RANDOM WALKS ON GROUPS: CHARACTERS AND GEOMETRY

PERSI DIACONIS

Departments of Mathematics and Statistics, Stanford University, Stanford, CA 94305 USA

Introduction

These notes tell two stories. The first is an overview of a general approach to studying random walk on finite groups. This involves the character theory of the group and the geometry of the group in various generating sets. The second is the life and times of a single example: random transpositions on the symmetric group. This was the first example where sharp estimates were obtained. We will prove it takes $\frac{1}{2}n\log n$ transpositions to mix up n cards. This example gives rise to a rich comparison theory allowing general walks to be studied. Its history goes back to Hurwitz's 1891 work on counting branched covers. There are generalizations to finite groups of Lie type, various deformations (Jack symmetric functions and Hecke algebras) and applications to diffusion, phylogeny and coagulation processes in chemistry.

Let us begin with a definition. Let G be a finite group. Let $\{Q(g)\}_{g\in G}$, be a probability distribution on G. Thus $Q(g)\geq 0$ and $\sum_g Q(g)=1$. This is the basic data. Define convolution by

$$Q\star Q(g)=\sum_h Q\left(gh^{-1}\right)Q(h), \qquad Q^{\star k}(g)=Q^{\star(k-1)}\star Q(g).$$

Thus $Q\star Q(g)$ is the chance that a random walk on G generated by picking elements repeatedly with weight $Q(g)$ and multiplying is at g after two steps; some element h must have been chosen first, followed by gh^{-1}. Similarly, $Q^{\star k}(g)$ is the chance that the walk is at g after k steps. All walks start at the identity.

Under a mild restriction (the support of Q is not contained in a coset of a subgroup) the random walk converges to the uniform distribution $u(g)=1/|G|$.

$$Q^{\star k}(g)\to\frac{1}{|G|} \qquad \text{as } k\to\infty.$$

The question is how fast? We will measure convergence by the total variation distance

$$\|Q^{\star k}-u\|=\frac{1}{2}\sum_g\left|Q^{\star k}(g)-\frac{1}{|G|}\right|=\max_{A\subset G}\left|Q^{\star k}(A)-\frac{|A|}{|G|}\right|. \tag{0.1}$$

We thus arrive at a well-posed math problem:

Given a finite group G, a probability measure $Q(g)$ and $c>0$, how large should k be so that

$$\|Q^{\star k}-u\|\leq c?$$

For group theorists, this may be rephrased as follows. Identify Q with an element $q = \sum_g g\,Q(g)$ of the group algebra. The coefficient of g in q^k is $Q^{\star k}(g)$. High powers of q converge to $\frac{1}{|G|}\sum_g g$. We are asking for the speed of convergence in the ℓ^1 norm.

Here is our leading example, said more carefully.

Example (Random transpositions). Imagine a deck of n playing cards face down in an ordered row with card 1 at the left, card n at the right. The cards are repeatedly mixed by the following operation. The left hand touches a random card. The right hand touches a random card (so left = right with probability $1/n$). These cards are transposed (nothing is done if left = right). It is intuitively clear that after many switches the row of cards is all mixed up.

More mathematically, on the symmetric group S_n, let

$$Q(\pi) = \begin{cases} 1/n \text{ if } \pi = id \\ 2/n^2 \text{ if } \pi \text{ is a transposition} \\ 0 \qquad \text{otherwise} \end{cases} \tag{0.2}$$

Repeated switches are modeled by $Q^{\star k}(\pi)$. The following theorem was proved in joint work with Shahshahani [26].

Theorem A *For the random walk generated by* (0.2) *on the symmetric group* S_n, *if* $k = \frac{1}{2}n(\log n + c)$ *for* $c > 0$, *then*

$$||Q^{\star k} - u|| \leq 6e^{-c} \tag{0.3}$$

conversely, if $k = \frac{1}{2}n(\log n - c)$ *there is* $\epsilon > 0$ *such that* $||Q^{\star k} - u|| > \epsilon$ *for all* n.

Remarks

1. When $n = 52$, $\frac{1}{2}n\log n \doteq 103$; to make the distance in (0.3) smaller than 1/100, requires $k \doteq 270$ switches.
2. The theorem shows that convergence to stationarity has a cutoff or threshold at $\frac{1}{2}n\log n$. The transition from order to chaos happens as c varies. This often happens for random walk on non-commutative groups. See [9], [65] for further discussion; it is one of the important open problems of this subject to understand the cutoff phenomenon.
3. The original motivation for studying random transpositions is worth mentioning. In the course of a large scale study of optimal strategies in experiments with feedback ([13]), a Monte Carlo study involving hundreds of millions of random permutations in S_{52} was carried out. The final results "looked funny." In the course of checking the programming, the method of generating the permutations came under scrutiny. The usual way of generating a random permutation on a computer is to put numbers $1, 2, \ldots, n$ into n-registers. Then choose a random number I_1 uniformly from one to n and transpose registers 1 and I_1. Next, choose I_2 randomly from two to n and

transpose registers 2 and I_2. Continuing $n-1$ steps gives a perfectly random permutation. This is the subgroup algorithm [27]. The programmer had thought this "too fussy" and instead carried out 60 random transpositions (as in (0.2)). The mathematicians involved complained and asked that the simulation be redone (three hours of computer time on a powerful processor). The programmer (and her boss) complained, but in the end, the simulation was redone using the proper algorithm and satisfactory results were obtained. All of this left me wanting to know how many random transpositions are required to mix up 52 cards.

Section 1 introduces a basic upper bound on the distance between Q^{*k} and the uniform distribution. The bound is most useful for class functions ($Q(s^{-1}ts) = Q(t)$) and involves a detailed knowledge of characters. This is illustrated for random walks on the symmetric group using random transpositions and for the Drunkard's walk on the circle. Section 2 introduces various norms and quadratic form used to bound eigenvalues via the minimax characterization. The basic bounds on distance to uniform using eigenvalues are set out. Then, comparison theorems are developed which allow analysis of a random walk based on one generating set (usually a small or messy set) in terms of a walk based on a nice generating set. The analysis involves relating the geometry of the two Cayley graphs. This is illustrated by giving a sharp analysis of the walk on the symmetric group using a transposition and an n-cycle.

In Section 3, developments of the random transpositions result are outlined. These go back to Hurwitz work on coverings of the sphere and deform to interesting walks on Hecke algebras with application to the Metropolis algorithm of statistical physics. The final section sets out some open problems.

Who cares about this stuff?

Of course, one can take the high road and answer that mathematics should be judged by its own internal naturalness and beauty. It is resoundingly true that mathematics developed 'just because it was there' has a remarkable record of turning out useful. For example, Frobenius's development of character theory was based on the strange question of understanding why the determinants of circulant matrices are a product of linear forms in the entries. The random transposition results that are the basis of the present paper are completely derivative of Frobenius's development.

Of course, certain random walks arise in daily life when people shuffle cards. I have written a survey of this subject in [10]. Ordinary random walk on $\mathbf{R}^d$ is a mainstay of parts of biology, chemistry, physics, and finance. It is natural to seek appropriate generalization to more general groups. Hughes [41, 42] and Saloff-Coste [66] show how natural these questions can be, even for mathematics.

A very satisfactory answer comes from problems generated from within group theory. Modern algorithms to manipulate and study large finite groups need a source of random elements. These are generated by a variety of random walk algorithms. One of the most popular is the product replacement algorithm. The mathematics in the present paper can be used to study these algorithms. Igor

Pak has done wonderful work along these lines. His survey Pak [59] contains an extensive review. On different lines, the mathematical questions that arise in studying random walk on groups are often quite different than classical questions. They have led to some interesting new group theory.

Perhaps the most compelling current motivation is the 'Markov chain Monte Carlo revolution'. Scientists in every walk of life are computing quantities of interest by running generalized random walks called Markov Chains. Liu [52] gives a nice introduction to this subject. Random walks on groups are special cases of Markov chains and the extra group structure along with years of hard work by group theorists can allow sharp results. The comparison theory explained below allows transfer to more general chains. More basically, techniques developed for groups can sometimes be extended to general Markov chains; again, the comparison theory is a good example. (See Diaconis and Saloff-Coste [18].)

