante $C > 0$ such that for any finite n there are n vertices $b_1, \ldots, b_n \in T$ whose average meeting height is less than C.*

Proof. Suppose there is a finite unit flow φ and let $x_1, \ldots, x_n, \ldots$ be random independent paths in ∂T with distribution μ_φ. Then for each n one has

$$\mathbb{E}\left[\frac{1}{\binom{n}{2}} \sum_{1 \le i < j \le n} m(x_i, x_j)\right] = \mathcal{E}(\varphi).$$

Therefore there is a positive probability that the average meeting distance between the first n random paths (or suitably chosen vertices on them) will be less than $C = 2\mathcal{E}(\varphi)$.

In the other direction suppose there exists C and for each n one has paths $x_{n,1}, \ldots, x_{n,n}$ in ∂T such that their average meeting distance is less than C. Then any weak limit[2] μ of the sequence of probabilities

$$\mu_n = \frac{1}{n} \sum \delta_{x_{n,i}}$$

will satisfy $\int -\log(d(x,y))\mathrm{d}\mu(x)\mathrm{d}\mu(y) \le C$ (here δ_x denotes the probability giving total mass to x). Which implies that the boundary has positive capacity and that the tree is transient. □

3.8 Positive Hausdorff dimension implies transience

Some readers might be familiar with Hausdorff dimension, which is a notion of size for a compact metric space for which one obtains for example that the dimension of the unit interval $[0, 1]$ is 1 while that of the middle thirds Cantor set is $\log(2)/\log(3) \approx 0.631$. The boundary of a tree usually looks somewhat like a Cantor set, so this gives a superficial reason why dimension might be relevant in our context.

Definition 3. *The d-dimensional Hausdorff content of a compact metric space (X, d) is the infimum of*

$$\sum r_i^d$$

over all sequences of positive diameters r_i such that X can be covered by balls with those diameters.

The Hausdorff dimension of X is the infiumum over all $d > 0$ such that the d-dimensional content of X is 0.

Consider the task of proving that the Hausdorff dimension of $[0, 1]$ is 1. First one notices that one can easily cover $[0, 1]$ by n balls of length $r_n = 100/n$. Hence if $d > 1$ one gets

$$nr_n^d = n(100/n)^d \to 0.$$

[2]We haven't discussed weak convergence in these notes. The things every person on the planet should know about this notion are: First, that it's the notion that appears in the statement of the Central Limit Theorem. And, second, that a good reference is Billinsgley's book.

This shows that the dimension of $[0,1]$ is less than or equal to 1 (the easy part). But how does one show that it is actually 1? This is harder because one needs to control all possible covers not just construct a particular sequence of them.

The trick is to use Lebesgue measure μ on $[0,1]$. The existence of Lebesgue measure (and the fact that it is countably additive) implies that if one covers $[0,1]$ by intervals of length r_i then $\sum r_i \geq 1$. Hence the 1-dimensional Hausdorff content of $[0,1]$ is positive (greater than or equal to 1) and one obtains that the dimension must be 1.

For compact subsets of $\mathbb{R}^n$ the above trick was carried out to its maximal potential in a Phd. thesis from the 1930s. The result works for general compact metric spaces and is now known by name of the author of said thesis.

Lemma 3 (Frostman's Lemma)**.** *The d-dimensional Hausdorff content of (X,d) is positive if and only if there exists a probability measure μ on X satisfying*

$$\mu(B_r) \leq Cr^d$$

for some constant C and all balls B_r of radius $r > 0$.

An immediate corollary is the following (the original proof is more elementary, the paper is very readable and recommended).

Corollary 8 (Russel Lyons 1990)**.** *Let T be a tree whose boundary has positive Hausdorff dimension. Then the simple random walk on T is transient.*

Proof. By Frostman's theorem there's a probability μ on the ∂T such that $\mu(B_r) \leq Cr^d$ for every ball B_r of radius $r > 0$. This allows one to estimate the integral defining capacity.

