BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 39, Number 1, Pages 51–85
S 0273-0979(01)00920-X
Article electronically published on October 5, 2001

RANDOM MATRIX THEORY OVER FINITE FIELDS

JASON FULMAN

ABSTRACT. The first part of this paper surveys generating functions methods in the study of random matrices over finite fields, explaining how they arose from theoretical need. Then we describe a probabilistic picture of conjugacy classes of the finite classical groups. Connections are made with symmetric function theory, Markov chains, Rogers-Ramanujan type identities, potential theory, and various measures on partitions.

CONTENTS

1. Introduction 51
2. Cycle index techniques 53
   2.1. The General Linear Groups 54
   2.2. Applications 55
   2.3. Generalization to the Classical Groups 62
   2.4. Limitations and Other Methods 64
3. Running example: General linear groups 65
   3.1. Measures on Partitions 65
   3.2. Symmetric Function Theory and Sampling Algorithms 67
   3.3. Sampling for a Given Size: Unipotent Elements 70
   3.4. Markov Chain Approach 73
   3.5. Rogers-Ramanujan Identities 76
4. Upper triangular matrices 78
   4.1. Growth Algorithm for Jordan Form 78
   4.2. Symmetric Functions and Potential Theory 79
Acknowledgements 81
References 81

1. INTRODUCTION

A natural problem is to understand what a typical element of the finite general linear group $GL(n,q)$ "looks like". Many of the interesting properties of a random matrix depend only on its conjugacy class. The following list of questions one could ask are of this type:

1. How many Jordan blocks are there in the rational canonical form of a random matrix?

Received by the editors April 2000, and in revised form April 24, 2001.
2000 *Mathematics Subject Classification.* Primary 60B15, 20G40.

2. What is the distribution of the order of a random matrix?
3. What is the probability that the characteristic polynomial of a random matrix has no repeated factors?
4. What is the probability that the characteristic polynomial of a random matrix is equal to its minimal polynomial?
5. What is the probability that a random matrix is semisimple (i.e. diagonalizable over the algebraic closure $\bar{F}_q$ of the field of q elements)?

As Section 2 will indicate, answers to these questions have applications to the study of random number generators, to the analysis of algorithms in computational group theory, and to other parts of group theory. Section 2 describes a unified approach to answering such probability questions using cycle index generating functions. As an example of its power, it is proved independently in [F1] and [W2] that the $n \to \infty$ limit of the answer to question 4 is $(1 - \frac{1}{q^5})/(1 + \frac{1}{q^3})$. Although algebraic geometry accounts for the fact that this is roughly $1 - \frac{1}{q^3}$ for large q, there is (at present) no other method for deriving this result and generating functions give effective bounds on the convergence rate to the limit. Extensions of the cycle index method to the set of all matrices and to other finite classical groups are sketched. Limitations of cycle indices and other techniques in random matrix theory over finite fields are discussed.

Section 3 gives a purely probabilistic picture of what the conjugacy class of a random element of $GL(n, q)$ looks like. The main object of study is a probability measure $M_{GL,u,q}$ on the set of all partitions of all natural numbers. This measure is connected with the Hall-Littlewood symmetric functions. Exploiting this connection leads to several methods for growing random partitions distributed as $M_{GL,u,q}$ and gives insightful probabilistic proofs of group theoretic results. We hope to convince the reader that the interplay between probability and symmetric functions is beautiful and useful. A method is given for sampling from $M_{GL,u,q}$ conditioned to live on partitions of a fixed size (which amounts to studying the Jordan form of unipotent elements) and for sampling from a q-analog of Plancherel measure (which is related to the longest increasing subsequence problem of random permutations).

Section 3 goes on to describe a probabilistic approach to $M_{GL,u,q}$ using Markov chains. This connection is quite surprising, and we indicate how it leads to a simple and motivated proof of the Rogers-Ramanujan identities. The measure $M_{GL,u,q}$ has analogs for the finite unitary, symplectic, and orthogonal groups. As this is somewhat technical these results are omitted and pointers to the literature are given. However we remark now that while the analogs of the symmetric function theory viewpoint are unclear for the finite symplectic and orthogonal groups, the connections with Markov chains carry over. Thus there is a coherent probabilistic picture of the conjugacy classes of the finite classical groups.

Section 4 surveys probabilistic aspects of conjugacy classes in $T(n, q)$, the group of $n \times n$ upper triangular matrices over the field F_q with 1's along the main diagonal. Actually a simpler object is studied, namely the Jordan form of randomly chosen elements of $T(n, q)$. From work of Borodin and Kirillov, one can sample from the corresponding measures on partitions. We link their results with symmetric function theory and potential theory on Bratteli diagrams.

The field surveyed in this article is young and evolving. The applications to computational group theory call for extensions of probability estimates discussed in Section 2 to maximal subgroups of finite classical groups. It would be marvellous

if the program surveyed here carries over; this happens for the finite affine groups [F9]. The first step is understanding conjugacy classes, and partial results can be found in the thesis [Mu].

We close with a final motivation for the study of conjugacy classes of random matrices over finite fields. The past few years have seen an explosion of interest in eigenvalues of random matrices from compact Lie groups. For the unitary group $U(n, C)$ over the complex numbers, two matrices are in the same conjugacy class if and only if they have the same set of eigenvalues. Hence, at least in this case, which is related to the zeroes of the Riemann zeta function [KeaSn], the study of eigenvalues is the same as the study of conjugacy classes.

As complements to this article, the reader may enjoy the surveys [Py1],[Py2], [Py3],[Sh2],[Sh3] on enumerative and probabilistic questions in group theory. The current article uses probabilistic language, but this is just enumeration in disguise. We do not describe the closely related field of computational group theory but refer the interested reader to the conference volume [FinkKa].

2. Cycle index techniques

Before describing cycle index techniques for the finite classical groups, we mention that the cycle index techniques here are modelled on similar techniques for the study of conjugacy class functions on the symmetric groups. For a permutation π, let $n_i(\pi)$ be the number of length i cycles of π. The cycle index of a subgroup G of S_n is defined as

$$\frac{1}{|G|} \sum_{\pi \in G} \prod_{i \geq 1} x_i^{n_i(\pi)}$$

and is called a cycle index because it stores information about the cycle structure of elements of G. Applications of the cycle index to graph theory and chemical compounds are exposited in [PoRe]. It is standard to refer to the generating function

$$1 + \sum_{n \geq 1} \frac{u^n}{n!} \sum_{\pi \in S_n} \prod_{i \geq 1} x_i^{n_i(\pi)}$$

as the cycle index or cycle index generating function of the symmetric groups. From the fact that there are $\frac{n!}{\prod_i n_i! i^{n_i}}$ elements in S_n with n_i cycles of length i, one deduces Polya's result that this generating function is equal to $\prod_{m \geq 1} e^{\frac{x_m u^m}{m}}$. This allows one to study conjugacy class functions of random permutations (e.g. number of fixed points, number of cycles, the order of a permutation, length of the longest cycle) by generating functions. We refer the reader to [Ko] for results in this direction using analysis and to [ShLl] for results about cycle structure proved by a probabilistic interpretation of the cycle index generating function. Historically important papers in random permutation theory are [ErT], [Gon], and [VeSc].

Subsection 2.1 reviews the conjugacy classes of $GL(n, q)$ and then discusses cycle indices for $GL(n, q)$ and $Mat(n, q)$, the set of all $n \times n$ matrices with entries in the field of q elements. Subsection 2.2 describes applications of cycle index techniques. Subsection 2.3 discusses generalizations of cycle indices to the finite classical groups.

It is useful to recall some standard notation. Let λ be a partition of some non-negative integer $|\lambda|$ into integer parts (row lengths) $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$. We will also write $\lambda \vdash n$ if λ is a partition of n. Let $m_i(\lambda)$ be the number of parts of λ of size i,

and let λ' be the partition dual to λ in the sense that $\lambda'_i = m_i(\lambda) + m_{i+1}(\lambda) + \cdots$. Let $n(\lambda)$ be the quantity $\sum_{i \geq 1}(i-1)\lambda_i$ and let $(\frac{u}{q})_i$ denote $(1-\frac{u}{q}) \cdots (1-\frac{u}{q^i})$.

2.1. **The General Linear Groups.** To begin we follow Kung [Kun] in defining a cycle index for $GL(n,q)$. First it is necessary to understand the conjugacy classes of $GL(n,q)$. As is explained in Chapter 6 of the textbook [Her], an element $\alpha \in GL(n,q)$ has its conjugacy class determined by its rational canonical form. This form corresponds to the following combinatorial data. To each monic non-constant irreducible polynomial ϕ over F_q, associate a partition (perhaps the trivial partition) λ_ϕ of some non-negative integer $|\lambda_\phi|$. Let $deg(\phi)$ denote the degree of ϕ. The only restrictions necessary for this data to represent a conjugacy class are that $|\lambda_z| = 0$ and $\sum_\phi |\lambda_\phi| deg(\phi) = n$. Note that given a matrix α, the vector space V on which it acts uniquely decomposes as a direct sum of spaces V_ϕ where the characteristic polynomial of α on V_ϕ is a power of ϕ and the characteristic polynomials on different summands are coprime. Each V_ϕ decomposes as a direct sum of cyclic subspaces, and the row lengths of λ_ϕ are the dimensions of the subspaces in this decomposition divided by the degree of ϕ.

An explicit representative of this conjugacy class may be given as follows. Define the companion matrix $C(\phi)$ of a polynomial $\phi(z) = z^{deg(\phi)} + \alpha_{deg(\phi)-1} z^{deg(\phi)-1} + \cdots + \alpha_1 z + \alpha_0$ to be:

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ -\alpha_0 & -\alpha_1 & \cdots & \cdots & -\alpha_{deg(\phi)-1} \end{pmatrix}.$$

Let $\phi_1, \cdots, \phi_k$ be the polynomials such that $|\lambda_{\phi_i}| > 0$. Denote the parts of λ_{ϕ_i} by $\lambda_{\phi_i,1} \geq \lambda_{\phi_i,2} \geq \cdots$. Then a matrix corresponding to the above conjugacy class data is

$$\begin{pmatrix} R_1 & 0 & 0 & 0 \\ 0 & R_2 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & R_k \end{pmatrix}$$

where R_i is the matrix

$$\begin{pmatrix} C(\phi_i^{\lambda_{\phi_i,1}}) & 0 & 0 \\ 0 & C(\phi_i^{\lambda_{\phi_i,2}}) & 0 \\ 0 & 0 & \cdots \end{pmatrix}.$$

For example, the identity matrix has λ_{z-1} equal to (1^n) and all other λ_ϕ equal to the empty set. An elementary transvection with $a \neq 0$ in the $(1,2)$ position, ones on the diagonal and zeros elsewhere has λ_{z-1} equal to $(2, 1^{n-2})$ and all other λ_ϕ equal to the empty set. For a given matrix only finitely many λ_ϕ are non-empty. Many algebraic properties of a matrix can be stated in terms of the data parameterizing its conjugacy class. For instance the characteristic polynomial of $\alpha \in GL(n,q)$ is equal to $\prod_\phi \phi^{|\lambda_\phi(\alpha)|}$, and the minimal polynomial of α is equal to $\prod_\phi \phi^{|\lambda_{\phi,1}(\alpha)|}$. Furthermore α is semisimple (diagonalizable over the algebraic closure $\bar{F}_q$) precisely when all $\lambda_\phi(\alpha)$ have largest part at most 1.

To define the cycle index for $Z_{GL(n,q)}$, let $x_{\phi,\lambda}$ be variables corresponding to pairs of polynomials and partitions. Define

$$Z_{GL(n,q)} = \frac{1}{|GL(n,q)|} \sum_{\alpha \in GL(n,q)} \prod_{\phi : |\lambda_\phi(\alpha)|>0} x_{\phi,\lambda_\phi(\alpha)}.$$

Note that the coefficient of a monomial is the probability of belonging to the corresponding conjugacy class and is therefore equal to one over the order of the centralizer of a representative. It is well known (e.g. easily deduced from page 181 of [Mac]) that one over the order of the centralizer of the conjugacy class of $GL(n,q)$ corresponding to the data $\{\lambda_\phi\}$ is

$$\frac{1}{\prod_\phi q^{deg(\phi)\cdot\sum_i (\lambda'_{\phi,i})^2} \prod_{i\geq 1} (\frac{1}{q^{deg(\phi)}})_{m_i(\lambda_\phi)}}.$$

The formulas given for conjugacy class size in [Kun] and [St1] are written in different form; for the reader's benefit they have been expressed here in the form most useful to us. It follows that

$$1 + \sum_{n=1}^{\infty} Z_{GL(n,q)} u^n = \prod_{\phi \neq z} \left[1 + \sum_{n\geq 1} \sum_{\lambda \vdash n} x_{\phi,\lambda} \frac{u^{n\cdot deg(\phi)}}{q^{deg(\phi)\cdot\sum_i (\lambda'_i)^2} \prod_{i\geq 1} \left(\frac{1}{q^{deg(\phi)}}\right)_{m_i(\lambda_\phi)}} \right].$$

This is called the cycle index generating function.