Acknowledgement

It is the greatest pleasure to acknowledge my co-authors Mehadad Shahshahani, David Aldous, Dave Bayer, Ken Brown, Lou Billera, Fan Chung, Ron Graham, Susan Holmes, Jim Fill, Arun Ram and Laurent Saloff-Coste for their co-development of the subject. I have had the good fortune to work on random walks with wonderful graduate students, Farid Bassiri, Eric Belsley, Carl Dou, Martin Hildebrand, Andy Greenhaugh, Jason Fulman, Nathan Lulov, Igor Pak, Peter Mathews, Robin Pemantle, Francis Su, Jeffrey Rosenthal, Elizabeth Wilmer and Thomas Yan. Finally, I'm truly thankful to the organizers and participants of Group St. Andrews, Oxford Branch, for making an outsider feel welcome.

1 Random walk and representation theory

Let G be a finite group. A representation

$$\rho : G \to \mathrm{GL}_{d_\rho}(V)$$

assigns matrices to group elements in such a way that $\rho(st) = \rho(s)\rho(t)$. Here the dimension of V is denoted d_ρ. Background in representation theory and the probabilistic developments discussed here may be found in my book [8].

If $Q(s)$ is a probability distribution on G, the Fourier transform of Q at ρ is defined as

$$\widehat{Q}(\rho) = \sum_s Q(s)\rho(s).$$

As usual, Fourier transforms turn convolution into products

$$\widehat{Q^{\star 2}}(\rho) = \left(\widehat{Q}(\rho)\right)^2.$$

Further, for irreducible representations, the uniform distribution $u(s) = 1/|G|$ is characterized by its Fourier transform

$$\widehat{u}(\rho) = \begin{cases} 1 \text{ if } \rho \text{ is the trivial representation} \\ 0 \text{ if } \rho \text{ is non-trivial irreducible.} \end{cases}$$

Repeated convolutions may be shown to converge to the uniform distribution by showing that $\left(\widehat{Q}(\rho)\right)^k \to 0$ for ρ non-trivial irreducible. A quantitative version of this follows from the Fourier inversion and Planchenel theorems.

Theorem 1.1 *Let f be a complex function defined on G. Then*

$$\text{a) } f(s) = \frac{1}{|G|} \sum_{\rho} d_\rho \, tr \left(\widehat{f}(\rho)\, \rho\left(s^{-1}\right)\right)$$

$$\text{b) } ||f||^2 - \frac{1}{|G|} \sum_{\rho} d_\rho ||\widehat{f}(\rho)||^2 \qquad \text{(trace norm).}$$

Proof For (a), both sides are linear in f. Take $f(s) = \delta_t(s)$. Then $\widehat{f}(\rho) = \rho(t)$ and (a) asserts

$$\delta_t(s) = \frac{1}{|G|} \sum_{\rho} d_\rho \chi_\rho(st^{-1}).$$

This assertion holds as the right hand sum is the character of the regular representation which is $|G|$ or zero as st^{-1} is the identity or not.

For (b), we show

$$\langle f_1, f_2 \rangle = \sum f_1(s) \bar{f}_2(s) = \frac{1}{|G|} \sum_{\rho} d_\rho tr \left(\widehat{f}_1(\rho) \widehat{f}_2(\rho)^*\right).$$

Again, both sides are linear in f_2. Taking $f_2 = \delta_t$, we must show

$$f_1(\tau) = \frac{1}{|G|} \sum_{\rho} d_\rho tr \left(\widehat{f}_1(\rho) \rho^*(\tau)\right).$$

This follows from (a) since without loss of generality, ρ^* is unitary so $\rho^*(\tau) = \rho(\tau^{-1})$ □

The basic upper bound lemma was developed to study random transpositions in joint work with Shahshahani.

Lemma 1.2 (Upper bound lemma) *Let Q be a probability on the finite group G. Then, for the distance defined in* (0.1),

$$4||Q^{*k} - u||^2 \leq \sum_{\rho \neq 1} d_\rho ||\widehat{Q}(\rho)||^{2k}.$$

Proof From the definitions

$$\begin{aligned} 4||Q^{\star k} - u||^2 &= \left(\sum_s \left|Q^{\star k}(s) - u(s)\right|\right)^2 \leq |G| \sum_s \left|Q^{\star k}(s) - u(s)\right|^2 \\ &= \sum_{\rho \neq 1} d_\rho tr\ (\widehat{Q}(\rho)^k \left(\widehat{Q}\,(\rho)^k\right)^* \leq \sum_{\rho \neq 1} d_\rho ||\widehat{Q}(\rho)||^k. \end{aligned}$$

There, the first inequality is from Cauchy-Schwarz, the second equality is from Plancherel, the last inequality is $||AB|| \leq ||A||\,||B||$. □

Example 1.3 (Drunkard's walk) Let C_n be the integers modulo n, and let $Q(\pm 1) = \frac{1}{2}$ with $Q(j) = 0$ otherwise. Thus $Q^{\star k}(j)$ is the chance that a simple random walk is at j after k steps. To avoid periodicity problems, suppose n is odd. The irreducible representations are one-dimensional and given by $\rho_h(j) = e^{2\pi ijh/n}$. Here $\widehat{Q}(h) = \frac{1}{2}e^{2\pi ijh/n} + \frac{1}{2}e^{-2\pi ih/n} = \cos(2\pi h/n)$. The upper bound lemma gives

$$4||Q^{\star \ell} - u||^2 \leq \sum_{h=1}^{n-1} \cos\left(\frac{2\pi h}{n}\right)^{2\ell} = 2 \sum_{h=1}^{(n-1)/2} \cos\left(\frac{2\pi h}{n}\right)^{2\ell}.$$

This last sum must be bounded by calculus arguments. One way to proceed is to use $\cos(x) \leq e^{\frac{-x^2}{2}}$ for $0 \leq x \leq \pi/2$. Thus

$$\begin{aligned} ||Q^{\star \ell} - u||^2 &\leq \frac{1}{2} \sum_{j=1}^{(n-1)/2} e^{-\pi^2 j^2 \ell/n^2} \\ &\leq \frac{1}{2} e^{-\pi^2 \ell/n^2} \sum_{i=1}^{\infty} e^{-\pi^2 (j^2-1)\ell/n^2} \\ &\leq \frac{1}{2} e^{-\pi^2 \ell/n^2} \sum_{j=0}^{\infty} e^{-3\pi^2 j\ell/n^2} \\ &= \frac{1}{2} \frac{e^{-\pi^2 \ell/n^2}}{1 - e^{-3\pi^2 \ell/n^2}}. \end{aligned}$$

If $\ell \geq n^2$, $\left[2\left(1 - e^{-3\pi^2 \ell/n^2}\right)\right]^{-1} < 1$ and we conclude

$$||Q^{\star \ell} - u|| \leq e^{\frac{-\pi^2}{2} \frac{\ell}{n^2}} \text{ for } n \text{ odd with } \ell \geq n^2.$$

The result shows that a multiple of n^2 steps suffice to drive the distance to zero exponentially fast. While not developed here, this result is sharp–for ℓ small with respect to n^2 the distance to uniformity is close to its maximum value of 1 (See [8] pg. 29).

It is instructive to view the bounding process in the light of representation theory. The sum is dominated by the representations close to the trivial representation. Thus when $h = 1$, $\widehat{Q}(h) = \cos\left(\frac{2\pi}{n}\right) = 1 - \frac{2\pi^2}{n^2} + O\left(\frac{1}{n^4}\right)$. This must be raised to the

power of n^2 or more to make it small. For more general groups, it also seems to hold that the representations close to the trivial representation control the rate of convergence. Problem 5 in Section 4 has more on this.

The next example is central to present developments.