To see this, fix x and notice that the set of y such that $-\log(d(x,y)) \geq t$ is exactly the ball $B_{e^{-t}}(x)$ (radius e^{-t} and centered at x). This implies that (by a variant of the trick we used in the proof of Theorem 13)

$$\int -\log(d(x,y))\mathrm{d}\mu(y) \leq \int_0^{+\infty} \mu(\{y : -\log(d(x,y)) \geq t\}\mathrm{d}t \leq \int_0^{+\infty} Ce^{-dt}\mathrm{d}t < +\infty.$$

Since this is valid for all x one may itegrate over x to obtain that ∂T has positive capacity. □

3.9 Perturbation of trees

Our final corollary, combined with the exercises below, gives a pretty good idea of how much one has to modify a tree to change the behavior of the simple random walk on it.

Corollary 9. *If T and T' are trees whose boundaries are Hölder homeomorphic (i.e. there is a homeomorphism $f : \partial T \to \partial T'$ satisfying $d(f(x), f(y)) \leq Cd(x,y)^\alpha$ for some constants $C, \alpha > 0$) then the random walk on T is transient if and only if the one on T' is.*

Exercise 6. *Prove that if T is a tree and $a, b \in T$ then the boundaries one obtains by considering a and b as the root vertex are Lipschitz homeomorphic and the Lipschitz constant is no larger than $e^{d(a,b)}$.*

Exercise 7. *Prove that if one subdivides the edges of a tree into no more than N smaller edges then the boundary of the resulting tree is homeomorphic to the original via a Hölder homeomorphism.*

Exercise 8. *Show that if T and T' are quasi-isometric trees then their boundaries are homeomorphic via a Hölder homeomorphism.*

3.10 Notes for further reading

The original article by Terry Lyons [Lyo83] gives a relatively short proof of the flow theorem which is based on the Martingale convergence theorem. Reading this article one also learns of the flow theorem's origins. Apparently, the theorem was discovered as the translation to the discrete setting of a result from the theory of Riemann surfaces. In short, it's not to wrong to say that the flow theorem for Brownian motion preceeded and inspired the corresponding theorem for discrete random walks.

Our treatment of the flow theorem beginnning with finite graphs is based on the very clearly written article by Levin and Peres [LP10]. Like them, we have avoided the language of electrical networks even though it is a very important topic in the subject (introduced by Crispin Nash-Williams in one of the first interesting works on random walks on general infinite graphs [NW59]).

The many interesting results about trees are due to Russell Lyons ([Lyo90] and [Lyo92], and Yuval Peres and Itai Benjamini [BP92].

The focus on quasi-isometry comes mostly from the more geometric works on Riemannian manifolds (here one should mention the Furstenberg-Lyons-Sullivan proceedure which allows one to associated a discrete random walk to a Brownian motion on a Riemannian manifold [LS84][3]). The original paper by Kanai is [Kan86].

For Frostman's lemma and much more geometric measure theory the reader can have a look at [Mat95].

[3] There's a story explaining why Furstenberg is usually included in the name of this discretization proceedure. It turns out Furstenberg once wanted to prove that $SL(2, \mathbb{Z})$ was not a lattice in $SL(3, \mathbb{R})$, and there was a very simple proof available using Kazhdan's property T, which was well known at the time. But Furstenberg didn't know about property T, so he concocted the fantastic idea of trying to translate the clearly distinct behavior of Brownian motion on $SL(2, \mathbb{R})$ and $SL(3, \mathbb{R})$ to the discrete random walks on the corresponding lattices. This was done mid-proof, but has all the essential elements later refined in Sullivan and Lyons' paper (see [Fur71]).

4 The classification of recurrent groups

4.1 Sub-quadratic growth implies recurrence

We've seen as a consequence of the flow theorem that the simple random walk behaves the same way (with respect to recurrence and transience) on all Cayley graphs of a given group G. Hence one can speak of recurrent or transient groups (as opposed to pairs (G, F)).