Let $Mat(n,q)$ be the set of all $n \times n$ matrices over the field F_q. Define

$$Z_{Mat(n,q)} = \frac{1}{|GL(n,q)|} \sum_{\alpha \in Mat(n,q)} \prod_{\phi : |\lambda_\phi(\alpha)|>0} x_{\phi,\lambda_\phi(\alpha)}.$$

Analogous arguments [St1] show that

$$1 + \sum_{n=1}^{\infty} Z_{Mat(n,q)} u^n = \prod_{\phi} \left[1 + \sum_{n\geq 1} \sum_{\lambda \vdash n} x_{\phi,\lambda} \frac{u^{n\cdot deg(\phi)}}{q^{deg(\phi)\cdot\sum_i (\lambda'_i)^2} \prod_{i\geq 1} \left(\frac{1}{q^{deg(\phi)}}\right)_{m_i(\lambda_\phi)}} \right].$$

This will be used in Subsection 2.2. Note that the denominator in $Z_{Mat(n,q)}$ is $|GL(n,q)|$, not $|Mat(n,q)|$, since the formula follows from a formula for the size of the orbits of $GL(n,q)$ acting on $Mat(n,q)$ by conjugation. This makes no essential difference for applications.

2.2. **Applications.** This subsection describes applications of cycle indices. The first example is treated in detail, and results for the other examples are sketched.

Example 1 (Cyclic and Separable Matrices)**.** Recall that a matrix $\alpha \in Mat(n,q)$ operating on a vector space V is called cyclic if there is a vector $v_0 \in V$ such that $v_0, v_0\alpha, v_0\alpha^2, \cdots$ span V. As is explained in [NP2], this is equivalent to the condition that the characteristic and minimal polynomials of α are equal.

The need to estimate the proportion of cyclic matrices arose from [NP1] in connection with analyzing the running time of an algorithm for deciding whether or not the group generated by a given set of matrices in $GL(n,q)$ contains the special linear group $SL(n,q)$. Cyclic matrices also arise in recent efforts to improve upon the MeatAxe algorithm for computing modular characters [NP4] and in Example 8 below. John Thompson has asked if every matrix is the product of a cyclic matrix

and a permutation matrix, suggesting that the answer could have applications to finite projective planes.

Letting $c_M(n,q)$ be the proportion of cyclic elements of $Mat(n,q)$, the paper [NP2] proves that

$$\frac{1}{q^2(q+1)} < 1 - c_M(n,q) < \frac{1}{(q^2-1)(q-1)}.$$

The cycle index approach is also informative, yielding a formula for the $n \to \infty$ limit of $C_M(n,q)$, denoted by $c_M(\infty,q)$, together with convergence rates. For the argument two lemmas are useful, as is some notation. Let $N_d(q)$ be the number of monic degree d irreducible polynomials over the field F_q. In all that follows ϕ will denote a monic irreducible polynomial over F_q. Given a power series $f(u)$, let $[u^n]f(u)$ denote the coefficient of u^n in $f(u)$.

Lemma 1.

$$\prod_{\phi}(1 - \frac{u^{deg(\phi)}}{q^{deg(\phi)}}) = 1 - u.$$

Proof. Expanding $\frac{1}{1-\frac{u^{deg(\phi)}}{q^{deg(\phi)}}}$ as a geometric series and using unique factorization in $F_q[x]$, one sees that the coefficient of u^d in the reciprocal of the left hand side is $\frac{1}{q^d}$ times the number of monic polynomials of degree d, hence 1. Comparing with the reciprocal of the right hand side completes the proof. □

Lemma 2. *If the Taylor series of f around 0 converges at $u = 1$, then*

$$lim_{n\to\infty}[u^n]\frac{f(u)}{1-u} = f(1).$$

Proof. Write the Taylor expansion $f(u) = \sum_{n=0}^{\infty} a_n u^n$. Then observe that $[u^n]\frac{f(u)}{1-u}$ $= \sum_{i=0}^{n} a_i$. □

Theorem 1 calculates $c_M(\infty,q)$.

Theorem 1. ([F1],[W2])

$$c_M(\infty,q) = (1 - \frac{1}{q^5})\prod_{r=3}^{\infty}(1 - \frac{1}{q^r}).$$

Proof. Recall that α is cyclic precisely when its characteristic polynomial and minimal polynomials are equal. From Subsection 2.1, these polynomials are equal when all λ_ϕ have at most one part. In the cycle index for $Mat(n,q)$ set $x_{\phi,\lambda} = 1$ if λ has at most 1 part and $x_{\phi,\lambda} = 0$ otherwise. It follows that

$$c_M(n,q) = \frac{|GL(n,q)|}{q^{n^2}}[u^n]\prod_{\phi}(1 + \sum_{j=1}^{\infty}\frac{u^{j\cdot deg(\phi)}}{q^{(j-1)deg(\phi)}(q^{deg(\phi)}-1)}).$$

By Lemma 1 this equation can be rewritten as

$$\begin{aligned} c_M(n,q) &= \frac{|GL(n,q)|}{q^{n^2}}[u^n]\frac{\prod_\phi(1-\frac{u^{deg(\phi)}}{q^{deg(\phi)}})(1+\sum_{j=1}^\infty \frac{u^{j\cdot deg(\phi)}}{q^{(j-1)deg(\phi)}(q^{deg(\phi)}-1)})}{1-u} \\ &= \frac{|GL(n,q)|}{q^{n^2}}[u^n]\frac{\prod_\phi(1+\frac{u^{deg(\phi)}}{q^{deg(\phi)}(q^{deg(\phi)}-1)})}{1-u} \\ &= \frac{|GL(n,q)|}{q^{n^2}}[u^n]\frac{\prod_{d\geq 1}(1+\frac{u^d}{q^d(q^d-1)})^{N_d(q)}}{1-u}. \end{aligned}$$

Recall that a product $\prod_{n=1}^\infty(1+a_n)$ converges absolutely if the series $\sum_{n\geq 1}|a_n|$ converges. Thus using the crude bound $N_d(q)\leq q^d$

$$\prod_{d\geq 1}(1+\frac{u^d}{q^d(q^d-1)})^{N_d(q)}$$

is analytic in a disc of radius greater than 1. Lemma 2 implies that

$$\begin{aligned} c_M(\infty,q) &= lim_{n\to\infty}\frac{|GL(n,q)|}{q^{n^2}}[u^n]\frac{\prod_{d\geq 1}(1+\frac{u^d}{q^d(q^d-1)})^{N_d(q)}}{1-u} \\ &= \prod_{r=1}^\infty(1-\frac{1}{q^r})\prod_{d\geq 1}(1+\frac{1}{q^d(q^d-1)})^{N_d(q)}. \end{aligned}$$

Applying Lemma 1 (with $u=\frac{1}{q}$, $u=\frac{1}{q^2}$ and then $u=\frac{1}{q^5}$) gives

$$\begin{aligned} c_M(\infty,q) &= \prod_{r=3}^\infty(1-\frac{1}{q^r})\prod_{d\geq 1}((1+\frac{1}{q^d(q^d-1)})(1-\frac{1}{q^{2d}})(1-\frac{1}{q^{3d}}))^{N_d(q)} \\ &= \prod_{r=3}^\infty(1-\frac{1}{q^r})\prod_{d\geq 1}(1-\frac{1}{q^{6d}})^{N_d(q)} \\ &= (1-\frac{1}{q^5})\prod_{r=3}^\infty(1-\frac{1}{q^r}). \end{aligned}$$

□

The next challenge is to bound the convergence rate of $c_M(n,q)$ to $c_\infty(n,q)$. Wall [W2] found a strikingly simple way of doing this by relating the cycle index of cyclic matrices to the cycle index of the set of matrices whose characteristic polynomial is squarefree (these matrices are termed separable in [NP2]). To state the result, let $s_M(n,q)$ be the probability that an $n\times n$ matrix is separable. Next let $C_M(u,q)$ and $S_M(u,q)$ be the generating functions defined as

$$C_M(u,q)=1+\sum_{n\geq 1}\frac{u^n q^{n^2}}{|GL(n,q)|}c_M(n,q)$$

$$S_M(u,q)=1+\sum_{n\geq 1}\frac{u^n q^{n^2}}{|GL(n,q)|}s_M(n,q).$$

Lemma 3. ([W2])

$$(1-u)C_M(u,q)=S_M(u/q,q).$$

Proof. The proof of Theorem 1 shows that

$$(1-u)C_M(u,q) = \prod_{d\geq 1}(1+\frac{u^d}{q^d(q^d-1)})^{N_d(q)}.$$

A matrix is separable if and only if all λ_ϕ have size 0 or 1. Hence

$$S_M(u,q) = \prod_{d\geq 1}(1+\frac{u^d}{q^d-1})^{N_d(q)}.$$

The result follows. □

Corollary 1. ([W2])

$$0 < |c_M(n,q) - c_M(\infty,q)| < \frac{1}{q^{n+1}(1-1/q)}.$$

Proof. Taking coefficients of u^{n+1} on both sides of Lemma 3 gives the relation

$$c_M(n+1,q) - c_M(n,q) = \frac{s_M(n+1,q) - c_M(n,q)}{q^{n+1}}.$$

Since $0 \leq |s_M(n+1,q) - c_M(n,q)| \leq 1$ for all n, it follows that

$$|c_M(n,q) - c_M(\infty,q)| \leq \sum_{i=n}^{\infty} |c_M(i+1,q) - c_M(i,q)| \leq \sum_{i=n}^{\infty} \frac{1}{q^{i+1}},$$

as desired. □

Remarks.

1. As mentioned in the introduction, an argument similar to that of Theorem 1 shows that the $n \to \infty$ probability that an element of $GL(n,q)$ is cyclic is $(1-\frac{1}{q^5})/(1+\frac{1}{q^3})$. For large q this goes like $1-1/q^3$. The reason for this is a result of Steinberg [Stei] stating that the set of non-regular elements in an algebraic group has co-dimension 3 (see also [GuLub]). In type A, regular (i.e. centralizer of minimum dimension) and cyclic elements coincide, but not always. For more discussion on this point, see [NP2], [FNP].
2. The generating functions $S_M(u,q)$ and $C_M(u,q)$ have intriguing analytical properties. It is proved in [W2] that

$$S_M(u,q) = \frac{\prod_{d=1}^{\infty}(1-\frac{u^d(u^d-1)}{q^d(q^d-1)})^{N_d(q)}}{1-u}.$$

   Thus $S_M(u,q)$ has a pole at 1 and $S_M(u,q) - \frac{1}{1-u}$ can be analytically extended to the circle of radius q. Analogous properties hold for $C_M(u,q)$ by means of Lemma 3.
3. The limits $s_M(\infty,q)$ and $s_{GL}(\infty,q)$ are in [F1],[W2]. Bounding the rate of convergence of $s_M(n,q)$ to $s_M(\infty,q)$ leads to interesting number theory. Let $p(d)$ be the number of partitions of d and let $p_2(d) = \sum_{i=0}^{d} p(d)$. It is proved in [W2] that

$$\left| \frac{s_M(n,q)q^{n^2}}{|GL(n,q)|} - 1 \right| \leq \sum_{d=n+1}^{\infty} (p_2(d) + qp(d-2))q^{-d} \leq \frac{1}{3}(\frac{4q+27}{2q-3})(\frac{2}{3}q)^{-n}.$$

4. Lehrer [Leh] expresses $s_M(n,q)$ and $s_{GL}(n,q)$ as inner products of characters in the symmetric group and proves a stability result about their expansions in powers of q^{-1}. See also [W2]. The paper [LehSe] gives a topological approach to these stability results.
5. The results of [F2] and [W2] surveyed above are extended to the finite classical groups in [FNP]. The paper [FlJ] gives (intractable) formulas for the chance of being separable in groups such as $SL(n,q)$ (i.e. semisimple and simply connected). Forthcoming work of John Britnell shows that the SL and GL limiting cyclic and separable probabilities coincide and estimates the convergence rate of the SL probabilities to their limits.

Example 2 (Eigenvalue-free matrices)**.** The paper [NP3] studies eigenvalue-free matrices (i.e. matrices without fixed lines) over finite fields as a step in obtaining estimates of cyclic probabilities in orthogonal groups [NP5]. It is interesting that the study of eigenvalue-free matrices was one of the motivations for the original $GL(n,q)$ cycle index papers [Kun],[St1], the latter of which proves that the $n,q \to \infty$ limit of the chance that an element of $GL(n,q)$ has no eigenvalues is $\frac{1}{e}$.

The $n \to \infty$ probability that a random element of S_n has no fixed points is also $\frac{1}{e}$. This is not coincidence; in general the $q \to \infty$ limit of the chance that the characteristic polynomial of a random element of $Mat(n,q)$ factors into n_i degree i irreducible factors is the same as the probability that an element of S_n factors into n_i cycles of degree i. This is proved at the end of [St1] and is extended to finite Lie groups in [F1] using the combinatorics of maximal tori. There is another interesting line of argument which should be mentioned. It is easy to see from the cycle index that the factorization type of the characteristic polynomial of a random element of $Mat(n,q)$ and the factorization type of a random degree n polynomial over F_q have the same distribution as $q \to \infty$. Now the factorization type of a random degree n polynomial over F_q has the same distribution as the cycle type of a random permutation distributed as a q-shuffle on n cards [DiaMcPi], and as $q \to \infty$ a q-shuffle converges to a random permutation. The connection of Lie theory with card shuffling may seem ad hoc, but is really the tip of a deep iceberg [F10].