Example 1.4 (Random transpositions) Let S_n be the symmetric group. The probability measure $Q(w)$ defined at (0.2) is invariant under conjugation $Q(w) = Q(v^{-1}wv)$. Thus its Fourier transform satisfies $\rho(v^{-1})\widehat{Q}(\rho)\rho(v) = \widehat{Q}(\rho)$. For irreducible ρ, Schurs lemma implies that $\widehat{Q}(\rho)$ is a constant multiple of the identity: $\widehat{Q}(\rho) = cI$. Take the trace of both sides to see that $c = \frac{1}{n} + \frac{n-1}{n}\frac{\chi_\rho(\tau)}{d_\rho}$ with $\chi_\rho(\tau)$ the character of ρ at a transposition and d_ρ the dimension. Thus the upper bound is

$$4||Q^{*k} - u||^2 \leq \sum_{\rho \neq 1} d_\rho^2 \left(\frac{1}{n} + \frac{n-1}{n}\frac{\chi_\rho(\tau)}{d_\rho}\right)^{2k}.$$

To make further progress, we must get our hands dirty and come to terms with the character ratio. Fortunately, Frobenius did most of the work. The irreducible representations of S_n are indexed by partitions λ of n. If $\lambda = (\lambda_1, \lambda_2, , \ldots, \lambda_r)$ with $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r > 0$ and $\lambda_1 + \ldots + \lambda_r = n$, Frobenius showed that

$$\frac{\chi_\lambda(\tau)}{d_\lambda} = \frac{1}{n(n-1)} \sum_{i=1}^{r} \lambda_i^2 - (2i-1)\lambda_i.$$

To see what is involved in the bounding, take the representation closest to the trivial representation—the $n-1$ dimensional 'permutation' representation. This corresponds to $(n-1, 1)$ and the term to be bounded is

$$(n-1)^2 \left(1 - \frac{2}{n}\right)^{2k}.$$

Using $1 - x \leq e^{-x}$ we see that this is smaller than e^{-2c} for $k = \frac{1}{2}n\log n + cn$. This lead term dominates and determines the rate of convergence stated in Theorem A of the Introduction. The details lean on years of work by group theorists and combinatorialists (Tableaux Combinatorics). See [8] pg. 36-47 for a detailed proof which is a streamlined version of the original. □

Other walks and conjugacy classes

There have been a number of further careful studies of random walks on groups which are constant on conjugacy classes. One of the earliest is random walk on the hypercube C_2^n. This can be recast as a successful analysis of the Ehrenfest Urn model of statistical physics [8] pg. 19. Of course, any walk on an abelian group is constant on conjugacy classes.

Hildebrand [39] gave a careful complete analysis of a random transvections walk on $\mathrm{SL}_n(\mathbf{F}_q)$ using character theory. Roughly, he showed that for fixed q and n large, $n + c$ steps are necessary and sufficient. A sweeping generalization was make by Gluck [34]. He studied random walks on finite groups of Lie type with

probability distribution Q supported on small generating conjugacy classes. In close to complete generality, he proves that order rank (G) steps are necessary and suffice. Gluck uses the upper bound lemma and must thus bound character ratios of the form $\chi(c)/\chi(id)$. As opposed to earlier efforts which used hairy but explicit formulae for character ratios, Gluck proved that the ratios are uniformly bounded above by terms of form $q^{-\chi(id)^\theta}$ for e.g. $\theta = 1/10$. These proofs work by induction and make careful use of the structure of Lie type groups. The argument is so powerful that it is worthwhile trying to abstract it; Gluck (personal communication) reports that this did not seem easy to do for the symmetric group. In a different development, Gluck [35] applied his character ratio bounds to get sharp results on first return times for these random walks.

Character ratios figure in the analysis of many other applications of group theory. Some of these are explained in Gluck [34]. His results described above omit some groups "in the corners" and have some restrictions on the size of the generating conjugacy class. An elegant rounding out and natural completion of Gluck's results was carried out by Liebeck-Shalev [52, 51]. As will appear later, these constant on conjugacy class walks are a backbone of the approach outlined here to studying general walks. We thus have a very solid base to build on—a comprehensive theory for the finite simple groups of Lie type.

There is one striking open problem here. Hildebrand's work proves a sharp cutoff ($n+c$ steps necessary and suffice, the transition to randomness happens as c varies). The work on other Lie type walks results in statements such as 'fewer than rank (G) steps are not enough and there is a constant $A > 1$ so that $A \cdot$ rank $(G) + c$ steps suffice'. One thus does not have a sharp cutoff. There are good reasons to conjecture that the lead terms in the upper and lower bounds of any walk constant on conjugacy classes match up for a sequence of Lie type groups of growing rank. This is carefully explained in [9] or [65] which give surveys of this cutoff phenomena.

Turning to walks on the symmetric group, Rousell [64] has proved sharp bounds on random walk generated by small conjugacy classes such as three-cycles or products of two transpositions. Her results are of the following form: let c be a conjugacy class in S_n with $f(c)$ fixed points. Then $\frac{n(\log n+\theta)}{n-f(c)}$ steps are necessary and suffice to achieve randomness (as θ varies). At the other end, Lulov-Pak [56] treat walks generated by a single large conjugacy class. Here is one of their results. Let Q be a probability supported on the conjugacy class of an ℓ-cycle, with $\ell > n/2$. Then the walk generated by Q has a cut off at $k = \log n/\log(n/(n-\ell))$. This is closely related to earlier work of Roichman [61] who derived useful bounds on character ratios for large conjugacy classes in S_n. Alas his work has not been pushed through to give sharp bounds in probability problems. This is a hard but potentially fruitful area of study. One striking result along these lines has been proved by Lulov [55] and Fomin-Lulov [32]. Fix h and take n a multiple of h. Consider the walk generated by the conjugacy class which is a product of n/h h-cycles. This random walk gets random after two steps (and one step will not do). Lulov and Pak [56] give a more comprehensive survey, several conjectures and many further results.

A nice development of the character theoretic approach to random walk on the hyperoctahedral group B_n appears in Schoolfield [67]. He gives careful bounds for

the walk on a small generating conjugacy class and applies the result to a problem that arose in gene shuffling. I find it easiest to state their final result in the language of playing cards. Imagine a row of face down playing cards in order on the table. To start, card one is at left, ..., card n is at the right. Pick a random block of k cards and turn them over, both end for end and face up. How many times should this be repeated to mix both the order and face up—face down pattern?

Character theory methods can also be applied to compact groups such as the orthogonal group O_n or $\mathrm{SL}_n(\mathbf{Z}_p)$. Following work of Rosenthal [62] and Porod [60]. I have worked on such problems in [25]. This last, joint work with Laurent Saloff-Coste, attacks a 50 year old math/physics problem posed by Mark Kac. This may be rephrased as an analysis of the walk on O_n generated by picking a random pair of coordinates and then rotating by a random angle in that 2-space (random Givens rotations). It has been wonderfully developed and completed in work of Carlin et al. [7], Janvresse [47] and Maslin [57]. The natural conjecture: "order $n \log n$ steps are necessary and suffice and there is a cutoff in total variation" is still open.

Random walks that are constant on conjugacy classes can be seen as a special case of bi-invariant walks on a Gelfand pair. This is specified by a subgroup $H \subset G$ and a probability Q so that $Q(y) = Q(h_1 y h_2)$. Then the tools of spherical functions can be usefully employed. Surveys of random walks on Gelfand pairs appear in Letac [48, 49], Diaconis [8] and Belsley [4]. Taking $G \subset G \times G$, the spherical functions are the characters of G. I have not found this connection particularly illuminating. The two languages have their own rhythm and classes of examples. In particular, I have never been able to use any of the classical Gelfand pairs as a useful comparison for less symmetric walks.

Other techniques

To conclude this section let me mention that there are *many* other ways of studying random walks on groups that do not involve character theory. A survey of analytical approaches can be found in Diaconis and Saloff-Coste [20]. These included Poincaré, Nash, Sobolev, and Log Sobolev inequalities. There are techniques for lifting walks to covering walks, volume growth considerations and much else.

One striking result may be mentioned here as an example of what a comprehensive theory might give. Let G be a finite p-group of bounded derived length and Frattini rank. For any random walk based on a symmetric minimal generating set, the squared diameter of the group in this generating set is necessary and sufficient to achieve randomness. For the easiest example, the simple ± 1 random walk on the integers (mod p) has diameter of order p in these generators and the example above shows that order p^2 steps are necessary and sufficient to achieve uniformity. For the next simplest example, consider one of the extra special groups of order p^3. Any set of two generators has diameter of order p and again order p^2 steps are necessary and suffice. There are three proofs of the general $(diam)^2$ result and many applications; in [19] it is proved using moderate geometric growth. In [22] it is proved using Nash inequalities. In [20] it is proved using coverings and Harnack inequalities.