In this section G will always denote a group and F a finite symmetric generator of G. The Cayley graph will be denoted by G abusing notation slightly and we keep with the notation $E(G)$ for the set of (directed) edges of the graph.

We introduce a new notation which will be very important in this section. Given a set $A \subset G$ (or in any graph) we denote by ∂A the edge boundary of A, i.e. the set of outgoing edges (edges starting in A and ending outside of A).

Let B_n denote the ball of radius n centered at the identity element of G. That is, it is the set of all elements of the group which can be written as a product of n or less elements of F. A first consequence of the flow theorem (which in particular implies the recurrence of $\mathbb{Z}^2$ and of all wallpaper groups such as *632) is the following:

Corollary 10. *If there exists $c > 0$ such that $|B_n| \le cn^2$ for all n then the group G is recurrent.*

Proof. Suppose φ is a unit source flow with source the identity element of G. Notice that for all n one has

$$\sum_{e \in \partial B_n} \varphi(e) = 1.$$

Using Jensen's inequality (in the version "average of squares is greater than square of the average") one gets

$$\frac{1}{|\partial B_n|} \sum_{e \in \partial B_n} \varphi(e)^2 \ge \left(\frac{1}{|\partial B_n|} \sum_{e \in \partial B_n} \varphi(e) \right)^2,$$

so that

$$\sum_{e \in \partial B_n} \varphi(e)^2 \ge \frac{1}{|\partial B_n|}$$

for all n. Hence it suffices to show that $\sum 1/|\partial B_n| = +\infty$ to conclude that the flow has infinite energy (notice that for the standard Cayley graph associated to $\mathbb{Z}^2$ one can calculate $|\partial B_n| = 4 + 8n$). Here the growth hypothesis must be used and we leave it to the reader (see the exercise below and notice that $|F|(|B_{n+1}| - |B_n|) \ge |\partial B_n|$). □

Exercise 9. *Let $a_k, k = 1, 2, \dots$ be a sequence of positive numbers such that for some positive constant c the inequality*

$$\sum_{k=1}^{n} a_n \le cn^2$$

holds for all n. Prove that

$$\sum_{k=1}^{+\infty} \frac{1}{a_n} = +\infty.$$

4.2 Growth of groups

The function $n \mapsto |B_n|$ is sometimes called the growth function of the group G. It depends on the chosen set of generators F but the functions associated to two distinct generators are comparable up to a multiplicative constant. Hence we can roughly classify groups into the following categories:

1. Polynomial growth: A group is said to have polynomial growth of degree d if $cn^d \le |B_n| \le Cn^d$ for all n and some constants $c, C > 0$. It is known that having only the upper bound is sufficient to have a lower bound with the same exponent d. Also, d is always an integer. This is a consequence of a famous theorem from the 1980s due to Gromov.

2. Exponential growth: A group is said to have exponential growth if $c\lambda^n \le |B_n|$ for all n and some constants $c > 0, \lambda > 1$. Notice that $|B_n| \le (1 + |F|)^n$ so no group can grow faster than exponentially.

3. Intermediate growth: The first examples of a group with subexponential but super-polynomial growth where discovered in the early 1980s by Grigorshuk. We will not discuss any of these examples in these notes.

We have seen that a group with polynomial growth of degree d less than or equal to 2 is recurrent. Our objective in this section is to prove that a group with $|B_n| \ge cn^3$ for some $c > 0$ and all n must be transitive. These two cases cover all posibilities by the theorem of Gromov classifying groups of polynomial growth. Recall that a subgroup H of a group G is said to have finite index k if there exist $g_1, \ldots, g_k \in G$ such that every element of G can be written as hg_i for some $h \in H$ and some $i = 1, \ldots, k$. The final result is the following:

Theorem 14 (Varopoulos+Gromov)**.** *A group G is recurrent if and only if it has polynomial growth of degree less than or equal to 2. This can only happen if the group is either finite, has a subgroup isomorphic to $\mathbb{Z}$ with finite index, or has a subgroup isomorphic to $\mathbb{Z}^2$ with finite index.*

This is an example of a result whose statement isn't very interesting (it basically says that one shouldn't study recurrence of Cayley graphs since it's too strong of a property) but for which the ideas involved in the proof are very interesting (the flow theorem, isoperimetric inequalities, and Gromov's theorem on groups of polynomial growth).