Example 3 (Characteristic polynomials)**.** The previous example is a special case of the problem of studying the degrees of the factors of the characteristic polynomial of a random matrix. Many results in this direction (all proved used cycle indices) can be found in Stong's paper [St1]. Hansen and Schmutz [HSchm] use cycle index manipulations to prove that if one ignores factors of small degree, then the factorization type of the characteristic polynomial of a random element of $GL(n,q)$ is close to the factorization type of a random degree n polynomial over F_q. More precisely, let $A_{n,l}$ be the set of sequences $(\alpha_{l+1}, \cdots, \alpha_n)$ where α_i is the number of degree i factors of a random polynomial chosen from some measure. Let $Q_n^{(1)}$ be the measure on polynomials arising from characteristic polynomials of random elements of $GL(n,q)$ and let $Q_n^{(2)}$ be the measure arising from choosing a degree n polynomial over F_q uniformly at random. They prove

Theorem 2. ([HSchm]) *There exist constants c_1, c_2 such that for all l with $c_1 log(n) \leq l \leq n$ and $B \subset N^{n-l}$,*

$$|Q_n^{(1)}(A_n(B)) - Q_n^{(2)}(A_n(B))| < c_2/l.$$

The final section of their paper uses this principle to prove results about characteristic polynomials of irreducible factors of random matrices using known results about random polynomials. A useful reference on the distribution of degrees of irreducible factors of random polynomials over finite fields is [ArBarT].

Example 4 (Generating transvections)**.** Recall that the motivation behind Example 1 was a group recognition problem, i.e. trying to determine whether or not the group generated by a given set X of matrices in $GL(n,q)$ contains the special general linear group $SL(n,q)$. However the problem still remains of making the recognition algorithm constructive. For instance if the group generated by X is $GL(n,q)$, it would be desirable to write any element of $GL(n,q)$ as a word in X.

The paper [CeLg] proposes such a constructive recognition algorithm. An essential step involves constructing a transvection, that is a non-identity element of $SL(n,q)$ which has an $n-1$ dimensional fixed space. This in turn is done in two steps. First, find an element α of $GL(n,q)$ conjugate to $\mathrm{diag}(C((z-\tau)^2), R)$ where C is the companion matrix as in Subsection 2.1 and R is semisimple without τ as an eigenvalue. Second, one checks that raising α to the least common multiple of the orders of τ and R gives a transvection.

Thus it is necessary to bound the number of feasible α in the first step. Such α have conjugacy class data $\lambda_{z-\tau} = (2)$, and all other λ_ϕ have largest part at most 1. The cycle index approach gives bounds improving on those in [CeLg]; see [FNP] for the details.

Example 5 (Semisimple matrices)**.** A fundamental problem in computational group theory is to construct an element of order p. Given a group element g with order a multiple of p, this can be done by raising g to an appropriate power. It is proved in [IsKanSp] that if G is a permutation group of degree n with order divisible by p, then the probability that a random element of G has order divisible by p is at least $\frac{1}{n}$.

Their proof reduces the assertion to simple groups and then uses the classification of simple groups. Let us consider the group $GL(n,q)$, which is close enough to simple to be useful for the applications at hand. When p is the characteristic of the field of definition of $GL(n,q)$, an element has order prime to p precisely when it is semisimple. Thus the problem is to study the probability that an element of $GL(n,q)$ is semisimple. The paper [GuLub] shows that if G is a simple Chevalley group, then the probability of not being semisimple is at most $3/(q-1)+2/(q-1)^2$ and thus at most c/q for some constant c as conjectured by Kantor.

As mentioned earlier, a matrix α is semisimple if and only if all $\lambda_\phi(\alpha)$ have largest part size at most 1. Stong [St1] used cycle indices to obtain crude asymptotic bounds for the probability that an element of $GL(n,q)$ is semisimple. The thesis [F1] used the Rogers-Ramanujan identities to prove that the $n \to \infty$ probability that an element of $GL(n,q)$ is semisimple is

$$\prod_{\substack{r=1 \\ r=0,\pm 2 (mod\ 5)}}^{\infty} \frac{(1-\frac{1}{q^{r-1}})}{(1-\frac{1}{q^r})}.$$

The paper [FNP] gives effective bounds for finite n.

Example 6 (Order of a matrix)**.** A natural problem is to study the order of a random matrix. This has been done in [St2] and [Schm]; see also the remarks in Subsection 3.3 and the very preliminary calculations for other classical groups in

[F1]. Shalev [Sh1] uses facts about the distribution of the order of a random matrix together with Aschbacher's study of maximal subgroups of classical groups [As] as key tools in studying the probability that a random element of $GL(n,q)$ belongs to an irreducible subgroup of $GL(n,q)$ that does not contain $SL(n,q)$. As explained in [Sh1] this has a number of applications; for instance it leads to a proof that if x is any non-trivial element of $PSL(n,q)$, then the probability that x and a randomly chosen element y generate $PSL(n,q)$ tends to 1 as $q \to \infty$. Shalev [Sh1] asks for extensions of these results to other finite classical groups.

It is also useful to count elements of given orders (e.g. 2 or 3) in classical groups and their maximal subgroups. The recent paper [CTY] uses cycle indices to perform such enumerations. One motivation for such enumerations is the study of finite simple quotients of $PSL(2,Z)$; a group G is a quotient of $PSL(2,Z)$ if and only if $G =< x,y >$ with $x^2 = y^3 = 1$. For further discussion, see [Sh2].

Example 7 (Random number generators)**.** We follow [Mar],[MarTs] in indicating the relevance of random matrix theory to the study of random number generators. Suppose one wants to test a mechanism for generating a random integer between 0 and $2^{33} - 1$. In base 2 these are length 33 binary vectors. Generating, say, n of these and listing them gives an $n \times 33$ matrix. If the random generator were perfect, the arising matrix would be random. One could choose a statistic such as the rank of a matrix and compare the generation method with theory. Marsaglia and Tsay [MarTs] report that shift-register generators will fail such tests but that congruential generators usually pass. It would be interesting to see how various random number generators perform when tested using other conjugacy class functions of random matrices.

Diaconis and Graham [DiaGr] analyze random walks of the form $X_n = AX_{n-1} + \epsilon_n$ where X_i is a length d $0-1$ vector, A is an element of $GL(n,2)$, and ϵ_n is a random vector of disturbance terms. For more general A (in $GL(n,q)$) this includes the problem of running a pseudo-random number generator with recurrence $Y_n = a_1 Y_{n-1} + \cdots + a_d Y_{n-d} + \epsilon_n$ with $Y_i \in F_q$ and $\epsilon_n \in \{0, \pm 1\}$. They show that the rational canonical form of A is related in a subtle way to the convergence rate of the walk. It would be interesting to understand what happens when A is a random matrix.

Example 8 (Product replacement algorithm)**.** In recent years finite group theory has become much more computational. Given a generating set S of a finite group G, it is natural to seek random elements of G. One approach, implemented in the computer systems GAP and MAGMA, is the product replacement algorithm [CeLgMuNiOb]. Fixing G and some k, one performs a random walk on k-tuples $(g_1, \cdots, g_k)$ of elements of G which generate the group. The walk proceeds by picking an ordered pair (i,j) with $1 \le i \ne j \le n$ uniformly at random and applying one of the following four operations with equal probability:

$$R_{i,j}^{\pm} : (g_1, \cdots, g_i, \cdots, g_k) \mapsto (g_1, \cdots, g_i \cdot g_j^{\pm}, \cdots, g_k)$$
$$L_{i,j}^{\pm} : (g_1, \cdots, g_i, \cdots, g_k) \mapsto (g_1, \cdots, g_j^{\pm} \cdot g_i, \cdots, g_k).$$

These moves map generating k-tuples to generating k-tuples. One starts from any generating k-tuple, applies the algorithm for r steps, and then outputs a random entry of the resulting k-tuple (i.e. a group element).

The product replacement algorithm has superb practical performance (often converging more rapidly than random walk on the Cayley graph), in spite of the

theoretical defects that a random entry of a random generating k-tuple does not have the same distribution as a random element of G and that the convergence rate of the chain on k-tuples to its stationary distribution is unknown. The paper [CeLgMuNiOb], aware of these issues, tests the algorithm against theory, using conjugacy class statistics such as the order of an element, the number of factors of the characteristic polynomial of a random matrix, the degree of the largest irreducible factor of the characteristic polynomial of a random matrix, and the proportion of cyclic matrices in the finite classical groups. In short, understanding properties of random matrices is crucial to their analysis.

A recent effort to understand the performance of the product replacement algorithm uses Kazhdan's property T from the representation theory of Lie groups [LubPa]; the paper [Pa] is a useful survey. Much remains to be done.

Example 9 (Running times of algorithms)**.** One of the main approaches to computing determinants and permanents of integer matrices involves doing the computations for reductions mod prime powers. Section 4.6.4 of [Kn] gives a detailed discussion with references to literature on upper bounds of running times. If one believes that typical matrices one encounters in the real world are like random matrices, this motivates studying random matrices over finite fields. In fact von Neumann's interest in eigenvalues of random matrices with independent normal entries arose from the same heuristic applied to questions in numerical analysis (the introduction of [Ed] gives further discussion of this point).

Examples of algorithms in which properties of random matrices were really needed to bound running times include recognizing when a group generated by a set of matrices contains $SL(n,q)$ [NP1] and the MeatAxe algorithm for computing modular characters [NP4].

Example 10 (Isometry classes of linear codes)**.** Fripertinger [Frip1], [Frip2] considers cycle indices (in the permutation sense) of matrix groups acting on lines. His interest was in understanding properties of random isometry classes of linear codes–a harder problem than understanding random linear codes. The cycle indices he obtains seem quite intractable for theorem proving, but are useful in conjunction with computers. He also gives references to the switching function literature.

Curiously, understanding the permutation action of random matrices of lines comes up in another context. Wieand [Wi] has shown that the eigenvalues of random permutation matrices possess a structure similar to the eigenvalues of matrices from compact Lie groups. Persi Diaconis has suggested that the eigenvalues of representations of finite groups of Lie type (such as the permutation action on lines) may possess similar structure; see [F5] for more in this direction.

2.3. **Generalization to the Classical Groups.** This subsection will focus on the finite unitary groups, with remarks about symplectic and orthogonal groups at the end. These cycle indices were derived in [F1],[F2] and were applied to the problem of estimating proportions of cyclic, separable, and semisimple matrices (these terms were defined in Subsection 2.2) in [FNP]. John Britnell (in preparation) has pushed cycle index techniques through for groups such as $SL(n,q)$.

The unitary group $U(n,q)$ can be defined as the subgroup of $GL(n,q^2)$ preserving a non-degenerate skew-linear form. Recall that a skew-linear form on an n dimensional vector space V over F_{q^2} is a bilinear map $<,>: V \times V \to F_{q^2}$ such that $< \vec{x}, \vec{y} >=< \vec{y}, \vec{x} >^q$ (raising to the qth power is an involution in a field of order

q^2). One such form is given by $< \vec{x}, \vec{y} >= \sum_{i=1}^{n} x_i y_i^q$. Any two non-degenerate skew-linear forms are equivalent, so that $U(n,q)$ is unique up to isomorphism.

Wall [W1] parametrized the conjugacy classes of the finite unitary groups and computed their sizes. To describe his result, an involution on polynomials with non-zero constant term is needed. Given a polynomial ϕ with coefficients in F_{q^2} and non-vanishing constant term, define a polynomial $\tilde{\phi}$ by:

$$\tilde{\phi} = \frac{z^{deg(\phi)} \phi^q(\frac{1}{z})}{[\phi(0)]^q}$$

where ϕ^q raises each coefficient of ϕ to the qth power. Writing this out, a polynomial $\phi(z) = z^{deg(\phi)} + \alpha_{deg(\phi)-1} z^{deg(\phi)-1} + \cdots + \alpha_1 z + \alpha_0$ with $\alpha_0 \neq 0$ is sent to $\tilde{\phi}(z) = z^{deg(\phi)} + (\frac{\alpha_1}{\alpha_0})^q z^{deg(\phi)-1} + \cdots + (\frac{\alpha_{deg(\phi)-1}}{\alpha_0})^q z + (\frac{1}{\alpha_0})^q$. An element $\alpha \in U(n,q)$ associates to each monic, non-constant, irreducible polynomial ϕ over F_{q^2} a partition λ_ϕ of some non-negative integer $|\lambda_\phi|$ by means of rational canonical form. The restrictions necessary for the data λ_ϕ to represent a conjugacy class are that $|\lambda_z| = 0$, $\lambda_\phi = \lambda_{\tilde{\phi}}$, and $\sum_\phi |\lambda_\phi| deg(\phi) = n$.

Using formulas for conjugacy class sizes from [W1] together with some combinatorial manipulations, one obtains the following unitary group cycle index generating function. The products in the theorem are as always over monic irreducible polynomials. The pair $\{\phi, \tilde{\phi}\}$ is unordered.