In addition to the analytic techniques above, there are purely probabilistic techniques– coupling and strong stationary times which give eloquent definitive results in special cases. The forthcoming book of Aldous and Fill [3] treats these. I have given a treatment in [8] and in joint work with Fill [12]. Igor Pak [58], has developed strong stationary times in marvelous ways.

There are also a variety of mixing schemes studied by methods not captured above. Chief among these are walks on S_n associated to riffle shuffling cards. After the original work of Bayer-Diaconis showing seven ordinary riffle shuffles suffice, Bidigare-Hanlon-Rockmore followed by Diaconis and Brown extended things to walks on the chambers of a hyperplane arrangement. This includes all finite reflexion groups. Then, Ken Brown extended things to walks on idempotent semi groups. This includes walks on spherical buildings. These shuffling walks have many connections with group theory and symmetric function theory. I have written a recent survey of these developments [10]. Finally, marvelous results which may be interpreted in the language of random walk on finite groups have been proved in studying expander graphs. These make deep contact with modern mathematics, drawing on work of Deligne and Selberg. I will not try to paint this picture here but refer to Lubotzky [54] for extensive pointers to the literature.

2 Analytic geometry

This section sets out the analytic tools concerning eigenvalues that allow geometric methods (paths between elements, diameter, covering numbers, volume growth) to be used to bound rates of convergence of random walk. Throughout, G is a finite group, $Q(s) = Q(s^{-1})$ a symmetric probability distribution. Often $Q(s) = 1/|S|$ or zero as $s \in S$ or not with $S = S^{-1}$ a symmetric set of generators. Associated to Q is a graph with vertex set G and an edge from s to t if $Q(st^{-1}) > 0$. Geometry refers to the geometry of this graph. It is also helpful to associate a $|G| \times |G|$ matrix, the transition matrix with $M_{st} = Q(ts^{-1})$. This gives the chance of going from s to t in one step of the walk. The matrix M is symmetric and doubly stochastic. It thus has real eigenvalues π_i and the Perron-Frobenius theorem says that $1 = \pi_0 \geq \pi_1 \ldots \geq \pi_{|G|-1} \geq -1$. Just to warm up, here is an easy (and useful) bound for the smallest eigenvalue.

Lemma 2.1 *For G and Q above, the smallest eigenvalue satisfies*

$$\pi_{|G|-1} \geq -1 + 2Q(id).$$

Proof If $Q(id) = 0$ this is certainly true. If $Q(id) > 0$, let $\tilde{Q}(s) = Q(s)/(1 - Q(id))$ for $s \neq id$, $\tilde{Q}(id) = 0$. This is symmetric with eigenvalues $\tilde{\pi}_i \geq -1$. Thus $\frac{1}{1-Q(s)}\left(\pi_{|G|-1} - Q(id)\right) \geq -1$. This gives the result. □

Remark It is not necessary to have $Q(id) > 0$ to bound negative eigenvalues (see Diaconis and Saloff-Coste [22]). This *is* the most commonly occurring special case.

The main bound on convergence to uniformity using eigenvalues is

Lemma 2.2 *For G and Q as above,*

$$4||Q^{*k} - u||^2 \le |G|\,||Q^{*k} - u||_2^2 = |G|\left(Q^{*2k}(id) - \frac{1}{|G|}\right) = \sum_{i=1}^{|G|-1} \pi_i^{2k}.$$

Proof From the definition of total variation distance in the Introduction,

$$\begin{aligned} 4||Q^{*k} - u||^2 &= \left(\sum \left|Q^{*k}(s) - \frac{1}{|G|}\right|\right)^2 \le |G| \sum \left|Q^{*k}(s) - \frac{1}{|G|}\right|^2 \\ &= |G|\left(\sum_s \left(Q^{*k}(s)\right)^2 - \frac{2}{|G|} + \frac{1}{|G|}\right) = |G|Q^{*2k}(id) - 1. \end{aligned}$$

From the matrix interpretation, $|G|Q^{*2k}(id) = \text{tr}\,\left(M^{2k}\right) - \sum_{i=0}^{|G|-1} \pi_i^{2k}$. This completes the proof. □

Corollary 2.3 *For G and Q as above*

$$4||Q^{*k} - u||^2 \le |G|\pi_\star^{2k}, \qquad \pi_\star = \max\left(\pi_1, |\pi_{|G|-1}|\right).$$

Remark The bound in Lemma 2.2 has been shown to be quite sharp in many examples. There is usually a reasonably close matching lower bound. The bound in the corollary is looser and usually off by factors of $\log|G|$. It is sometimes all that is available.

How to bound eigenvalues Define an inner product on real valued functions on G by $\langle f_1|f_2\rangle = \sum_s f_1(s)f_2(s)$. The linear space of all functions (the group algebra) is denoted $L^2(G)$. A symmetric probability Q defines a Laplace operator $(I-Q)f(s) = f(s) - \sum f(t)Q(ts^{-1})$. This has eigenvalues $1 - \pi_i$. The associated quadratic form is called the *Dirichlet Form* in probabilistic circles:

$$\mathcal{E}(f|f) = \langle (I-Q)f, f\rangle = \frac{1}{2}\sum_{s,t} (f(s) - f(st))^2\, Q(t).$$

As usual, we may bound eigenvalues by the minimax principle: let V be a real linear space, T a symmetric linear map on V with eigenvalues $q_0 < q_1 < \ldots$. For a subspace W, set $M(W) = \max\langle Tv, v\rangle/||v||^2$, $v \in W$, $v \neq 0$. Then $q_i = \min\{M(W) : \dim(W) = i+1\}$. See Horn and Johnson [40] for a nice development. The minimax characterization implies

Corollary 2.4 *Let Q, $\tilde{Q}$ be symmetric probabilities on G with eigenvalues $\pi_i, \tilde{\pi}_i$ and associated forms $\mathcal{E}, \tilde{\mathcal{E}}$. If for some constant $A > 0$, $\tilde{\mathcal{E}} \le A\mathcal{E}$ then for all i, $\pi_i < 1 - (1 - \tilde{\pi}_i)/A$.*

Several applications of this result are given later in this section. Usually, $\tilde{Q}$ is a walk about which we know everything and Q is a walk we want to study. The main use of Corollary 2 is the following basic upper bound.

Theorem 2.5 *Let $Q, \widetilde{Q}$ be symmetric probabilities on G. If $\widetilde{\mathcal{E}} \leq A\mathcal{E}$ then*

$$||Q^{\star k} - u||_2^2 \leq \pi_{\min}^{2k} + e^{-k/A} + ||\widetilde{Q}^{\star \lfloor k/2A \rfloor} - u||_2^2.$$

Proof From Lemma 2, $||Q^{\star k} - u||_2^2 = \frac{1}{|G|} \sum_{i=1}^{|G|-1} \pi_i^{2k} \leq \pi_{\min}^{2k} + \frac{1}{|G|} \sum_{i:\pi_i>0} \pi_i^{2k}$.

Now use the calculus bounds $1 - x < e^{-x}$, $1 - x > e^{-2x}$ for $0 < x \leq \frac{1}{2}$. Thus $\pi_i > 0$ gives $\pi_i^{2k} \leq \left(1 - \frac{(1-\pi_i)}{A}\right)^{2k} \leq e^{\frac{-(1-\pi_i)2k}{A}} \leq \pi_i^{k/A}$. □

Theorem 1 gives bounds on a probability of interest in terms of a known probability in the presence of a comparison between forms.

How to compare forms (and first examples).

If S is a symmetric generating set of G let $|t| = \min_k : t = s_1 \dots s_k$. $|id| = 0$. For such a representation of t, let $N(s,t) = \#$ (times s appears). Fix a minimal representation for each t.