We will not prove the theorem of Gromov which leads to the final clasification. Only the fact that growth larger than cn^3 implies transience.

4.3 Examples

Recall our list of examples from the first section: The d-dimensional grid $\mathbb{Z}^d$, the free group in two generators $\mathbb{F}_2$, the Modular group of fractional linear transformations Γ, the wallpaper group $*632$, the Heisenberg group Nil, and the lamplighter group $Lamplighter(\mathbb{Z})$.

It is relatively simple to establish that $\mathbb{Z}^d$ has polynomial growth of degree d, that the free group in two generators has exponential growth, that the wallpaper group has polynomial growth of degree $d = 2$, and that the lamplighter group has exponential growth (with $2n$ moves the lamplighter can light the first n-lamps in any of the 2^n possible on-off combinations).

It turns out that the modular group has exponential growth. To see this it suffices to establish that the subgroup generated by $z \mapsto z + 2$ and $z \mapsto \frac{-2}{2z+1}$ is free (this subgroup is an example of a "congruence subgroup" which are important in number theory, or so I've been told... by wikipedia). We leave it to the interested reader to figure out a proof (several are possible, either by trying to find an explicit fundamental domain on the action of the upper half plane of $\mathbb{C}$, by combinatorial analysis of the coefficients of compositions, or by a standard argument for establishing freeness of a group called the Ping-Pong argument which the reader can Google and learn about quite easily, there's even a relevant post in Terry Tao's blog).

In view of Gromov's theorem (which again, we will neither state fully nor prove), working out the growth of the Heisenberg group and variants of it (which was first done by Bass and Guivarc'h) turns out to be a key point in the proof of Theorem 14. Hence any time the reader spends thinking about this issue is well spent.

Exercise 10. *Show that there exists $c > 0$ such that $cn^4 \leq |B_n|$ for all n on the Heisenberg group Nil.*

4.4 Isoperimetric inequalities

There exist graphs with exponential growth (for the natural definition) which are recurrent. The simplest example is perhaps the Canopy tree (notice that the ball of radius $2n$ centered at a leaf of the Canopy tree contains 2^n leaves). So clearly if growth is to characterize recurrence of Cayley graphs there must be some special property of these graphs to explain this which does not hold in general. The relevant property is characterized by the so-called isoperimetric inequalities.

A Cayley graph is said to satisfy the strong isoperimetric inequality if there is a positive constant c such that

$$c|A| \leq |\partial A|$$

for all finite subsets $A \subset G$. Notice that the strong isoperimetric inequality implies exponential growth. The free group satisfies the strong isoperimetric inequality. The existence of a finite index free subgroup implies that the Modular group also satisfies the strong isoperimetric inequality (there are other ways of showing this, in fact the strong isoperimetric inequality is equivalent to a property called non-amenability which has a near-infinite list of equivalent definitions which are not trivially equivalent, some of these definitions are simple to verify on the Modular group).

A Cayley graph is said to satisfy the d-dimensional isoperimetric inequality if there is a positive constant c such that

$$c|A|^{d-1} \leq |\partial A|^d$$

for all finite subsets $A \subset G$. An analogous relationship with $d = 3$ is satisfied between surface area and volume of a three dimensional submanifold in $\mathbb{R}^3$.