Theorem 3.

$$1 + \sum_{n=1}^{\infty} \frac{u^n}{|U(n,q)|} \sum_{\alpha \in U(n,q)} \prod_{\phi : |\lambda_\phi(\alpha)| > 0} x_{\phi, \lambda_\phi(\alpha)}$$

$$= \prod_{\phi \neq z, \phi = \tilde{\phi}} \left[1 + \sum_{n \geq 1} \sum_{\lambda \vdash n} x_{\phi,\lambda} \frac{u^{n \cdot deg(\phi)}}{q^{deg(\phi) \cdot \sum_i (\lambda_i')^2} \prod_{i \geq 1} \left(\frac{1}{q^{deg(\phi)}} \right)_{m_i(\lambda)}} \right]_{u \mapsto -u, q \mapsto -q}$$

$$\cdot \prod_{\{\phi, \tilde{\phi}\}, \phi \neq \tilde{\phi}} \left[1 + \sum_{n \geq 1} \sum_{\lambda \vdash n} x_{\phi,\lambda} x_{\tilde{\phi},\lambda} \frac{u^{n \cdot deg(\phi)}}{q^{deg(\phi) \cdot \sum_i (\lambda_i')^2} \prod_{i \geq 1} \left(\frac{1}{q^{deg(\phi)}} \right)_{m_i(\lambda)}} \right]_{u \mapsto u^2, q \mapsto q^2}.$$

One interesting theoretical result concerning the cycle index of $U(n,q)$ is the following functional equation. Letting $C_{GL}(u,q)$ and $C_U(u,q)$ be the cycle index generating functions for cyclic matrices in the general linear and unitary groups respectively, the functional equation states that

$$C_{GL}(u,q) C_U(-u,-q) = C_{GL}(u^2, q^2).$$

The paper [FNP] proves that this relation holds whenever the condition on the partitions λ_ϕ is independent of the polynomial ϕ. In the current example, a matrix is cyclic if and only if all λ_ϕ have at most one row. This condition is independent of ϕ.

Cycle indices for the symplectic and orthogonal groups are a bit trickier to establish from Wall's formulas. To the treatment in [F1],[F2] we add a remark which should be very helpful to anyone trying to use those cycle indices. Those papers only wrote out an explicit formula for the cycle index for the sum of +,− type orthogonal groups. To solve for an individual orthogonal group, it is necessary to average that formula with a formula for the difference of +,− type orthogonal

groups (this procedure is carried out in a special case in [FNP]). In general, the formula for the difference of orthogonal groups is obtained from the formula for the sum of orthogonal groups as follows. First, for the polynomials $z \pm 1$, replace terms corresponding to partitions with an odd number of odd parts by their negatives. Second, for polynomials invariant under $\tilde{}$, replace terms corresponding to partitions of odd size by their negatives.

2.4. **Limitations and Other Methods.** Cycle index techniques, while very useful, also have their limitations and are not always the best way to proceed, as the following examples demonstrate.

Example 1 (Primitive prime divisor elements)**.** For integers $b, e > 1$ a primitive prime divisor of $b^e - 1$ is a prime dividing $b^e - 1$ but not dividing $b^i - 1$ for any i with $1 \leq i < e$. An element of $GL(n, q)$ is called a primitive prime divisor (ppd) element if its order is divisible by a primitive prime divisor of $q^e - 1$ with $n/2 < e \leq n$. This is a conjugacy class function. The analysis in [NiP] derives elegant bounds on the proportions of ppd elements in the finite classical groups and applies them to the group recognition problem for classical groups over finite fields (determining when a group generated by a set of matrices contains $SL(n, q)$). We do not see how to get comparable bounds using generating function techniques.

Example 2 (Proportions of semisimple elements in exceptional groups)**.** Although Example 5 of Section 2.2 was estimating proportions of semisimple matrices, this was only for the finite classical groups, where the index n can take an infinite number of values. Cycle indices don't seem useful unless there is a tower of groups of varying rank available.

Fortunately the computer package CHEVIE permits calculations precisely in finite rank cases such as the exceptional groups. Indeed this is how [GuLub] obtained estimates of the proportions of semisimple elements in the exceptional groups.

Example 3 (Non-uniform distributions on matrices)**.** The cycle indices give useful information about conjugacy class functions when the matrix is chosen uniformly at random. However there are other distributions on matrices which one could study and for which cycle index methods (at present) cannot be applied.

One example is random $n \times n$ matrices where the matrix entries are chosen independently according to a given probability distribution on F_q. Charlap, Rees, and Robbins [ChReRo] show that if the probability distribution is not concentrated on any proper affine subspace of F_q, then as $n \to \infty$ the probability that the matrix is invertible is the same as for a uniform matrix. They use Moebius inversion on the lattice of subspaces of an n dimensional vector space and the Poisson summation formula. Is the same true for other natural conjugacy class functions? We expect that the answer is yes, which can be regarded as a type of "universality" result for the asymptotic description of random elements of $GL(n, q)$ to be given in Subsection 3.1. Analogous universality results are known for matrices with complex entries [So]. For further information on the rank of random $0-1$ matrices, see [BKW] for sparse matrices, [Bo] for a survey of results on the rank over the real numbers, and also the discussion of work of Rudvalis and Shinoda in Subsection 3.2. Elkies [El] studies the rank of Hankel random matrices with non-uniform entries, and Chapter 15 of [MaSl] relates the rank of random matrices to the weight distribution of Reed-Muller codes.

It is conceivable that cycle index techniques will be able to handle certain natural non-uniform distributions on $GL(n,q)$. Non-uniform distributions which depend only on the conjugacy class of an element can be handled by rescaling. (This happens for the Ewens distributions on permutations–important in population genetics–which picks a permutation with probability proportional to $\theta^{n(\pi)}$ where $0 < \theta \leq 1$ and $n(\pi)$ is the number of cycles of π). For the symmetric groups one can make a cycle index for distributions such as a q-riffle shuffle on a deck of cards [DiaMcPi],[F10], even though the underlying distribution on permutations is not a conjugacy class function!

Example 4 (Random generation and maximal subgroups)**.** As the survey [Sh3] explains, understanding the maximal subgroups of a finite group G gives one approach to studying quantities such as the probability $P_{2,3}(G)$ that x, y generate G where x is a random element of order 2 and y is a random element of order 3. Letting $i_r(G)$ denote the number of elements of order r in a group G, one can easily see that

$$1 - P_{2,3}(G) \leq \sum_{\substack{M \subset G \\ M \ maximal}} \frac{i_2(M) i_3(M)}{i_2(G) i_3(G)},$$

since if x, y do not generate G, then they lie in some maximal subgroup M. Using this, Liebeck and Shalev [LiSh] show that if G is a simple classical group other than $PSp_4(q)$, then $P_{2,3}(G) \to 1$ as $|G| \to \infty$. The paper [LiSh2] is another excellent example in which understanding the maximal subgroups of finite classical groups leads to powerful results.

3. Running example: General linear groups

The purpose of this section is to give different ways of understanding the conjugacy class of a random element of $GL(n,q)$. The analogous theory for other finite classical groups is mentioned in passing but is not treated in detail as many of the main ideas can be communicated using $GL(n,q)$. Subsection 3.1 will show how this leads naturally to the study of certain probability measures $M_{GL,u,q}$ on the set of all partitions of all natural numbers. Connections with symmetric function theory lead to several ways of growing random partitions distributed according to $M_{GL,u,q}$. One consequence is a motivated proof of the Rogers-Ramanujan identities.

3.1. Measures on Partitions.

The goal is to obtain a probabilistic description of the conjugacy class of a random element of $GL(n,q)$. The ideas are based on [F1]. The following definition will be fundamental.

Definition. The measure $M_{GL,u,q}$ on the set of all partitions of all natural numbers is defined by

$$M_{GL,u,q}(\lambda) = \prod_{r=1}^{\infty} (1 - \frac{u}{q^r}) \frac{u^{|\lambda|}}{q^{\sum_i (\lambda_i')^2} \prod_i (\frac{1}{q})_{m_i(\lambda)}}.$$

The motivation for this definition will be clear from Theorem 4. The measure $M_{GL,u,q}$, while seemingly complicated, does have some nice combinatorial properties. For instance for partitions of a fixed size, this measure respects the dominance order on partitions (in this partial order $\lambda \geq \mu$ if and only if $\lambda_1 + \cdots + \lambda_i \geq \mu_1 + \cdots + \mu_i$ for all $i \geq 1$). In work with Bob Guralnick we actually needed this property.

Lemma 4 proves that for $q > 1$ and $0 < u < 1$, the measure $M_{GL,u,q}$ is in fact a probability measure. There are at least three other proofs of this fact: an argument using q series, specializing an identity about Hall-Littlewood polynomials, or a slick argument using Markov chains and an identity of Cauchy. This third argument will be given in Subsection 3.4.

Lemma 4. *If $q > 1$ and $0 < u < 1$, then $M_{GL,u,q}$ defines a probability measure.*

Proof. $M_{GL,u,q}$ is clearly non-negative when $q > 1$ and $0 < u < 1$. Stong [St1] established an equation which is equivalent to the sought identity

$$\sum_{\lambda} \frac{u^{|\lambda|}}{q^{\sum_i (\lambda_i')^2} \prod_i (\frac{1}{q})_{m_i(\lambda)}} = \prod_{r=1}^{\infty} (\frac{1}{1 - \frac{u}{q^r}}).$$

As some effort is required to see this equivalence, we derive the identity directly using Stong's line of reasoning.

First observe that unipotent elements of $GL(n,q)$ correspond to nilpotent $n \times n$ matrices (subtract the identity matrix), and that the number of nilpotent $n \times n$ matrices is $q^{n(n-1)}$ by the Fine-Herstein theorem [FineHer]. The number of unipotent elements in $GL(n,q)$ can be evaluated in another way using the cycle index of the general linear groups. Namely set $x_{\phi,\lambda} = 1$ if $\phi = z - 1$ and set $x_{\phi,\lambda} = 0$ otherwise. One concludes that

$$\sum_{\lambda \vdash n} \frac{1}{q^{\sum_i (\lambda_i')^2} \prod_i (\frac{1}{q})_{m_i(\lambda)}} = \frac{q^{n(n-1)}}{|GL(n,q)|}.$$

Now multiply both sides by u^n, sum in n, and apply Euler's identity (page 19 of [A1]):

$$\sum_{n=0}^{\infty} \frac{u^n q^{\binom{n}{2}}}{(q^n - 1) \cdots (q-1)} = \prod_{r=1}^{\infty} (\frac{1}{1 - \frac{u}{q^r}}).$$

□

The measure $M_{GL,u,q}$ is a fundamental object for understanding the probability theory of conjugacy classes of $GL(n,q)$. This emerges from Theorem 4.

Theorem 4.

1. *Fix u with $0 < u < 1$. Then choose a random natural number N with probability of getting n equal to $(1-u)u^n$. Choose α uniformly in $GL(N,q)$. Then as ϕ varies, the random partitions $\lambda_\phi(\alpha)$ are independent random variables, with λ_ϕ distributed according to the measure $M_{GL,u^{deg(\phi)},q^{deg(\phi)}}$.*
2. *Choose α uniformly in $GL(n,q)$. Then as $n \to \infty$, the random partitions $\lambda_\phi(\alpha)$ converge in finite dimensional distribution to independent random variables, with λ_ϕ distributed according to the measure $M_{GL,1,q^{deg(\phi)}}$.*

Proof. Recall the cycle index factorization

$$1 + \sum_{n=1}^{\infty} Z_{GL(n,q)} u^n = \prod_{\phi \neq z} \left[1 + \sum_{n \geq 1} \sum_{\lambda \vdash n} x_{\phi,\lambda} \frac{u^{n \cdot deg(\phi)}}{\prod_\phi q^{deg(\phi) \cdot \sum_i (\lambda_i')^2} \prod_{i \geq 1} (\frac{1}{q^{deg(\phi)}})_{m_i}} \right].$$

Setting all $x_{\phi,\lambda}$ equal to 1 and using Lemma 4 show that

$$\frac{1}{1-u} = \prod_{\phi \neq z} \prod_{r=1}^{\infty} (\frac{1}{1 - \frac{u^{deg(\phi)}}{q^{r \cdot deg(\phi)}}}).$$

Taking reciprocals and multiplying by the cycle index factorization show that

$$(1-u)+\sum_{n=1}^{\infty} Z_{GL(n,q)}(1-u)u^n$$
$$=\prod_{\phi\neq z}\left(M_{GL,u^{deg(\phi)},q^{deg(\phi)}}(\emptyset)+\sum_{\lambda:|\lambda|>0} M_{GL,u^{deg(\phi)},q^{deg(\phi)}}(\lambda)x_{\phi,\lambda}\right).$$

This proves the first assertion of the theorem. For the second assertion, use Lemma 2 from Subsection 2.2. □

Remarks. 1. Theorem 4 has an analog for the symmetric groups [ShLl]. The statement is as follows. *Fix u with $0<u<1$. Then choose a random natural number N with probability of getting n equal to $(1-u)u^n$. Choose π uniformly in S_N. Letting n_i be the number of i-cycles of π, any finite number of the random variables n_i are independent, with n_i distributed as a Poisson with mean $\frac{u^i}{i}$. Furthermore if one chooses π uniformly in S_n and lets $n\to\infty$, then the random variables n_i are independent random variables, with n_i distributed as a Poisson($\frac{1}{i}$).*

2. The idea of performing an auxiliary randomization of n is a mainstay of statistical mechanics, known as the grand canonical ensemble. For a clear discussion see Sections 1.7, 1.9, and 4.3 of [Fey].