Theorem 2.6 *Let $Q, \widetilde{Q}$ be symmetric probabilities on G. Let S be a symmetric generating set with $Q(s) > 0$ for $s \in S$. Then*

$$\widetilde{\mathcal{E}} \leq A\mathcal{E} \text{ for } A = \max_{s \in S} \frac{1}{Q(s)} \sum_{\iota \in G} |t| N(s,t) \widetilde{Q}(t).$$

Proof For x, t in G, write $t = s_1 \dots s_k$ as above. Then, for any function f,

$$f(x) - f(xt) = (f(x) - f(xs_1)) + f(xs_1) - f(xs_1s_2) + \dots + (f(xs_1 \dots s_{k-1}) - f(xt)).$$

Squaring both sides and using the Cauchy-Schwarz inequality gives

$$(f(x) - f(xt))^2 \leq |t| \left\{ (f(x) - f(xs_1))^2 + \dots + (f(xs_1 \dots s_{k-1}) - f(xt))^2 \right\}.$$

Summing in x

$$\sum_x (f(x) - f(x\iota))^2 \leq |\iota| \sum_{\substack{x \in G \\ s \in S}} (f(x) - f(xt))^2 N(s,t).$$

Multiply both sides by $\widetilde{Q}(t)/2$ and sum in t. The left side is $\widetilde{\mathcal{E}}(f|f)$. The right side is

$$\frac{1}{2} \sum_{\substack{x \in G \\ s \in S}} (f(x) - f(xs)) \frac{Q(s)}{Q(s)} \sum_t |t| N(s,t) \widetilde{Q}(t) \leq A\mathcal{E}(f|f).$$

□

The first time one looks at Theorem 2.6, things look hopeless. As will emerge, it is often possible to get useful bounds on A. This is illustrated in a series of examples below.

First examples We begin with a very basic example which 'works' for all walks on all groups. This is then specialized to the walk on the symmetric group based on the generating set of a transposition and an n-cycle.

3 Other appearances of random transpositions

When Shahshahani and I completed our analysis of random transpositions it seemed very delicate; the method of proof breaks down if the generating set is not a union of conjugacy classes. The last example of section 3 shows that transpositions plus comparison can handle examples with no relation to conjugacy. In joint work with Saloff-Coste [17] we treated many further examples: consider a connected graph on $\{1, 2, \ldots, n\}$, take a transposition for each edge. This gives a generating set. If the graph is a path, we get the usual Coxeter generators. If the graph is a star, the walk becomes 'transpose random with one'. Of course, for the complete graph, we get random transpositions. The comparison theory is easy and we derived reasonably sharp answers for this general set of problems. For a path the answer is order $n^3 \log n$. This had been done earlier by Aldous [1] using coupling. The lower bound for this case was only done recently by Wilson [75]. It introduces a powerful lower bound technique that seems very useful. For a star, $n(\log n + c)$ steps are necessary and suffice. Again, the example had been done earlier by Flatto-Odlyzko-Wales using the fact that the restriction of an irreducible representation of S_n restricts S_{n-1} in a multiplicity free way. All other graphs give new examples, handled by a uniform method. The paper with Saloff-Coste treats overhand shuffles, an open problem of Borel-Levy (choose a random packet from the center of the deck and cut it to the top.).

While I will not expand on it here, the comparison approach is *not* restricted to symmetric random walk ($Q(w) = Q(w^{-1})$). This is developed in great detail in Diaconis and Saloff-Coste [22]. In particular, for n odd, the generating set $\{(1,2), (1,2,\ldots,n)\}$ (no identity, no inverses) yields a random walk that equilibriates in $n^3 \log n$ steps. Another technical improvement developed in joint work with Saloff-Coste [22] uses averages over random paths.

I've been very pleased to see two applications in biology: DNA sequences evolve by a variety of transmutations (substitutions, insertions, deletions). There are also 'translocations', exchanging genetic material between chromosomes. This can be studied by picturing a row of symbols, picking a pair (i, j) from some distribution, and reversing the order $i \leftrightarrow j \quad i+1 \leftrightarrow j-1, \ldots$. Fill and Schoolfield studied this (keeping track of face up and face down symbols which have a biological meaning) using the analog of random transpositions on the hyperoctahedral group B_n. Durrett [31] studied the process on the symmetric group and includes a serious comparison between data and model. A very different application, to phylogenetic trees, is also based on random transpositions; see Diaconis and Holmes [15] which also contains pointers to coagulation processes in chemistry.

All of this development leans on Frobenius's neat character formula of example 1.4. I was surprised to see an earlier application of this formula in work of Rothaus-Thompson [63]. They were studying perfect codes on the symmetric group. This is a set of permutations $w_1, w_2, \ldots, w_k$ such that the balls of radius h in the metric based on transpositions exactly partition the group. The problem can be phrased as convolution with the random transpositions measure and the character formula (and some clever number theory) yields some partial results. There is much left

open as well.

The earliest Application A most surprising discovery was made by Basil Gordon (personal communication). The Basic random transposition model appeared in work of Hurwitz [43, 44]. Hurwitz was counting the number of n sheeted covers of the Riemann sphere with d-simple branch points. Riemanns existence theorem says that such covers are in one to one correspondence with ways of writing the identity permutation in S_n as a product of d transpositions. Here the individual transpositions specify how to glue the covering sheets over the various branch points. Transpositions occur because of the (assumed) square root singularities. Up to multiplying by a factor of $\binom{n}{2}^d$ this number of ways equals $Q^{*d}(id)$. Using the Fourier inversion formula

$$Q^{*d}(id) = \sum_{\lambda \vdash n} d_\lambda^2 \left(\frac{\chi_\lambda(\tau)}{d_\lambda} \right)^d$$

Now Frobenius's formula gives an explicit result.

In fact, Hurwitz wanted to count *irreducible* covers. In group theory terms this means that he requires that the transpositions that appear must generate S_n. He originally solved this problem by a form of exclusion-inclusion (if the transpositions don't generate S_n they generate a subgroup). His version of this is one of the earliest appearances of what is now called the exponential formula of enumerative combinatorics (see Stanley [70], Chapter 6)). By looking at examples computed in this way for small n, Hurwitz guessed that all the mess clears away and the final answer is simple. Hurwitz made the conjecture:

Let σ have cycle lengths $k_1, k_2, \ldots, k_\rho$. *The number of factorizations*

$$\tau_1 \ldots \tau_{n+\rho-2} = \sigma$$

of the permutation σ into a product of transpositions τ_i, where $\langle \tau_1, \ldots, \tau_{n+\rho-2} \rangle = S_n$ is

$$(n+\rho-2)!\, n^{\rho-3} \prod_{i=1}^{\rho} \frac{k_i^{k_i+1}}{k_i!}.$$

The full result was proved by Goulden and Jackson [36] . In fact, Hurwitz outlined a proof which was completed by Strehl [72].

The enumerative theory of surfaces and their covers and triangulations has a rich development. Useful surveys are given by Jones [46], Zvonkin [76] and in the article by Condor that appears in this volume. There are amazing recent connections to modern theoretical physics through Gromov-Witten invariants. See the article by Goulden-Jackson-Vakil [37] for pointers. I have no doubt that some of these developments can be turned into shuffling theorems but this lies in the future.

Deformations My most recent encounter with random transpositions comes through joint work with Arun Ram on analysis of 'Systemic scan Metropolis algorithms'. This involves a deformation into the Iwahori-Hecke algebra. In the end, the familiar analysis based on Frobenius's character formula allowed sharp estimates. I think there is much further work to be done here and will include a high level overview.

The Metropolis algorithm is a mainstay of scientific computing. It is used in physics, chemistry, biology, statistics and business applications. Billera and Diaconis [6] and Diaconis and Saloff-Coste [24] are surveys with extensive pointers to the literature. The present example involves comparing two variants (random versus systematic scan). The problem is to draw repeated random samples from a *non-uniform* distribution on S_n. Fix $0 < \theta < 1$. Let

$$\pi(w) = \theta^{-\ell(w)}/P(\theta^{-1}) \text{ where } P(\theta^{-1}) = \sum_w \theta^{-\ell(w)}$$

is the normalizing constant. Here $\ell(w)$ is the length of the permutation w in the usual generating set $s_i = (i, i+1)$, $1 \le i \le n-1$. Of course, for $\theta = 1$ this becomes the uniform distribution but, for *e.g.* $\theta = \frac{1}{2}$, this concentrates on permutations with larger lengths. These non-uniform distributions are known as Mallows models and are widely applied.