The main point we will use which distinguishes Cayley graphs from general graphs is the following:

Lemma 4. *If there exists $c > 0$ such that $|B_n| \geq cn^d$ for all n then G satisfies an isoperimetric inequality of dimension d.*

Proof. Fix a finite set $A \subset G$ and let $f = 1_A$ be the indicator function of A (i.e. $f = 1$ on A and $f = 0$ on $G \setminus A$). Notice that

$$|\nabla f|_1 = \sum_{e \in E(G)} |\nabla f| = 2|\partial A|,$$

so the point is to bound the L^1 norm of the gradient of f (i.e. the above sum) from below in terms of $|A|$.

On the other hand

$$|f|_1 = \sum_{x \in G} |f(x)| = |A|$$

so the objective is to bound $|\nabla f|_1$ from below in terms of $|f|_1$.

In order to use the growth hypothesis we will introduce an auxiliary function $\tilde{f}$ constructed by averaging f over suitable balls and in fact bound $|\tilde{f} - f|$ from above and below in terms of $|f|_1$ and $|\nabla f|_1$ respectively.

The function $\tilde{f}$ is defined by letting n be the smallest integer such $|B_n| \geq 2|A|$ and setting

$$\tilde{f}(x) = \frac{1}{|B_n|} \sum_{g \in B_n} f(xg).$$

The point of the choice of n is that if $x \in A$ then $f(x) = 1$ but $\tilde{f}(x) \leq 1/2$. Hence the L^1 norm of $\tilde{f} - f$ is bounded from below as follows

$$\frac{1}{2}|A| \leq |\tilde{f} - f|_1.$$

If we can bound $|\tilde{f} - f|_1$ from above using $|\nabla f|_1$ then we're done.

To acomplish this notice that if $g \in F$ then

$$\sum_{x \in G} |f(xg) - f(x)| \leq |\nabla f|_1.$$

A slight generalization (left to the reader; Hint: triangle inequality) is that if $g = g_1 \cdots g_k$ for some sequence of $g_i \in F$ then

$$\sum_{x \in G} |f(xg) - f(x)| \leq k|\nabla f|_1.$$

Using this one can bound the L^1 norm of $\tilde{f} - f$ from above in terms of $|\partial A|$ as follows

$$\sum_{x \in G} |\tilde{f}(x) - f(x)| \leq \frac{1}{|B_n|} \sum_{g \in B_n} \sum_{x \in G} |\tilde{f}(xg) - f(x)| \leq n|\nabla f| = 2n|\partial A|.$$

Hence we have obtained

$$\frac{1}{2}|A| \leq |\tilde{f} - f|_1 \leq 2n|\nabla f|_1 = 4n|\partial A|.$$

The result follows by noticing that the growth hypothesis implies that $n \leq C|A|^{1/d}$ where $C = c^{-1/d}$. $\square$

4.5 Growth of order 3 or more implies transience

Lemma 5. *Let G be a group satisfying a 3-dimensional isoperimetric inequality. Then G is transient.*

Proof. Let a be the identity element of the group G and define for each n a function $f_n : G \to [0, 1]$ by

$$f(x) = \mathbb{P}\left[d(a, x_n) \geq n \text{ before } x_n = a\right]$$

where x_n is a simple random walk on the Cayley graph starting at x.

Recall that (from the finite flow theorem) one had that ∇f_n is a flow with source at 1 and sinks at the points with $d(a, x) = n$ and that

$$0 < D_n = \operatorname{div}(\nabla f_n)(a) = \mathcal{E}(\nabla f_n).$$

Notice that since f_n takes values in $[0, 1]$ one has $D_n \leq \deg(a) = |F|$ (recall F is a finite symmetric generating set which was used to define the Cayley graph). If one could bound D_n from below by a positive constant then this would imply that (taking a limit of a subsequence of ∇f_n) there is a finite energy source flow on G and hence G is transient as claimed.

To acomplish this let n be fixed and define a finite sequence of subsets beginning with $A_1 = \{a\}$ using the rule that if $A_k \subset B_{n-1}$ then

$$A_{k+1} = A_k \cup \left\{ e_+ : e \in \partial A_k, \nabla f_n(e) \leq \frac{2D_n}{|\partial A_k|} \right\}.$$

The sequence stops the first time A_k contains a point at distance n from A, let N be the number of sets in the thus obtained sequence.