3.2. **Symmetric Function Theory and Sampling Algorithms.** The aim of this subsection is two-fold. First, the measures $M_{GL,u,q}$ are connected with the Hall-Littlewood symmetric functions. Then we indicate how this connection can be exploited to give probabilistic methods for growing random partitions distributed as $M_{GL,u,q}$. The purpose is not to drown the reader in formulas, but rather to show that the connection between symmetric functions and probability is deep, beautiful, and useful in both directions. The results on this section are based on [F1] and [F3], except for the remark on how to make the algorithms terminate in finite time, which is joint with Mark Huber.

To begin, we recall the Hall-Littlewood symmetric functions, which arise in many parts of mathematics: enumeration of p groups, representation theory of $GL(n,q)$, and counting automorphisms of modules. The basic references for Hall-Littlewood polynomials P_λ is Chapter 3 of [Mac], which offers the following definition for a partition λ with at most n parts:

$$P_\lambda(x_1,\cdots,x_n;t)=\left[\frac{1}{\prod_{i\geq 0}\prod_{r=1}^{m_i(\lambda)}\frac{1-t^r}{1-t}}\right]\sum_{w\in S_n} w\left(x_1^{\lambda_1}\cdots x_n^{\lambda_n}\prod_{i<j}\frac{x_i-tx_j}{x_i-x_j}\right).$$

Here w is a permutation acting on the x-variables by sending x_i to $x_{w(i)}$. Recall that $m_i(\lambda)$ is the number of parts of λ of size i. Also $m_0(\lambda)$ is defined as $n-\lambda_1'$. At first glance it is not obvious that these are polynomials, but the denominators cancel out after the symmetrization. The Hall-Littlewood polynomials interpolate between the Schur functions ($t=0$) and the monomial symmetric functions ($t=1$).

Theorem 5 relates the measures $M_{GL,u,q}$ to the Hall-Littlewood polynomials. Recall that $n(\lambda)=\sum_i(i-1)\lambda_i=\sum_i\binom{\lambda_i'}{2}$.

Theorem 5.

$$M_{GL,u,q}(\lambda) = \prod_{i=1}^{\infty}(1-\frac{u}{q^i})\frac{P_\lambda(\frac{u}{q},\frac{u}{q^2},\cdots;\frac{1}{q})}{q^{n(\lambda)}}.$$

Proof. From the above formula for Hall-Littlewood polynomials, it is clear that the only surviving term in the specialization $P_\lambda(\frac{u}{q},\frac{u}{q^2},\cdots;\frac{1}{q})$ is the term when w is the identity. The rest is a simple combinatorial verification. (Alternatively, one could use "principal specialization" formulas for Macdonald polynomials on page 337 of [Mac].) □

Remark. The paper [F3] gives symmetric function theoretic generalizations of the measure $M_{GL,u,q}$ on partitions. In the case of Schur functions s_λ, this measure depends on two infinite sets of variables x_i, y_i and assigns a partition λ mass equal to $s_\lambda(x_i)s_\lambda(y_i)\prod_{i,j}(1-x_iy_j)$. It is remarkable that this measure arose independently in work of the random matrix community relating the distribution of the lengths of increasing subsequences of random permutations to the distribution of eigenvalues of random GUE matrices (these matrices have complex entries). To elaborate, the Robinson-Schensted-Knuth correspondence associates a random partition of size n to a random permutation of size n, and the shape of the partition encodes information about the longest increasing subsequence of the permutation. Choosing the size of the symmetric group randomly (according to a Poisson distribution) gives a probability measure on the set of all partitions of all natural numbers which is a special case of the above Schur function measure. Then the coordinate change $h_j = \lambda'_1 + \lambda_j - j$ maps the set of row lengths $\{\lambda_j\}$ of the partition to a set of distinct integers $\{h_j\}$. These h_j can be viewed as positions of electrostatic charges repelling each other, and from this viewpoint the measure on subsets of the integers bears a striking resemblance to the eigenvalues of a random GUE matrix. This fantastic heuristic can be made precise and led to a solution of the long-standing conjecture relating lengths of increasing subsequences of permutations to eigenvalues of random matrices. For these developments see [BOOl],[Jo] and the many references therein.

Now we return to the measure $M_{GL,u,q}$ and describe an algorithm for growing random partitions according to this measure.

The Young Tableau Algorithm

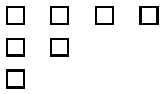
Step 0: Start with $N = 1$ and λ the empty partition. Also start with a collection of coins indexed by the natural numbers, such that coin i has probability $\frac{u}{q^i}$ of heads and probability $1 - \frac{u}{q^i}$ of tails.

Step 1: Flip coin N.

Step 2a: If coin N comes up tails, leave λ unchanged, set $N = N + 1$ and go to Step 1.

Step 2b: If coin N comes up heads, choose an integer $S > 0$ according to the following rule. Set $S = 1$ with probability $\frac{q^{N-\lambda'_1}-1}{q^N-1}$. Set $S = s > 1$ with probability $\frac{q^{N-\lambda'_s}-q^{N-\lambda'_{s-1}}}{q^N-1}$. Then increase the size of column s of λ by 1 and go to Step 1.

As an example of the Young Tableau Algorithm, suppose we are at Step 1 with λ equal to the following partition:



Suppose also that $N = 4$ and that coin 4 had already come up heads once, at which time we added to column 1, giving λ. We flip coin 4 again and get heads, going to Step 2b. We add a box to column 1 with probability $\frac{q-1}{q^4-1}$, to column 2 with probability $\frac{q^2-q}{q^4-1}$, to column 3 with probability $\frac{q^3-q^2}{q^4-1}$, to column 4 with probability 0, and to column 5 with probability $\frac{q^4-q^3}{q^4-1}$. We then return to Step 1.

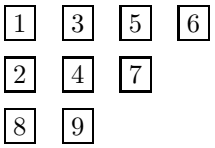
Theorem 6. *For $0 < u < 1$ and $q > 1$, the Young Tableau Algorithm generates partitions which are distributed according to the measure $M_{GL,u,q}$.*

To give insight into the proof of Theorem 6, we remark that it was deduced by proving a stronger result (Theorem 7) inductively and then taking the $N \to \infty$ limit. As is clear from the statement of Theorem 7, the connection with Hall-Littlewood polynomials (in particular the ability to truncate them) was crucial. It is unlikely that the Young Tableau Algorithm would have been discovered without this connection.

Theorem 7. *Let $P^N(\lambda)$ be the probability that the algorithm outputs λ when coin N comes up tails. Then*

$$P^N(\lambda) = \begin{cases} \frac{u^{|\lambda|}(\frac{u}{q})_N(\frac{1}{q})_N}{(\frac{1}{q})_{N-\lambda_1'}} \frac{P_\lambda(\frac{1}{q},\cdots,\frac{1}{q^N},0,\cdots;0,\frac{1}{q})}{q^{n(\lambda)}} & \text{if } \lambda_1' \le N \\ 0 & \text{if } \lambda_1' > N. \end{cases}$$

Next we explain why the Young Tableau Algorithm is called that. A standard Young tableau T of size n is a partition of n with each box filled by one of $\{1, \cdots, n\}$ such that each of $\{1, \cdots, n\}$ appears exactly once and the numbers increase in each row and column of T. For instance,



is a standard Young tableau. Standard Young tableaux are important in combinatorics and representation theory. The Young Tableau Algorithm is so named because numbering the boxes in the order in which they are created gives a standard Young tableau. Thus although our initial interest was in the measure $M_{GL,u,q}$ on partitions, the Young Tableau Algorithm yields more: a probability measure on standard Young tableaux. One consequence of this is a (new) representation of prinicipally specialized Hall-Littlewood polynomials as a sum of certain weights over standard Young tableaux.

Let us indicate an application of this probability measure on standard Young tableaux. Rudvalis and Shinoda [RuShi] studied the distribution of fixed vectors for the classical groups over finite fields. Let $G = G(n)$ be a classical group (i.e. one of GL, U, Sp, or O) acting on an n dimensional vector space V over a finite field F_q (in the unitary case F_{q^2}) in its natural way. Let $P_{G,n}(k,q)$ be the chance that

an element of G fixes a k dimensional subspace and let $P_{G,\infty}(k,q)$ be the $n \to \infty$ limit of $P_{G,n}(k,q)$. They found (in a 76-page unpublished work) beautiful formulas for $P_{G,\infty}(k,q)$. Their formulas are (setting $x = \frac{1}{q}$):

1. $P_{GL,\infty}(k,q) = [\prod_{r=1}^{\infty}(1-x^r)] \frac{x^{k^2}}{(1-x)^2 \cdots (1-x^k)^2}$
2. $P_{U,\infty}(k,q) = \left[\prod_{r=1}^{\infty} \frac{1}{1+x^{2r-1}}\right] \frac{x^{k^2}}{(1-x^2)\cdots(1-x^{2k})}$
3. $P_{Sp,\infty}(k,q) = \left[\prod_{r=1}^{\infty} \frac{1}{1+x^r}\right] \frac{x^{\frac{k^2+k}{2}}}{(1-x)\cdots(1-x^k)}$
4. $P_{O,\infty}(k,q) = \left[\prod_{r=0}^{\infty} \frac{1}{1+x^r}\right] \frac{x^{\frac{k^2-k}{2}}}{(1-x)\cdots(1-x^k)}$.

From a probabilistic perspective, it is very natural to try to interpret the factorizations in these formulas as certain random variables being independent (the paper [RuShi] gives no insight as to why these formulas have a product form). The Young Tableau Algorithm leads to such an understanding for the finite general linear and unitary groups; see [F3] for details.

Remarks. 1. A skew diagram is the set theoretic difference between partitions μ, λ with $\mu \subseteq \lambda$, and a horizontal strip is a skew diagram with at most one square in each column. There is another algorithm for growing random partitions distributed according to $M_{GL,u,q}$ in which one tosses coins and adds horizontal strips (as opposed to a box at a time). Details are in [F3].

2. (Joint with Mark Huber) We indicate how to make the Young Tableau Algorithm run on a computer, so as to terminate in finite time (clearly one can't flip infinitely many coins). Let a_N be the number of times that coin N comes up heads; the idea is to first determine the random vector $(a_1, a_2, \cdots)$ and then grow the partitions as in Step 2b of the Young Tableau Algorithm. So let us explain how to determine $(a_1, a_2, \cdots)$. For $N \geq 1$ let $t^{(N)}$ be the probability that all tosses of all coins numbered N or greater are tails. For $N \geq 1$ and $j \geq 0$ let $t_j^{(N)}$ be the probability that some toss of a coin numbered N or greater is a head and that coin N comes up heads j times. It is simple to write down expressions for $t^{(N)}, t_0^{(N)}, t_1^{(N)}, \cdots$ and clearly $t^{(N)} + \sum_{j \geq 0} t_j^{(N)} = 1$.

   The basic operation a computer can perform is to produce a random variable U distributed uniformly in the interval $[0,1]$. By dividing $[0,1]$ into intervals of length $t^{(1)}, t_0^{(1)}, t_1^{(1)}, \cdots$ and seeing where U is located, one arrives at the value of a_1. Furthermore, if U landed in the interval of length $t^{(1)}$, then all coins come up tails and the algorithm is over. Otherwise, move on to coin 2, dividing $[0,1]$ into intervals of length $t^{(2)}, t_0^{(2)}, t_1^{(2)}, \cdots$ and so on.

   For $0 < u < 1$ and q the size of a finite field, this algorithm terminates quickly. The probability of the algorithm stopping after the generation of the first uniform in $[0,1]$ is $\prod_{i=1}^{\infty}(1-u/q^i) \geq \prod_{i=1}^{\infty}(1-1/q^i) > 1-1/q-1/q^2 \geq 1/4$ where the second inequality is Corollary 3.6 of [NP2]. Should it be necessary to generate future uniforms, the same argument shows that the algorithm stops after each one with probability at least $1/2$.

3.3. **Sampling for a Given Size: Unipotent Elements.** An element of $GL(n,q)$ is called unipotent if all of its eigenvalues are 1; a theorem of Steinberg asserts that the number of unipotent elements in $GL(n,q)$ is $q^{n(n-1)}$ (this is the square of the

order of a q-Sylow subgroup if q is prime). Unipotent elements are interesting because any element α in $GL(n,q)$ can be written uniquely as the product $\alpha_s\alpha_u$ where α_s is semisimple and α_u is unipotent.

Thus it is natural to study the random partition λ_{z-1} for unipotent elements in $GL(n,q)$. This is the same as conditioning the measure $M_{GL,u,q}$ to live on partitions of size n. This subsection explains how to modify the sampling method of Subsection 3.2 to sample from this conditioned version of $M_{GL,u,q}$ and also from a q-analog of Plancharel measure (related to the longest increasing subsequence problem). These results are joint with Mark Huber.