A standard way to sample from $\pi(w)$ involves a random walk on S_n which may be called the *Random Scan Metropolis Algorithm*. It is simple to state. The walk starts at some fixed permutation (say id) and proceeds by making random pairwise adjacent transpositions according to the following scheme. Suppose the walk is currently at w. Choose i uniformly in $1 \le i \le n-1$. If $\ell(s_i w) > \ell(w)$ the walk moves to $s_i w$. If $\ell(s_i w) < \ell(w)$, flip a coin with probability of heads θ. If this coin comes up heads, move to $s_i w$. If the coin comes up tails, the walks stays at w. This generates a random sequence $w_0 = id$, $w_1, w_2, \ldots$. Simple theory shows that the probability of $(w_k = w) \to \pi(w)$ as $k \to \infty$.

In our second variant, the walk proceeds as above but instead of choosing the proposal transpositions at random, things proceed systematically; first try s_1 then $s_2, \ldots$, then s_{n-1} then $s_{n-1}, s_{n-2}, \ldots$, then s_1. At each stage one compares the length and makes an auxiliary coin toss if needed. Call the result of one pass through (based on $2(n-1)$ steps) the *systematic scan* Metropolis algorithm. Such systematic scans are widely employed in applications to Ising-like models in physics and image analysis. Again, simple theory shows that they converge to $\pi(w)$.

It is natural to ask how long each of the algorithms takes to converge and which or when one is better. The problems are largely open for the usual applications of the Metropolis algorithm. For the special permutation case discussed here, the analysis can be pushed through. Roughly stated, the systematic scan procedure takes order n passes (and so order n^2 steps). This is stated more carefully below. In very recent work, Benjamini, et al. [5] have shown that the random scan version also takes order n^2 steps to converge. This seems surprising, since the systematic procedure builds in some extra structure. This is the first and only example where such a comparison has been made. The paper with Ram [16] carries this out for general finite reflection groups with similar findings.

The reason for mentioning the subject here is that the analysis rests on a novel probabilistic interpretation of multiplication in the Hecke algebra. Let W be a finite Coxeter group generated by simple reflections $s_1, \ldots, s_n$. These define a length function with $\ell(id) = 0$, $\ell(s_i) = 1$ and $\ell(s_i w) = \ell(w) \pm 1$. The Iwahori-Hecke algebra H corresponding to W is the vector space with basis T_w for $w \in W$

and multiplication given by

$$T_i' T_w' = \begin{cases} T_{s_i w} & \text{if } \ell(s_i w) = \ell(w) + 1 \\ (q-1)T_w + q T_{s_i w}' & \text{if } \ell(s_i w) = \ell(w) - 1 \end{cases}$$

where $T_i = T_{s_i}$. We have $T_i^2 = (q-1)T_i + q$ or equivalently $(T_i' - q)(T_i + 1) = 0$. I have always found this multiplication intriguing and tried to find a stochastic interpretation. In the work with Ram we proved

Theorem 3.1 *Let W be a finite Coxeter group as above. Set*

$$q = \theta^{-1}, \quad \widetilde{T}_i = T_i/q, \quad \widetilde{T}_w = q^{-\ell(w)} T_w \text{ for } w \in W.$$

Then the systematic scan Metropolis chain (symbolically $K_1 K_2 \dots K_n K_n \dots K_1$) has the same transition matrix as multiplication by $\widetilde{T}_1 \widetilde{T}_2 \dots \widetilde{T}_n \widetilde{T}_n \dots \widetilde{T}_1$ in the Iwahori-Hecke algebra with basis $\widetilde{T}_i$.

Central to our work is the fact due to Breiskorn and Deligne that the long systematic scan $(K_1 K_2 \dots K_n K_n \dots K_1) \cdots (K_1 K_2 K_2 K_1)(K_1 K_1)$ corresponds to multiplying by an element in the center of H. This is the 'q analog', or deformation of the fact that the sum of all transpositions is in the center of the group algebra. The action by this element can be explicitly diagonalized. The eigenvalues are closely related to Frobenius's formula and it was possible to push through a successful analysis. Let me end this sea of exposition with a clear mathematical result from my work with Ram.

Theorem 3.2 *Let K be the transition matrix for one pass of the short systematic scan algorithm on S_n. For $\ell = n/2 - (\log n / \log \theta) + c$ with $c > 0$, for all n*

$$4 \| K_1^\ell - \pi \|^2 \leq (e^{\theta^{2c+1}} - 1) + n! \theta^{n^2/8 - n \log n / \log \theta + n(c+1/4)}.$$

Conversely, for $\ell < n/4$, for fixed θ, $\| K_1^\ell - \pi \|$ tends to 1 as $n \to \infty$.

I must not leave this part of the world without mentioning that Diaconis and Hanlon [14] analyzed the Metropolis deformation of the random transpositions chain on S_n with stationary distribution $\sigma(w) = Z^{-1} \theta^{|w|}$ where $|w|$ is the length if *all* transpositions are used. This deformation led to the Jack symmetric functions. It is an intriguing problem to see if the two parameter Macdonald polynomials can be obtained in this way.

4 Some open problems

There are an infinite variety of open problems. These range from specific to fairly general. Specific problems can be helpful in pointing to the need for new tools and understanding.

1. It is well known (Suzuki [73]) that for p a prime, $\mathrm{SL}_n(\mathbf{F}_p)$ is generated by elementary row operations, adding or subtracting one row from another. Let E_{ij} be the $n \times n$ matrix with ones down the diagonal and a one in position (i, j) (for fixed $i \neq j$). Then these elementary transvections generate. It is natural to ask for the speed of convergence. This problem appears in a variety of guises; Diaconis and Saloff-Coste [23] relate it to a particle system and to a special case of the product replacement algorithm. Using the method outlined in section three they were able to show that order $n^4(\log p)^2$ steps suffice. The best lower bound available is of order $n^2 \log p$. This is a natural enough problem that an answer should be sought. The approach taken in [23] is to use comparison with the walk generated by random transvections (the whole conjugacy class). Hildebrand [39] analyzed this walk by character theory. Kai Magaard showed us that any transvection can be written as a product of at most $5n$ elementary transvections so we were off and running. Igor Pak has obtained some improvements and extensions but the general problem of determining the right rate (and perhaps showing there is a cut off) remains open.
2. Steinberg [71] showed that any finite group of Lie type can be generated by two generators. This offers a list of problems: pick your favorite group (or one you'd like to get to know), figure out what Steinberg's generators are and get to work. To make things easier, you might begin with $p = 2$. Glucks bounds and some basic geometry should suffice to show that order a small polynomial in rank(G) steps suffice. Even this would require honest work. Finding the right answer would be a major achievement. Here is a specific case which I find interesting. Take $\mathrm{SL}_n(\mathbf{F}_p)$. A Singer cycle is an element of maximal order. It can be explicitly constructed as the $n \times n$ companion matrix of a primitive polynomial. Let A be a Singer cycle and $B = E_{12}$ (elementary transvection). These two matrices generate $\mathrm{SL}_n(\mathbf{F}_p)$. They are the appropriate analogs of n-cycle and transposition in S_n. Bound the rate of convergence of this random walk.
3. Towards generality; all finite simple groups are generated by almost all pairs of elements. For the alternating group, this is a theorem of Dixon. It has been improved extended and refined by a generation of group theorists. See Shalev [68] for the latest results. This suggests a class of problems: pick two elements of such a group at random and study the expected relaxation time for the random walk. I conjecture that it is bounded by a small polynomial in Rank(G). In the case of the alternating group, I conjecture it is bounded by $n^3 \log n$. The best that is known rigorously is order $e^{c\sqrt{n}}$ (Babai). Sticking to A_n, it is possible that there are a bounded number of generators such that the relaxation time is of order $n \log n$; a tantalizing conjecture of Lubotzky

says this can't happen see Gamburd-Pak [33] for more on this.