The fact that the sequence stops follows because if $A_k \subset B_{n-1}$ one has

$$\sum_{e \in \partial A_k} \nabla f(e) = D_n,$$

so at least half of the edges in A_k must lead to a point in A_{k+1}.

The point of the definition is that there is a point $x \in A_N$ with $d(a,x) = n$ so that choosing a path $e_1, \ldots, e_{N-1}$ with $e_{1-} = a$, $e_{(N-1)+} = x$ and $e_i \in \partial A_i$ for $i = 1, \ldots, N-1$ one obtains

$$1 = f(x) \le \sum_{i=1}^{k} \frac{2D_n}{|\partial A_k|}.$$

Hence it suffices to bound

$$\sum_{k=1}^{N} \frac{1}{|\partial A_k|}$$

from above (in a way which doesn't depend on n) to obtain a uniform lower bound for D_n and hence prove that G is transient.

Here we use the 3-dimensional isoperimetric inequality, the fact we had noted before that if $k < N$ then at least half the edges of ∂A_k lead to A_{k+1}, and the fact that at most $|F|$ edges can lead to any given vertex. Combining these facts we obtain

$$|A_{k+1}| - |A_k| \ge \frac{1}{2|F|}|\partial A_k| \ge c|A_k|^{\frac{2}{3}}$$

where $c > 0$ is the constant in the isoperimetric inequality divided by $2|F|$.

This implies (see the exercise below) that $|A_k| \ge c^3k^3/343$. Hence

$$\sum_{k=1}^{N} \frac{1}{|\partial A_k|} \le \sum_{k=1}^{+\infty} \frac{49}{c^3k^2} = \frac{49\pi^2}{6c^3}$$

where the exact sum in the final equation was included on a whim since only convergence is needed to obtain the desired result. □

Exercise 11. *Let $1 = x_1 < x_2 < \cdots < x_N$ be a sequence satisfying $x_{k+1} - x_k \ge cx_k^{\frac{2}{3}}$ for all $k = 1, \ldots, N-1$. Prove that $x_k \ge c^3k^3/343$ for all $k = 1, \ldots, N$, or at least prove that there exists a constant $\lambda > 0$ depending only on c such that $x_k \ge \lambda^3k^3$.*

4.6 Notes for further reading

As far as I know the first serious study of random walks on discrete groups was carried out by Kesten in his thesis [Kes59b]. He was also the first to prove that the strong isoperimetric inequality implies exponential decay of return probabilities and hence transience (see [Kes59a] though this article is hard to find and even harder to read, at least for me). This idea was later on refined and improved by several people, the reader might want to check out Gerl's article [Ger88] and the more recent work by Virág [Vir00].

Kesten was also the first to introduce the idea that growth might determine recurrence or transience (see [Kes67]) and the idea that polynomial growth of order 2 was equivalent to recurrence is sometimes known as Kesten's conjecture.

I find it interesting that Kesten's conjecture was first established for continuous groups (see [Bal81] which is the culmination of a series of works by several people including Baldi, Guivarc'h, Keane, Roynette, Peyrière, and Lohoué).

The idea of using d-dimensional isoperimetric inequalities for estimating return probabilities was introduced quite successfully into the area of random walks on discrete groups by Varopoulos in the mid-80s. The main result is that a d-dimensional isoperimetric inequality implies a decay of return probabilities of the order of $n^{-d/2}$ (in particular if $d \geq 3$ the series converges and the walk is transient) which was proved in [Var85].

Instead of proving Varopoulos' decay estimates we borrowed the proof given in Mann's excellent book [Man12] that growth implies isoperimetric inequalities on finitely generated groups, and then proved that an isoperimetric inequality of degree 3 or more implies transience using an argument from a paper by Benjamini and Kozma [BK05].