Algorithm for sampling from $M_{GL,u,q}$ given that $|\lambda| = n$

Step 0: Start with $N = 1$ and λ the empty partition.
Step 1: If $n = 0$, then stop. Otherwise set $h = 1 - \frac{1}{q^n}$.
Step 2: Flip a coin with probability of heads h.
Step 2a: If the toss of Step 2 comes up tails, increase the value of N by 1 and go to Step 2.
Step 2b: If the toss of Step 2 comes up heads, decrease the value of n by 1, increase λ according to the rule of Step 2b of the Young Tableau Algorithm (which depends on N), and then go to Step 1.

Theorem 8 will show that the above algorithm samples from $M_{GL,u,q}$ conditioned to live on partitions of size n. It is perhaps surprising that unlike the Young Tableau Algorithm, the probability of a coin coming up heads is independent of the coin number; it depends only on the number of future boxes needed to get a partition of size n.

Lemma 5. *Let N_i be the number of times that coin i comes up heads in the Young Tableau Algorithm with $u = 1$, and let $\vec{N_i}$ be the infinite vector with ith component N_i.*

1. *The probability that $\vec{N_i} = \vec{n_i}$ is $\frac{\prod_{r=1}^{\infty}(1-\frac{1}{q^i})}{q^{\sum_i in_i}}$.*
2. $$\sum_{\vec{n_i}:\sum n_i = a} \frac{1}{q^{\sum_i in_i}} = \frac{1}{q^a(\frac{1}{q})_a}.$$

Proof. The first assertion is clear. The second assertion is well known in the theory of partitions, but we argue probabilistically. Multiply both sides by $\prod_{r=1}^{\infty}(1-\frac{1}{q^i})$. Then note from the first assertion that the left hand side is the $M_{GL,1,q}$ chance of having a partition of size a. Now use the second equation in the proof of Lemma 4 in Subsection 3.1. □

For Theorem 8 the notation Prob. is shorthand for the probability of an event.

Theorem 8. *The algorithm for sampling from $M_{GL,u,q}$ conditioned to live on partitions on size n is valid.*

Proof. From the formula for $M_{GL,u,q}$, the conditioned measure for $M_{GL,u,q}$ is the same as for $M_{GL,1,q}$. Now let n_i be the number of times that coin i comes up heads in the Young Tableau Algorithm. Letting $|$ denote conditioning, it suffices to show that

$$Prob.(n_i \geq 1 | \sum_{j \geq i} n_j = s) = 1 - \frac{1}{q^s}.$$

In fact (for reasons to be explained later) we compute a bit more, namely the conditional probability that $n_i = a$ given that $\sum_{j \geq i} n_j = s$. By definition this conditional probability is the ratio

$$\frac{Prob.(n_i = a, \sum_{j \geq i} n_j = s)}{Prob.(\sum_{j \geq i} n_j = s)}.$$

The numerator and denominator are computed using Lemma 6 as follows:

$$\begin{aligned} Prob.(n_i = a, \sum_{j \geq i} n_j = s) &= \sum_{a_{i+1} + \cdots = s-a} \frac{\prod_{r=i}^{\infty}(1-1/q^r)}{q^{ia} q^{\sum_{j \geq i+1} j a_j}} \\ &= \sum_{a_{i+1} + \cdots = s-a} \frac{\prod_{r=i}^{\infty}(1-1/q^r)}{q^{is} q^{\sum_{j \geq i+1} (j-i) a_j}} \\ &= \sum_{a_1 + \cdots = s-a} \frac{\prod_{r=i}^{\infty}(1-1/q^r)}{q^{is} q^{\sum_{j \geq 1} a_j}} \\ &= \frac{\prod_{r=i}^{\infty}(1-1/q^r)}{q^{is} q^{s-a} (\frac{1}{q})_{s-a}}. \end{aligned}$$

$$\begin{aligned} Prob.(\sum_{j \geq i} n_j = s) &= \sum_{a_i + \cdots = s} \frac{\prod_{r=i}^{\infty}(1-1/q^r)}{q^{\sum_{j \geq i} j a_j}} \\ &= \sum_{a_i + \cdots = s} \frac{\prod_{r=i}^{\infty}(1-1/q^r)}{q^{(i-1)s + \sum_{j \geq i} (j-(i-1)) a_j}} \\ &= \sum_{a_1 + \cdots = s} \frac{\prod_{r=i}^{\infty}(1-1/q^r)}{q^{(i-1)s + \sum_{j \geq 1} j a_j}} \\ &= \frac{\prod_{r=i}^{\infty}(1-1/q^r)}{q^{is} (\frac{1}{q})_s}. \end{aligned}$$

Thus $Prob.(n_i = 0 | \sum_{j \geq i} n_j = s) = \frac{1}{q^s}$ and the result follows. □

As mentioned in Subsection 3.2 there is a natural measure $M_{Pl,q}$ on the set of all partitions of all integers which when conditioned to live on partitions of a given size gives a q-analog of Plancherel measure, which is related to longest increasing subsequences in non-uniform random permutations [F3]. It is beyond the scope of this paper to survey the literature on longest increasing subsequences, but an accessible introduction is [AlDia]. In what follows $J_a(q)$ is the polynomial discussed on pages 52-54 of [F1], $h(s)$ denotes the hook-length of a box in λ [Mac] and $[n] = \frac{q^n - 1}{q-1}$ is the q-analog of the number n. Recall that a skew diagram is the set theoretic difference between partitions μ, λ with $\mu \subseteq \lambda$ and that a horizontal strip is a skew diagram with at most one square in each column.

Algorithm for sampling from $M_{Pl,q}$ for $q > 1$ given that $|\lambda| = n$

Step 0: Start with λ the empty partition.

Step 1: If $n = 0$, then stop. Otherwise choose a with $0 \leq a \leq n$ with probability

$$\frac{q^{n^2}(1 - \frac{1}{q^{n-a+1}})^2 \cdots (1 - \frac{1}{q^n})^2}{q^{(n-a)^2 + n} (\frac{1}{q})_a} \frac{J_{n-a}(q)}{J_n(q)}.$$

Then increase λ to Λ with probability

$$(1-\frac{1}{q})\cdots(1-\frac{1}{q^a})\frac{q^{n(\lambda)}\prod_{s\in\lambda}(1-\frac{1}{q^{h(s)}})}{q^{n(\Lambda)}\prod_{s\in\Lambda}(1-\frac{1}{q^{h(s)}})}$$

if $\Lambda-\lambda$ is a horizontal strip of size a and with probability 0 otherwise. Finally replace n by $n-a$ and repeat Step 1.

Using Lemma 6, Theorem 9 proves that the algorithm for sampling from $M_{Pl,q}$ conditioned to live on $|\lambda|=n$ works. We omit the details, which (given the background material in [F1]) are analogous to the case of $M_{GL,u,q}$.

Lemma 6. *Let N_i be the number of times that coin i comes up heads in the algorithm from* [F3] *for sampling from the measure $M_{Pl,q}$ and let $\vec{N}_i$ be the infinite vector with ith component N_i.*

1. *The probability that $\vec{N}_i=\vec{n_i}$ is $\frac{\prod_{r=1}^{\infty}\prod_{j=r}^{\infty}(1-\frac{1}{q^j})}{q^{\sum_i in_i}\prod_i(\frac{1}{q})_{n_i}}$.*
2. $$\sum_{\vec{n_i}:\sum n_i=a}\frac{1}{q^{\sum_i in_i}\prod_i(\frac{1}{q})_{n_i}}=\frac{J_a(q)}{q^{a^2}(1-\frac{1}{q})^2\cdots(1-\frac{1}{q^a})^2}.$$

Theorem 9. *The algorithm given for sampling from $M_{Pl,q}$ with $q>1$ conditioned to live on $|\lambda|=n$ is valid.*

3.4. **Markov Chain Approach.** The main result in this subsection is a third method for understanding the measure $M_{GL,u,q}$ probabilistically [F7]. The idea is to build up the random partition a column at a time; if the current column has size a, then the next column will have size b (with $b\le a$) with probability $K(a,b)$. The surprise is that this transition rule turns out to be independent of the columns, yielding a Markov chain on the natural numbers. This Markov chain is diagonalizable with eigenvalues $1,\frac{u}{q},\frac{u^2}{q^4},\cdots$. It will be used to give a probabilistic proof of the Rogers-Ramanujan identities in Subsection 3.5.

It is convenient to set λ'_0 (the height of an imaginary zeroth column) equal to ∞. For the entirety of this subsection, let $P(a)$ be the $M_{GL,u,q}$ probability that $\lambda'_1=a$. Theorem 10, which makes the connection with Markov chains, is proved in a completely elementary way. The argument re-proves that $M_{GL,u,q}$ is a probability measure (Lemma 4 of Subsection 3.1), shows that the asserted Markov transition probabilities add to one, and gives a formula for $P(a)$.

Theorem 10. *Starting with $\lambda'_0=\infty$, define in succession $\lambda'_1,\lambda'_2,\cdots$ according to the rule that if $\lambda'_i=a$, then $\lambda'_{i+1}=b$ with probability*

$$K(a,b)=\frac{u^b(\frac{1}{q})_a(\frac{u}{q})_a}{q^{b^2}(\frac{1}{q})_{a-b}(\frac{1}{q})_b(\frac{u}{q})_b}.$$

Then the resulting partition is distributed according to $M_{GL,u,q}$.

Proof. Suppose we know that $M_{GL,u,q}$ is a probability measure and that

$$P(a)=\frac{u^a(\frac{u}{q})_\infty}{q^{a^2}(\frac{1}{q})_a(\frac{u}{q})_a}.$$

Then the $M_{GL,u,q}$ probability of choosing a partition with $\lambda'_i = r'_i$ for all i is

$$Prob.(\lambda'_0 = \infty)\frac{Prob.(\lambda'_0 = \infty, \lambda'_1 = r_1)}{Prob.(\lambda'_0 = \infty)}$$
$$\times \prod_{i=1}^{\infty} \frac{Prob.(\lambda'_0 = \infty, \lambda'_1 = r_1, \cdots, \lambda'_{i+1} = r_{i+1})}{Prob.(\lambda'_0 = \infty, \lambda'_1 = r_1, \cdots, \lambda'_i = r_i)}.$$

Thus it is enough to prove that

$$\frac{Prob.(\lambda'_0 = \infty, \lambda'_1 = r_1, \cdots, \lambda'_{i-1} = r_{i-1}, \lambda'_i = a, \lambda'_{i+1} = b)}{Prob.(\lambda'_0 = \infty, \lambda'_1 = r_1, \cdots, \lambda'_{i-1} = r_{i-1}, \lambda'_i = a)} = \frac{u^b (\frac{1}{q})_a (\frac{u}{q})_a}{q^{b^2} (\frac{1}{q})_{a-b} (\frac{1}{q})_b (\frac{u}{q})_b},$$

for all $i, a, b, r_1, \cdots, r_{i-1}$. One calculates that

$$\sum_{\substack{\lambda : \lambda'_1 = r_1, \cdots, \lambda'_{i-1} = r_{i-1} \\ \lambda'_i = a}} M_{GL,u,q}(\lambda)$$
$$= \frac{u^{r_1 + \cdots + r_{i-1}}}{q^{r_1^2 + \cdots + r_{i-1}^2} (\frac{1}{q})_{r_1 - r_2} \cdots (\frac{1}{q})_{r_{i-2} - r_{i-1}} (\frac{1}{q})_{r_{i-1} - a}} P(a).$$

Similarly, observe that

$$\sum_{\substack{\lambda : \lambda'_1 = r_1, \cdots, \lambda'_{i-1} = r_{i-1} \\ \lambda'_i = a, \lambda'_{i+1} = b}} M_{GL,u,q}(\lambda)$$
$$= \frac{u^{r_1 + \cdots + r_{i-1} + a}}{q^{r_1^2 + \cdots + r_{i-1}^2 + a^2} (\frac{1}{q})_{r_1 - r_2} \cdots (\frac{1}{q})_{r_{i-2} - r_{i-1}} (\frac{1}{q})_{r_{i-1} - a} (\frac{1}{q})_{a-b}} P(b).$$

Thus the ratio of these two expressions is

$$\frac{u^b (\frac{1}{q})_a (\frac{u}{q})_a}{q^{b^2} (\frac{1}{q})_{a-b} (\frac{1}{q})_b (\frac{u}{q})_b},$$

as desired. Note that the transition probabilities must sum to 1 because

$$\sum_{b \leq a} \frac{\sum_{\substack{\lambda : \lambda'_1 = r_1, \cdots, \lambda'_{i-1} = r_{i-1} \\ \lambda'_i = a, \lambda'_{i+1} = b}} M_{GL,u,q}(\lambda)}{\sum_{\substack{\lambda : \lambda'_1 = r_1, \cdots, \lambda'_{i-1} = r_{i-1} \\ \lambda'_i = a}} M_{GL,u,q}(\lambda)} = 1$$

for any measure $M_{GL,u,q}$ on partitions.