4. The question of understanding the cutoff phenomenon is perhaps more abstract. If a_n, b_n are two sequences tending to infinity with b_n/a_n tending to zero and $k_n = \lfloor a_n + cb_n \rfloor$ then a sequence of probabilities Q_n on groups G_n satisfies a cut off if there are real functions $f(c), g(c) \geq 0$ such that $g(c) \to 0$, $f(c) \to 1$ and

$$f(c) \leq \|Q_n^{*k_n} - u\| \leq g(c).$$

A variety of our examples above satisfy this; for random transpositions, one may take $a_n = \frac{n}{2}\log n$, $b_n = n$. The fact that the phenomena was discovered at all suggest that it is generic. There may be some soft way to prove this along the lines of concentration phenomena in modern combinatorics. Failing this, one can try to establish it for sets of examples along the lines of Rousell [64], Lulov-Pak [56]. In [19] we show that Nilpotent groups of low derived length with small generating sets do not satisfy cutoffs. Random walks on groups like C_p^n for n large do satisfy cutoffs. What's going on? It is tempting to try to relate these cutoffs to other types of phase transitions as in Dubois et al. [30].

5. In many examples there is a natural ordering on the irreducible representations ρ such that for "simple" probability measures Q, $\|\widehat{Q}(\rho)\|$ is monotone decreasing as ρ moves away from the trivial representation. For example, for $G = C_n$ with n odd and $Q(1) = Q(-1) = 1/2$, $|\widehat{Q}(h)| = |\cos\left(\frac{2\pi h}{n}\right)|$ this is one when $h = 0$ and decreases for increasing $h, 0 \leq h \leq n/2$. For random transpositions $\widehat{Q}(\rho_\lambda) = \left(\frac{1}{n} + \frac{x_\lambda(j)}{d_\lambda}\right)$ I. In [8] the character ratio was shown to be monotone in the usual partial order on partitions. This question can be asked in purely group theoretic terms. To be specific, let G be a finite simple group. Let c be a conjugacy class not in the center. Show that the character ratio $|x_\rho(c)|/x_\rho(id)$ is monotone decreasing in the dimension $x_\rho(id)$. Put this precisely the conjecture may be false but since many special cases have been found, some mild weakening must hold. The Riemann-Lebesgue lemma is actually an asymptotic version of this: let G be a compact group. Let $f : G \to C$ be in L^1 then, $\widehat{f}(\rho)$ tends to zero as $\dim(\rho)$ tends to infinity.

6. One of the exciting developments of group theory is the modern theory of p-groups. At long last a theory of these monsters is beginning to emerge. This is summarized in recent books by Dixon et al. [29]. The theory focuses on groups of size p^n with large class. One may study random walk on these. For example, the groups of maximal class are all generated by two generators. Pick a group and a generating set and start working. To be specific, the Nottingham group may be represented as polynomials $f(x) = x + a_2x^2 + \ldots + a_nx^n$, with a_i in $\mathbf{F}_p$, taken mod x^{n+1}. These form a group under composition. A generating set is x, $x + x^2$. What is the rate of convergence as a function of p and n? Here is one further specific example. Let $G = C_p wr C_p$. This is a group of order p^{p+1}. It appeared early in Philip Hall's study of regular p-groups (G is a "smallest" example of a non-regular group) if G is represented

as C_p acting on C_p^p by cyclic shift then a generator of C_p and any non-zero vector in C_p^p generate. Uyemura-Reyes [74] gives upper and lower bounds of order p^3 and $p^3 \log p$ respectively but the right answer is unknown. Reyes shows how random walk on such wreath products arise in card shuffling and relates these questions to recent work on the lamplighter group along with work of Grigorchuk et al. [38] who used the eigenvalues of such walks to disprove a question of Atiyah.

7. It is very natural to study Markov Chains on finite rings. The meat ax falls into this class. Here is one simple, completely open example. In F_p, consider the walk which moves from x to $x+1$ or x^2 with probability $1/2$. I'm morally certain that this takes order $\log p$ steps. At present, I don't even have a rough description of the stationary distribution (it is certainly *not* uniform).

Further questions are in [24]. I am pleased to report that almost all the questions I posed in my book [8] have been usefully settled. The most annoying open one is Thorp's model of shuffling cards [8] pg. 90.

Added in proof. A very recent application of random transpositions occurs in Diaconis, P., Mayer-Wolf, E., Zeitouni, O., and Zerner. M. (2002). Uniqueness of invariant measures for split-merge transformations and the Poisson-Dirichlet Law. To appear, Ann. Probab. They prove a conjecture of Vershik on the stationary distribution of a coagulation-fragmentation process occurring in chemistry by showing that it reduces to following the cycle structure under random transpositions.

References

[1] Aldous, D. (1983). Random walk on finite groups and rapidly mixing Markov Chains. In *Seminaire de Probabilités* **XVII**, 243-297. Springer Lecture notes in Math. 986.

[2] Aldous, D. and Diaconis, P. (1986). Shuffling cards and stopping times, *Amer. Math. Month.* **43**, 333-348.

[3] Aldous, D. and Fill, J. (2002). Reversible Markov Chains and random walk on graphs. Forthcoming book.

[4] Belsley, E. (1998). Rates of convergence of random walk on distance regular graphs, *Prob. Th. Related Fields* **112**, 493-533.

[5] Benjamini, I., Beger, W., Hoffman, Z. Mossell, E. (2002). Mixing time for biased card shuffling. Preprint, Dept. of Mathematics, University of Washington.

[6] Billera, L. and Diaconis, P. (2001). A geometric interpretation of the Metropolis-Hastings algorithm, *Statistical Science* **16**, 335-339.

[7] Carlin, E., Carvollio, M. and Loss, M. (2002). Determination of the spectral gap for Kac's master equation. To appear *Acta Math.*

[8] Diaconis, P. (1988) *Group representations in Probability and Statistics*, Institute of Mathematical Statistics, Hayward, CA.

[9] Diaconis, P. (1996). The cutoff phenomenon in finite Markov Chains, *Proc. Nat. Acad. Sci.* **43**, 1659-1664.

[10] Diaconis, P. (2002). Mathematical developments from the analysis of riffle shuffling. To appear, M. Liebeck (ed.). Proc. Durham Conference on Groups.

[11] Diaconis, P. and Brown, K. (1998). Random walk and hyperplane arrangements, *Ann. Probab.* **26**, 1813 1854.

[12] Diaconis, P. and Fill, J (1990). Strong stationary times via a new form of duality, *Ann. Prob.* **18**, 1483-1522.