There are several good references for Gromov's theorem including Gromov's original paper (where the stunning idea of looking at a discrete group from far away to obtain a continuous one is introduced), Mann's book [Man12], and even Tao's blog.

References

[Ash00] Robert B. Ash. *Probability and measure theory.* Harcourt/Academic Press, Burlington, MA, second edition, 2000. With contributions by Catherine Doléans-Dade.

[Bal81] P. Baldi. Caractérisation des groupes de Lie connexes récurrents. *Ann. Inst. H. Poincaré Sect. B (N.S.)*, 17(3):281–308, 1981.

[Ben13] Itai Benjamini. *Coarse geometry and randomness*, volume 2100 of *Lecture Notes in Mathematics.* Springer, Cham, 2013. Lecture notes from the 41st Probability Summer School held in Saint-Flour, 2011, Chapter 5 is due to Nicolas Curien, Chapter 12 was written by Ariel Yadin, and Chapter 13 is joint work with Gady Kozma, École d'Été de Probabilités de Saint-Flour. [Saint-Flour Probability Summer School].

[BK05] Itai Benjamini and Gady Kozma. A resistance bound via an isoperimetric inequality. *Combinatorica*, 25(6):645–650, 2005.

[BL85] Philippe Bougerol and Jean Lacroix. *Products of random matrices with applications to Schrödinger operators*, volume 8 of *Progress in Probability and Statistics.* Birkhäuser Boston, Inc., Boston, MA, 1985.

[Bla55] David Blackwell. On transient Markov processes with a countable number of states and stationary transition probabilities. *Ann. Math. Statist.*, 26:654–658, 1955.

[BP92] Itai Benjamini and Yuval Peres. Random walks on a tree and capacity in the interval. *Ann. Inst. H. Poincaré Probab. Statist.*, 28(4):557–592, 1992.

[Doo89] J. L. Doob. Kolmogorov's early work on convergence theory and foundations. *Ann. Probab.*, 17(3):815–821, 1989.

[Doo96] Joseph L. Doob. The development of rigor in mathematical probability (1900–1950) [in *development of mathematics 1900–1950 (luxembourg, 1992)*, 157–170, Birkhäuser, Basel, 1994; MR1298633 (95i:60001)]. *Amer. Math. Monthly*, 103(7):586–595, 1996.

[Ein56] Albert Einstein. *Investigations on the theory of the Brownian movement*. Dover Publications, Inc., New York, 1956. Edited with notes by R. Fürth, Translated by A. D. Cowper.

[Fur63] Harry Furstenberg. Noncommuting random products. *Trans. Amer. Math. Soc.*, 108:377–428, 1963.

[Fur71] Harry Furstenberg. Random walks and discrete subgroups of Lie groups. In *Advances in Probability and Related Topics, Vol. 1*, pages 1–63. Dekker, New York, 1971.

[Ger88] Peter Gerl. Random walks on graphs with a strong isoperimetric property. *J. Theoret. Probab.*, 1(2):171–187, 1988.

[Gri09] Alexander Grigor'yan. *Heat kernel and analysis on manifolds*, volume 47 of *AMS/IP Studies in Advanced Mathematics*. American Mathematical Society, Providence, RI; International Press, Boston, MA, 2009.

[Hsu02] Elton P. Hsu. *Stochastic analysis on manifolds*, volume 38 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2002.

[Kai92] Vadim A. Kaimanovich. Measure-theoretic boundaries of Markov chains, 0-2 laws and entropy. In *Harmonic analysis and discrete potential theory (Frascati, 1991)*, pages 145–180. Plenum, New York, 1992.

[Kan86] Masahiko Kanai. Rough isometries and the parabolicity of Riemannian manifolds. *J. Math. Soc. Japan*, 38(2):227–238, 1986.

[Kes59a] Harry Kesten. Full Banach mean values on countable groups. *Math. Scand.*, 7:146–156, 1959.

[Kes59b] Harry Kesten. Symmetric random walks on groups. *Trans. Amer. Math. Soc.*, 92:336–354, 1959.