Thus to complete the proof, it must be shown that $M_{GL,u,q}$ is a probability measure and that

$$P(a) = \frac{u^a (\frac{u}{q})_\infty}{q^{a^2} (\frac{1}{q})_a (\frac{u}{q})_a}.$$

Since

$$\frac{\sum_{\substack{\lambda : \lambda'_1 = r_1, \cdots, \lambda'_{i-1} = r_{i-1} \\ \lambda'_i = a, \lambda'_{i+1} = b}} M_{GL,u,q}(\lambda)}{\sum_{\substack{\lambda : \lambda'_1 = r_1, \cdots, \lambda'_{i-1} = r_{i-1} \\ \lambda'_i = a}} M_{GL,u,q}(\lambda)} = \frac{P(b) u^a}{P(a) q^{a^2} (\frac{1}{q})_{a-b}}$$

it follows that

$$\sum_{b \leq a} \frac{P(b) u^a}{P(a) q^{a^2} (\frac{1}{q})_{a-b}} = 1.$$

From this recursion and the fact that $P(0) = (\frac{u}{q})_\infty$, one solves for $P(a)$ inductively, finding that

$$P(a) = \frac{u^a(\frac{u}{q})_\infty}{q^{a^2}(\frac{1}{q})_a(\frac{u}{q})_a}.$$

(We remark that a more probabilistic understanding of this formula for $P(a)$ is available [F3].) Cauchy's identity (page 20 of [A1]) gives that $\sum_a P(a) = 1$, so that $M_{GL,u,q}$ is a probability measure. □

Theorem 11 diagonalizes the transition matrix K, finding a basis of eigenvectors, which is fundamental for understanding the Markov chain (part 3 is stated as a lemma in [A2]). Since the matrix K is upper triangular with distinct eigenvalues, this is straightforward.

Theorem 11. 1. *Let C be the diagonal matrix with (i,i) entry $(\frac{1}{q})_i(\frac{u}{q})_i$. Let M be the matrix $\left(\frac{u^j}{q^{j^2}(\frac{1}{q})_{i-j}}\right)$. Then $K = CMC^{-1}$, which reduces the problem of diagonalizing K to that of diagonalizing M.*

2. *Let A be the matrix $\left(\frac{1}{(\frac{1}{q})_{i-j}(\frac{u}{q})_{i+j}}\right)$. Then the columns of A are eigenvectors of M for right multiplication, the jth column having eigenvalue $\frac{u^j}{q^{j^2}}$.*

3. *The inverse matrix A^{-1} is $\left(\frac{(1-u/q^{2i})(-1)^{i-j}(\frac{u}{q})_{i+j-1}}{q^{\binom{i-j}{2}}(\frac{1}{q})_{i-j}}\right)$.*

Corollary 2 (immediate from Theorem 11) will be useful for the proof of the Rogers-Ramanujan identities in Section 3.5. In the case $L \to \infty$ and $j = 0$, it is the so-called Rogers-Selberg identity.

Corollary 2. *Let E be the diagonal matrix with (i,i) entry $\frac{u^i}{q^{i^2}}$. Then $K^r = CAE^rA^{-1}C^{-1}$. More explicitly,*

$$K^r(L,j) = \frac{(\frac{1}{q})_L(\frac{u}{q})_L}{(\frac{1}{q})_j(\frac{u}{q})_j}\sum_{n=0}^{\infty}\frac{u^{rn}(1-u/q^{2n})(-1)^{n-j}(\frac{u}{q})_{n+j-1}}{q^{rn^2}(\frac{1}{q})_{L-n}(\frac{u}{q})_{L+n}q^{\binom{n-j}{2}}(\frac{1}{q})_{n-j}}.$$

Proof. This is immediate from Theorem 11. □

Remarks. 1. One of our motivations for seeking a Markov chain description of $M_{GL,u,q}$ is work of Fristedt [Fris], who had a Markov chain approach for the measure P_q on the set of all partitions of all natural numbers defined by $P_q(\lambda) = \prod_{i=1}^{\infty}(1-q^i)q^{|\lambda|}$ where $q < 1$. Fristedt's interest was in studying what a uniformly chosen partition of an integer looks like, and conditioning P_q to live on partitions of size n gives a uniform partition. The measure P_q is related to the vertex operators [O1] and to the enumeration of ramified coverings of the torus [Dij]. In this regard the papers [O1] and [BlO] prove that the k point correlation function

$$F(t_1,\cdots,t_k) = \sum_\lambda q^{|\lambda|}\prod_{k=1}^{n}\sum_{i=1}^{\infty}t_k^{\lambda_i - i + \frac{1}{2}}$$

is a sum of determinants involving genus 1 theta functions and their derivatives and give connections with quasi-modular forms. It would be marvellous if the measure $M_{GL,u,q}$ (being related to modular forms via the Rogers-Ramanujan identities) is also related to enumerative questions in algebraic geometry.

2. As mentioned in the introduction, the Markov chain approach gives a unified description of conjugacy classes of the finite classical groups. For the symplectic and orthogonal groups it is necessary to use two Markov chains, K_1 and K_2. For the symplectic case, steps with column number i odd use K_1 and steps with column number i even use K_2. For the orthogonal case, steps with column number i odd use K_2 and steps with column number i even use K_1. The Markov chains K_1, K_2 are the same for both cases! (This construction is reminiscent of transfer matrices in statistical mechanics.) Details are in [F6]. The Markov chain approach is also related to quivers [F7].

3.5. **Rogers-Ramanujan Identities.** The Rogers-Ramanujan identities [Ro]

$$1 + \sum_{n=1}^{\infty} \frac{q^{n^2}}{(1-q)(1-q^2)\cdots(1-q^n)} = \prod_{n=1}^{\infty} \frac{1}{(1-q^{5n-1})(1-q^{5n-4})}$$

$$1 + \sum_{n=1}^{\infty} \frac{q^{n(n+1)}}{(1-q)(1-q^2)\cdots(1-q^n)} = \prod_{n=1}^{\infty} \frac{1}{(1-q^{5n-2})(1-q^{5n-3})}$$

are among the most interesting partition identities in number theory and combinatorics, with connections to Lie theory and statistical mechanics (see the discussions in [A2] and [F7] for many references). In terms of partitions, they are often stated as

1. The partitions of an integer n in which the difference between any two parts is at least 2 are equinumerous with the partitions of n into parts congruent to 1 or 4 mod 5.
2. The partitions of an integer n in which each part exceeds 1 and the difference between any two parts is at least 2 are equinumerous with the partitions of n into parts congruent to 2 or 3 mod 5.

One ongoing challenge in the subject (posed by Hardy) has been to find a proof of the Rogers-Ramanujan identities which is both motivated and simple. The purpose of this subsection is to describe such a proof ([F7]), which is also the first probabilistic proof of the Rogers-Ramanujan identities.

To illustrate the idea we give the proof of the following generalization of the first Rogers-Ramanujan identity (called the Andrews-Gordon identity [A3],[Gor]):

$$\sum_{n_1,\cdots,n_{k-1} \geq 0} \frac{1}{q^{N_1^2+\cdots+N_{k-1}^2}(1/q)_{n_1}\cdots(1/q)_{n_{k-1}}} = \prod_{\substack{r=1 \\ r \neq 0, \pm k (mod\ 2k+1)}}^{\infty} \frac{1}{1-(1/q)^r}$$

where $N_i = n_i + \cdots + n_{k-1}$. Setting $k = 2$ and replacing q by its reciprocal specialize to the first Rogers-Ramanujan identity.

The idea is simple. We study the distribution of the length of the first row of a random partition distributed as $M_{GL,1,q}$. From the definition of $M_{GL,1,q}$ the probability that the first row has length less than k is equal to

$$\prod_{r=1}^{\infty}(1-\frac{1}{q^r}) \sum_{\lambda : \lambda'_k = 0} \frac{1}{q^{(\lambda'_1)^2+\cdots+(\lambda'_{k-1})^2}(1/q)_{\lambda'_1-\lambda'_2}\cdots(1/q)_{\lambda'_{k-1}-\lambda'_k}}.$$

Letting n_i denote $\lambda'_i - \lambda'_{i+1}$ and N_i denote λ'_i, this becomes

$$\prod_{r=1}^{\infty}(1-\frac{1}{q^r}) \sum_{n_1,\cdots,n_{k-1}\geq 0} \frac{1}{q^{N_1^2+\cdots+N_{k-1}^2}(1/q)_{n_1}\cdots(1/q)_{n_{k-1}}},$$

which is essentially the left hand side of the Andrews-Gordon identity. On the other hand the probability that the first row has length less than k is equal to the probability that the Markov chain of Section 3.4 is 0 at time k. Since we diagonalized the matrix associated to this Markov chain, it is straightforward to compute this probability. To get it into product form it is necessary to apply Jacobi's triple product identity, which has a simple combinatorial proof [A1]. Further details are in [F7].

Next we argue that this proof is motivated. Certainly the measure $M_{GL,1,q}$ is a natural object to study, given that it is the $n \to \infty$ limit law of λ_{z-1} for a random element of $GL(n,q)$. It was natural to try to build up the random partitions λ column by column as in Section 3.4. Observing that the resulting Markov chain is absorbing at 0 with probability one, the time to absorption (equivalent to the distribution of the length of the first row) is the most natural quantity one could examine. The final step is applying Jacobi's triple product identity, and thus going from a "sum = sum" identity to a "sum = product" identity. As mentioned above Jacobi's triple product identity is easy to verify, but one still wants a motivation for trying to write the left hand side of the Andrews-Gordon identity in product form. The best motivation is Baxter's work on statistical mechanics (surveyed in [A2],[Bax1],[Bax2]) in which he really needed "sum = product" identities and was led to conjecture analogs of Rogers-Ramanujan type identities. Although a proof of the Rogers-Ramanujan identities doesn't emerge from his work, it is clearly one of the truly great accomplishments in mathematical physics, and his book [Bax1] has been very influential. A second motivation is our work on the $n \to \infty$ asymptotic probability that an element of $GL(n,q)$ is semisimple. The argument, recorded in [F1] or the more readily available [F4], needed a "sum = product" identity. The corresponding computation in [F9] for the finite affine groups needed both Rogers-Ramanujan identities.

Andrews' paper [A4] notes that many proofs of the Rogers-Ramanujan identities make use of the following mysterious result called Bailey's Lemma, alluded to in [Bai] and stated explicitly in [A3]. A pair of sequences $\{\alpha_L\}$ and $\{\beta_L\}$ is called a Bailey pair if

$$\beta_L = \sum_{r=0}^{L} \frac{\alpha_r}{(1/q)_{L-r}(u/q)_{L+r}}.$$

Bailey's Lemma states that if $\alpha'_L = \frac{u^L}{q^{L^2}}\alpha_L$ and $\beta'_L = \sum_{r=0}^{L} \frac{u^r}{q^{r^2}(1/q)_{L-r}}\beta_r$, then $\{\alpha'_L\}$ and $\{\beta'_L\}$ form a Bailey pair. From the viewpoint of Markov chains, this case of Bailey's Lemma is clear. To explain, let A, D, M be as in Theorem 11 (recall that $M = ADA^{-1}$). Viewing $\alpha = \vec{\alpha_L}$ and $\beta = \vec{\beta_L}$ as column vectors, the notion of a Bailey pair means that $\beta = A\alpha$. This case of Bailey's Lemma follows because

$$\beta' = M\beta = ADA^{-1}\beta = AD\alpha = A\alpha'.$$

As Andrews explains in [A2], the power of Bailey's Lemma lies in its ability to be iterated. This gives a short but unmotivated proof of the Rogers-Selberg identity (Corollary 2 in Section 3.4). Partition theorists refer to this iteration of Bailey's

Lemma as a Bailey chain. From the remarks of the preceding paragraph it is clear that the Bailey chain under consideration is really the Markov chain K, stripped of its probabilistic origin. The fact that the Markov chain approach has analogs for other finite classical groups and for quivers is further evidence of its naturality.

4. Upper triangular matrices

This section surveys probabilistic aspects of conjugacy classes in the group $T(n,q)$ of upper triangular matrices over finite fields with 1's along the main diagonal. At present little is known about conjugacy in $T(n,q)$. For instance the number of conjugacy classes, their size, or even a natural indexing are provably elusive (partial results on numbers of conjugacy classes can be found in [VAr],[VArV]). It seems as if the best description is "it's a mess". This makes a probabilistic description natural.

Kirillov's survey [Kir] calls for an extension of his method of coadjoint orbits for groups over real, complex, or p-adic fields to the group $T(n,q)$ and gives preliminary connections with statistical physics; the paper [IsKar] gives a counterexample to one of his conjectures. As we do not see how to further develop those results or improve on their exposition, we instead focus on a simpler problem: the probabilistic study of Jordan form of elements of $T(n,q)$.

Subsection 4.1 describes a probabilistic growth algorithm for the Jordan form of upper triangular matrices over a finite field. This is linked with symmetric function theory and potential theory on Bratteli diagrams in Subsection 4.2.

4.1. Growth Algorithm for Jordan Form.

Theorem 12 gives a probabilistic growth algorithm for the Jordan form of random elements of $T(n,q)$. Its proof uses elementary reasoning from linear algebra.

Theorem 12. ([Kir],[B]) *The Jordan form of a uniformly chosen element of $T(n,q)$ can be sampled by stopping the following procedure after n steps: Starting with the empty partition, at each step transition from a partition λ to a partition Λ by adding a box to column i chosen according to the rules*

- $i = 1$ *with probability* $\frac{1}{q^{\lambda'_1}}$,
- $i = j > 1$ *with probability* $\frac{1}{q^{\lambda'_j}} - \frac{1}{q^{\lambda'_{j-1}}}$.