[13] Diaconis, P. and Graham, R. (1981). The analysis of sequential experiments with feedback to subjects, *Ann. Statist.* **9**, 3-23.
[14] Diaconis, P. and Hanlon, P. (1992). Eigen-analysis for some examples of the Metropolis algorithm, *Contemp. Math.* **138**, 99-117.
[15] Diaconis, P. and Holmes, S. (2002). Random walk on trees and matchings. In *Electronic Jour. Probab.*
[16] Diaconis, P. and Ram, A. (2000). Analysis of systematic scan Metropolis algorithms using Iwahori-Hecke algebra techniques, *Mich. Math. Jour.* **48**, 157-190.
[17] Diaconis, P. and Saloff-Coste, L. (1993). Comparison techniques for random walk on finite groups, *Ann. Prob.* **21**, 2131-2156.
[18] Diaconis, P. and Saloff-Coste, L. (1993). Comparison theorems for reversible Markov Chains, *Ann. Appl. Prob.* **3**, 696-730.
[19] Diaconis, P. and Saloff-Coste, L. (1994). Moderate growth and random walk on finite groups, *GAFA* **4**, 1-36.
[20] Diaconis, P. and Saloff-Coste, L. (1995). An application to Harnack inequalities to random walk on nilpotent quotients, *Jour. Four. Anal. Applications, Special Kahane Issue*, 189-207.
[21] Diaconis, P. and Saloff-Coste, L. (1995). Random walks on finite groups: A survey of analytic techniques. In *Probability Measures on Groups* **XI** (H. Heyer ed.). *World Scientific*, Singapore, 44-77.
[22] Diaconis, P. and Saloff-Coste, L. (1996). Nash inequalities for finite Markov Chains, *Jour. Th. Probab.* **9**, 459-510.
[23] Diaconis, P. and Saloff-Coste, L. (1996). Walks on generating sets of Abelian groups, *Prob. Th. Related Fields* **105**, 393-421.
[24] Diaconis, P. and Saloff-Coste, L. (1998). What do we know about the Metropolis algorithm? *Jour. Comp. System Sci.* **5**, 20-36.
[25] Diaconis, P. and Saloff Coste, L. (2000). Bounds for Kac's master equation, *Commun. Math. Phys.* **209**, 729-755.
[26] Diaconis, P. and Shahshahani, M. (1981). Generating a random permutation with random transpositions, *Z. Wahr. Verw. Gebete* **57**, 159-179.
[27] Diaconis, P. and Shahshahani, M. (1987). The subgroup algorithm for generating uniform random variables, *Prob. Eng. Infosci.* **1**, 15-32.
[28] Diaconis, P. and Shahshahani, M. (1987). Time to reach stationality in the Bernoulli-Laplace diffusion model, *SIAM Jour. Math. Analysis* **18**, 208-218.
[29] Dixon, J., Dusautoy, M., Mann, A., Segal, D. (1999). Analytic pro-P groups, 2nd ed., Cambridge University Press, Cambridge.
[30] Dubois, O., Monassor, R., Selman, B. and Zecclina (eds.). Special issue on phase transitions in combinatorial problems, *Theoretical Computer Science* **265**.
[31] Durrett, R. (2002). Shuffling Chromosomes, Technical Report, Dept. of Mathematics, Cornell University.
[32] Fomin, S. and Lulov, N. (1995). On the number of rim hook tableaux, *Zap. Nauchn. Sem. Pomi.* **223**, 218-226. (Also, http://www.math.lsa.umich.edu/ Fomin/papers/).
[33] Gambaurd, A. and Pak, I. (2002). Expansion of product replacement groups, *S.O.D.A. 2002.*
[34] Gluck, D. (1994). Characters and random walk on finite classical groups, *Adv. Math.* **129**, 46 72.
[35] Gluck, D. (1999). First hitting times for some random walks on finite groups, *Jour. Th. Probab.* **12**, 739-756.
[36] Goulden, I. and Jackson, D. (1997). Transitive factorization into transpositions and holomorphic mappings on the sphere, *Proc. Amer. Math. Soc.* **125**, 51-60.
[37] Goulden, I. Jackson, D. and Vakil, R. (1999). The Gromov-Witten potential of a

point, Hurwitz numbers, and Hodge integrals. To appear *Proc. Lond. Math. Soc.*
[38] Grigorchuk, R. Linell, P., Schick, T. and Zuk, A. (2000). On a conjecture of Atiyah, *C.R. Acad. Sci. Paris. t331*, Serie I, 663-668.
[39] Hildebrand, M. (1992). Generating random elements in $SL_n(\mathbf{F}_q)$ by random transvections, *J. Alg. Combin.* **1**, 133-150.
[40] Horn, R. and Johnson, C. (1985). Matrix analysis, Cambridge Press. Cambridge.
[41] Hughes, B. (1995). Random walks and random environments, Vol. 1, Oxford Press, Oxford.
[42] Hughes, B. (1996). Random walks and random environments, Vol. 2, Oxford Press, Oxford.
[43] Hurwitz, A. (1891). Über Riemann' sche Flächen mit Gegebenen Verzweigungs Punkten, *Math. Ann.* **39**, 1-66.
[44] Hurwitz, A. (1902). Über Die Anzahl Der Riemann'schen Flächen Mit Gegebenen Veazweigungspunkten, *Math. Analen.* **55**, 55-66.
[45] Ingram, R. (1950). Some characters of the symmetric group, *Proc. Amer. Math. Soc.* **1**, 358-369.
[46] Jones, G. (1995). Enumeration of homomorphisms and surface-coverings, *Quarter. J. Math.* **46**, 485-507.
[47] Janvresse, E. (2002). Spectral gap for Kac's model of Boltzmann equation, *Ann. Probab.*
[48] Letac, G. (1981). Problems classiques de probabilitie sur un couple de Gelfand. In Springer *Lecture Notes in Math 861.*
[49] Letac, G. (1982). Les Functions spheriques d'un couple de Gelfand symetrique et les chaincs dc Markov, *Adv. Appl. Probab.* **14**, 272-294.
[50] Liebeck, M. and Shalev, A. (2001). Diameters of finite simple groups: Sharp bounds and applications, *Ann. Math.* **154**, 383-406.
[51] Liebeck, M. and Shalev, A. (2001). Random r s genera tion of finite classical groups, *Bull. Lond. Math. Soc.* **34**, 185-188.
[52] Liu, J. (2001). *Monte Carlo Strategies in Scientific Computing*, Springer, N.Y.
[53] Lovász, L. and Winkler, P. (1998). Mixing times, *Ams. Dimacs. Series* **14**, 189-204.
[54] Lubotzky, A. (1995). Cayley graphs: eigenvalues, expanders and random walks, LMS Lecture Note Series 218, Cambridge University Press, Cambridge.
[55] Lulov, N. (1996). Random walk on the symmetric group generated by conjugacy classes, Ph.D. Dissertation, Dept. of Math., Harvard University.
[56] Lulov, N. and Pak, I. (2002). Rapidly mixing random walks and bounds on characters of the symmetric group. Technical Report, Dept. of Mathematics, M.I.T.
[57] Maslin, D. (2002). On the eigenvalues of Kac's mastser equation. To appear, *Composito Math.*
[58] Pak, I. (1997). Random walks on permutations: Strong uniform time approach. Ph.D. Dissertation, Dept. of Mathematics, Harvard University.
[59] Pak, I. (2000). What do we know about the product replacement algorithm. In *groups and computation* **III**, 301-347, De Gruyter, Berlin.
[60] Porod, U. (1996). The cut-off phenomenon for random reflections, *Ann. Probab.* **24**, 74-99.
[61] Roichman, Y. (1996). Upper bound on characters of the symmetric groups, *Inven. Math.* **125**, 451 485.
[62] Rosenthal, J. (1994). Random rotations: Characters and random walks on $SO(n)$, *Ann. Probab.* **22**, 398-423.
[63] Rothaus, O. and Thompson, J. (1966). A combinatorial problem in the Symmetric Group, *Pacific Jour. Math.* **18**, 175-178.
[64] Roussel, S. (2000). Phenomene de cutoff pour certaines marches aleatoires sur le

groupe symetrique, *Coll. Math.* **86**, 111-135.
[65] Saloff-Coste, L. (1997). Lectures on finite Markov Chains, *Lecture Notes in Math.* **1665**, Springer, Berlin.
[66] Saloff-Coste, L. (2001). Probability on groups: Random walks and invariant diffusions, *Notices of Amer. Math. Soc.* **48**, 968-977.
[67] Schoolfield, C. (1998). Random walks on wreath products of groups and Markov Chains on related Homogeneous spaces. Ph.D. Dissertation, Johns Hopkins University.
[68] Shalev, A. (2000). Asymptotic group theory, *Notices Amer. Math. Soc.* **48**, 383-389.
[69] Silver, J. (1996). Weighted Poincare and exhaustive approximation techniques for scaled Metropolis-Hasings and spectral total variation convergence bounds in infinite commutable Markov chain theory. Ph.D. Dissertation, Dept. of Mathematics, Harvard University.
[70] Stanley, R. (1999). *Enumerative Combinatorics*, Vol. II, Cambridge University Press, Cambridge.
[71] Steinberg, R. (1962). Generators for simple groups, *Canadian Jour. Math.* **14**, 277-283.
[72] Strehl, V. (1996). Minimal transitive products of transpositions – The Reconstruction of a proof by A. Hurwitz. *Seminar Lothaire Combin. 37 Art 537C.*
[73] Suzuki, M. (1982). *Croup Theory*, I., Springer, New York.
[74] Uyemura-Reyes (2002). Random walks, perfect shuffles, and BHR shuffles, Ph.D. Thesis, Dept. of Mathematics, Stanford University.
[75] Wilson, D. (1997). Mixing times of Lozenge tiling and card shuffling Markov Chains. Preprint, Microsoft Research.
[76] Zvonkin, A. (1997). Matrix integrals and map enumeration: An accessible introduction, *Math. Comput. Modeling* **26**, 281-304.