[Kes67] H. Kesten. The Martin boundary of recurrent random walks on countable groups. In *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66)*, pages Vol. II: Contributions to Probability Theory, Part 2, pp. 51–74. Univ. California Press, Berkeley, Calif., 1967.

[Kin93] J. F. C. Kingman. *Poisson processes*, volume 3 of *Oxford Studies in Probability*. The Clarendon Press, Oxford University Press, New York, 1993. Oxford Science Publications.

[Kol50] A. N. Kolmogorov. *Foundations of the Theory of Probability*. Chelsea Publishing Company, New York, N. Y., 1950.

[Law13] Gregory F. Lawler. *Intersections of random walks*. Modern Birkhäuser Classics. Birkhäuser/Springer, New York, 2013. Reprint of the 1996 edition.

[LG13] Jean-François Le Gall. *Mouvement brownien, martingales et calcul stochastique*, volume 71 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer, Heidelberg, 2013.

[Lov96] L. Lovász. Random walks on graphs: a survey. In *Combinatorics, Paul Erdős is eighty, Vol. 2 (Keszthely, 1993)*, volume 2 of *Bolyai Soc. Math. Stud.*, pages 353–397. János Bolyai Math. Soc., Budapest, 1996.

[LP10] David A. Levin and Yuval Peres. Pólya's theorem on random walks via Pólya's urn. *Amer. Math. Monthly*, 117(3):220–231, 2010.

[LP15] Russel Lyons and Yuval Peres. *Random walks on infinite graphs and groups*. 2015.

[LS84] Terry Lyons and Dennis Sullivan. Function theory, random paths and covering spaces. *J. Differential Geom.*, 19(2):299–323, 1984.

[Lyo83] Terry Lyons. A simple criterion for transience of a reversible Markov chain. *Ann. Probab.*, 11(2):393–402, 1983.

[Lyo90] Russell Lyons. Random walks and percolation on trees. *Ann. Probab.*, 18(3):931–958, 1990.

[Lyo92] Russell Lyons. Random walks, capacity and percolation on trees. *Ann. Probab.*, 20(4):2043–2088, 1992.

[Man12] Avinoam Mann. *How groups grow*, volume 395 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2012.

[Mat95] Pertti Mattila. *Geometry of sets and measures in Euclidean spaces*, volume 44 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995. Fractals and rectifiability.

[MP10] Peter Mörters and Yuval Peres. *Brownian motion*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2010. With an appendix by Oded Schramm and Wendelin Werner.

[NW59] C. St. J. A. Nash-Williams. Random walk and electric currents in networks. *Proc. Cambridge Philos. Soc.*, 55:181–194, 1959.

[Pet06] Peter Petersen. *Riemannian geometry*, volume 171 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2006.

[Spi76] Frank Spitzer. *Principles of random walk*. Springer-Verlag, New York-Heidelberg, second edition, 1976. Graduate Texts in Mathematics, Vol. 34.

[Str00] Daniel W. Stroock. *An introduction to the analysis of paths on a Riemannian manifold*, volume 74 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2000.

[SV06] Glenn Shafer and Vladimir Vovk. The sources of Kolmogorov's *grundbegriffe*. *Statist. Sci.*, 21(1):70–98, 2006.

[Var85] N. Th. Varopoulos. Isoperimetric inequalities and Markov chains. *J. Funct. Anal.*, 63(2):215–239, 1985.

[Var01] S. R. S. Varadhan. *Probability theory*, volume 7 of *Courant Lecture Notes in Mathematics*. New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2001.

[Var07] S. R. S. Varadhan. *Stochastic processes*, volume 16 of *Courant Lecture Notes in Mathematics*. Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2007.

[Vir00] B. Virág. Anchored expansion and random walk. *Geom. Funct. Anal.*, 10(6):1588–1605, 2000.

[Woe00] Wolfgang Woess. *Random walks on infinite graphs and groups*, volume 138 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2000.

Index