Theorem 12 leads to the following central limit theorem about the asymptotic Jordan form of an element of $T(n,q)$.

Theorem 13. ([B]) *Let λ be the partition corresponding to the Jordan form of a random element of $T(n,q)$. Let $Prob^n$ denote probability under the uniform measure on $T(n,q)$ and let $p_i = \frac{1}{q^{i-1}} - \frac{1}{q^i}$. Then*

$$lim_{n\to\infty} Prob^n(\frac{\lambda_i - p_i n}{\sqrt{n}} \leq x_i, i = 1, \cdots, k) = (2\pi)^{-\frac{k}{2}} \int_{-\infty}^{x_1} \cdots \int_{\infty}^{x_k} e^{-\frac{1}{2} < Qt,t >} dt$$

for any $(x_1, \cdots, x_k) \in R^k$, where the covariance matrix equals

$$Q = diag(p_1, \cdots, p_k) - (p_i p_j)_{i,j=1}^k.$$

4.2. **Symmetric Functions and Potential Theory.** Given the usefulness of symmetric functions in the probabilistic study of the measure $M_{GL,u,q}$, it is natural to seek an analogous understanding of Theorem 12. That is the topic of the present subsection. The ideas here are from the report [F8].

The first step is to link the probability that an element of $T(n,q)$ has Jordan form of type Λ with symmetric function theory. For the rest of this section, $P_\Lambda(q,t)$ denotes a Macdonald polynomial, $K_{\mu\Lambda}(q,t)$ denotes a Kostka-Foulkes polynomial, and f^μ is the dimension of the irreducible representation of S_n corresponding to the partition μ (see [Mac] for background). Note that when $q=0$ the Macdonald polynomial is our friend, a Hall-Littlewood polynomial.

Theorem 14. ([F5]) *The probability that a random element of $T(n,q)$ has Jordan form of type Λ is*

$$P_\Lambda(1-\frac{1}{q},\frac{1}{q}-\frac{1}{q^2},\cdots;0,\frac{1}{q})\sum_{\mu\vdash n} f^\mu K_{\mu\Lambda}(0,1/q).$$

Next we give some background on potential theory on Bratteli diagrams. This is a beautiful subject, with connections to probability and representation theory. We recommend [Ke1] for background on potential theory with many examples and [BOl] for a survey of recent developments. The basic set-up is as follows. One starts with a Bratteli diagram, that is an oriented graded graph $\Gamma = \cup_{n\geq 0}\Gamma_n$ such that

1. Γ_0 is a single vertex $\emptyset$.
2. If the starting vertex of an edge is in Γ_i, then its end vertex is in Γ_{i+1}.
3. Every vertex has at least one outgoing edge.
4. All Γ_i are finite.

For two vertices $\lambda, \Lambda \in \Gamma$, one writes $\lambda \nearrow \Lambda$ if there is an edge from λ to Λ. Part of the underlying data is a multiplicity function $\kappa(\lambda,\Lambda)$. Letting the weight of a path in Γ be the product of the multiplicities of its edges, one defines the dimension $dim(\Lambda)$ of a vertex Λ to be the sum of the weights over all maximal length paths from $\emptyset$ to Λ (this definition clearly extends to intervals). Given a Bratteli diagram with a multiplicity function, one calls a function ϕ *harmonic* if $\phi(0)=1$, $\phi(\lambda)\geq 0$ for all $\lambda\in\Gamma$, and

$$\phi(\lambda) = \sum_{\Lambda:\lambda\nearrow\Lambda} \kappa(\lambda,\Lambda)\phi(\Lambda).$$

An equivalent concept is that of coherent probability distributions. Namely a set $\{M_n\}$ of probability distributions M_n on Γ_n is called *coherent* if

$$M_{n-1}(\lambda) = \sum_{\Lambda:\lambda\nearrow\Lambda} \frac{dim(\lambda)\kappa(\lambda,\Lambda)}{dim(\Lambda)} M_n(\Lambda).$$

The formula allowing one to move between the definitions is $\phi(\lambda) = \frac{M_n(\lambda)}{dim(\lambda)}$.

One reason the set-up is interesting from the viewpoint of probability theory is the fact that every harmonic function can be written as a Poisson integral over the set of extreme harmonic functions. For the Pascal lattice (vertices of Γ_n are pairs (k,n) with $k=0,1,\cdots,n$ and (k,n) is connected to $(k,n+1)$ and $(k+1,n+1)$), this fact is the simplest instance of de Finetti's theorem, which says that an infinite exchangeable sequence of $0-1$ random variables is a mixture of coin toss sequences for different probabilities of heads. When the multiplicity function κ is integer valued, one can define a sequence of algebras A_n associated to the Bratteli diagram,

and harmonic functions correspond to certain characters of the inductive limit of the algebras A_n.

Next we define a branching for which the probability that an element of $T(n,q)$ has Jordan type Λ is a harmonic function. First some notation is needed. For $\lambda \nearrow \Lambda$, let $R_{\Lambda/\lambda}$ (resp. $C_{\Lambda/\lambda}$) be the boxes of λ in the same row (resp. column) as the boxes removed from Λ to get λ. This notation differs from that in [Mac]. Let $a_\lambda(s)$, $l_\lambda(s)$ be the number of boxes in λ strictly to the east and south of s, and let $h_\lambda(s) = a_\lambda(s) + l_\lambda(s) + 1$.

Definition 1. For $0 \leq q < 1$ and $0 < t < 1$, the underlying Bratteli diagram Γ has as level Γ_n all partitions λ of n. Letting i be the column number of the box removed to go from λ to Λ, for $\lambda \nearrow \Lambda$, define the multiplicity function as

$$\kappa(\lambda, \Lambda) = \frac{1}{t^{\Lambda'_i - 1}} \prod_{s \in R_{\Lambda/\lambda}} \frac{1 - q^{a_\Lambda(s)+1} t^{l_\Lambda(s)}}{1 - q^{a_\lambda(s)+1} t^{l_\lambda(s)}} \prod_{s \in C_{\Lambda/\lambda}} \frac{1 - q^{a_\Lambda(s)} t^{l_\Lambda(s)+1}}{1 - q^{a_\lambda(s)} t^{l_\lambda(s)+1}}.$$

Equation I.10 of [GarsH] proves that

$$dim(\Lambda) = \frac{1}{t^{n(\Lambda)}} \sum_{\mu \vdash n} f^\mu K_{\mu\Lambda}(q,t).$$

Definition 2. For $0 \leq q < 1, 0 < t < 1$ and $0 \leq x_1, x_2, \cdots$ such that $\sum x_i = 1$, define a family $\{M_n\}$ of probability measures on partitions of size n by

$$\begin{aligned} M_n(\Lambda) &= \frac{(1-q)^{|\Lambda|} P_\Lambda(x;q,t) \sum_{\mu \vdash n} f^\mu K_{\mu\Lambda}(q,t)}{\prod_{s \in \Lambda} (1 - q^{a_\Lambda(s)+1} t^{l_\Lambda(s)})} \\ &= \frac{(1-q)^{|\Lambda|} P_\Lambda(x;q,t) t^{n(\Lambda)} dim(\Lambda)}{\prod_{s \in \Lambda} (1 - q^{a_\Lambda(s)+1} t^{l_\Lambda(s)})}. \end{aligned}$$

Consider the specialization that $q = 0$ and $t = \frac{1}{q}$, where this second q is the size of a finite field. Further, set $x_i = \frac{1}{q^{i-1}} - \frac{1}{q^i}$. Then Theorem 14 implies that $M_n(\Lambda)$ is the probability that a uniformly chosen element of $T(n,q)$ has Jordan type Λ. The multiplicities have a simple description; letting i be the column to which one adds in order to go from λ to Λ, it follows that $\kappa(\lambda, \Lambda) = q^{\lambda'_i} + q^{\lambda'_i - 1} + \cdots + q^{\lambda'_{i+1}}$. Second, $dim(\Lambda)$ reduces to a Green's polynomial $Q^\Lambda(q) = Q^\Lambda_{(1^n)}(q)$ as in Section 3.7 of [Mac]. These polynomials are important in the representation theory of the finite general linear groups. This specialization was the motivation for Definition 2.

The connection with potential theory is given by the following result.

Theorem 15 ([F8])**.** *The measures of Definition 2 are harmonic with respect to the branching of Definition 1.*

It is elementary and well known that if one starts at the empty partition and transitions from λ to Λ with probability $\frac{\kappa(\lambda,\Lambda) M_n(\Lambda) dim(\lambda)}{M_{n-1}(\lambda) dim(\Lambda)}$, one gets samples from any coherent family of measures $\{M_n\}$. Applying this principle to the above specialization in which $M_n(\Lambda)$ is $T(n,q)$ and using Macdonald's principal specialization formula (page 337 of [Mac]) give the advertised proof of Theorem 12 by means of symmetric functions and potential theory.

Remarks. 1. The example of Schur functions ($q = t < 1$) is also interesting. The measure $M_n(\Lambda)$ reduces to $s_\Lambda f^\Lambda$, where s_Λ is a Schur function. Setting

$x_1 = \cdots = x_n = \frac{1}{n}$ and letting $n \to \infty$, one obtains Plancherel measure, which is important in representation theory and random matrix theory. Letting $x_1 = \cdots = x_n$ satisfy $\sum x_i = 1$ (all other $x_j = 0$) gives a natural deformation of Plancherel measure, studied for instance by [ItTWi]. Stanley [Sta] shows that this measure on partitions also arises by applying the Robinson-Schensted-Knuth algorithm to a random permutation distributed after a biased riffle shuffle (in other words, this measure encodes information about the longest increasing subsequences of permutations distributed as shuffles).

2. It has been pointed out to the author that the branchings $\kappa(\lambda, \Lambda)$ of Definition 1 are related to the branchings $\tau(\lambda, \Lambda)$ of [Ke2] by the formula

$$\kappa(\lambda, \Lambda) = f(\lambda)\tau(\lambda, \Lambda)f(\Lambda)^{-1},$$

for a certain positive function $f(\lambda)$ on the set of vertices, which implies by [Ke3] that the boundaries of these two branchings are homeomorphic and that the branchings of Definition 1 are multiplicative. Kerov [Ke2] has a conjectural description of the boundary. It has been verified for Schur functions [T], Kingman branching [Kin], and Jack polynomials [KeOOl], but remains open for the general case of Macdonald polynomials. In particular, it is open for Hall-Littlewood polynomials, the case related to $T(n, q)$. It is interesting that the $\kappa(\lambda, \Lambda)$ of Definition 1 are integers for Hall-Littlewood polynomials, whereas the $\tau(\lambda, \Lambda)$ of [Ke2] are not.

Acknowledgements

The author's greatest thanks go to Persi Diaconis (his former thesis advisor) for years of friendship, encouragement, and inspiration. He was very helpful in the preparation of this article. We thank Peter M. Neumann and Cheryl E. Praeger for countless conversations about conjugacy classes and computational group theory, Roger Carter for correspondence, and Mark Huber for permission to survey some joint unpublished results. The author received the financial support of an NSF Postdoctoral Fellowship and wrote this article at Stanford University.

References

[AlDia] Aldous, D. and Diaconis, P., Longest increasing subsequences: from patience sorting to the Baik-Deift-Johansson theorem, *Bull. AMS (N.S.)* **36** (1999), 413-432. MR **2000g:**60013

[A1] Andrews, G., *The theory of partitions. Encyclopedia of Mathematics and its Applications, Vol. 2.* Addison-Wesley Publishing Co., Reading, Mass.-London-Amsterdam, 1976. MR **58:**27738

[A2] Andrews, G., *q-Series: Their development and application in analysis, number theory, combinatorics, physics, and computer algebra.* American Math Society, 1986. MR **88b:**11063

[A3] Andrews, G., *Multiple series Rogers-Ramanujan type identities*, *Pacific J. Math.* **114** (1984), 267-283. MR **86c:**11084

[A4] Andrews, G., On the proofs of the Rogers-Ramanujan identities, in *q-series and partitions*, IMA Vol. Math. Appl. 18, 1989. MR **91e:**11112

[ArBarT] Arratia, R., Barbour, A.D., and Tavare, S., On random polynomials over finite fields, *Math. Proc. Cambridge Phil. Soc.* **114** (1993), 347-368. MR **95a:**60011

[As] Aschbacher, M., On the maximal subgroups of the finite classical groups, *Invent. Math.* **76** (1984), 469-514. MR **86a:**20054

[Bai] Bailey, W.N., Identities of the Rogers-Ramanujan type, *Proc. London Math. Soc.* (2) **50** (1948), 1-10. MR **9:**585b

[Bax1] Baxter, R., Exactly solved models in statistical mechanics, Academic Press, London/New York, 1982. MR **86i:**82002a

[Bax2] Baxter, R., Ramanujan's identities in statistical mechanics, in *Ramanujan revisited* (1988), 69-84. MR **89h:**82043

[BlO] Bloch, S. and Okounkov, A., *The character of the infinite wedge representation*, Adv. Math. **149** (2000), 1–60. MR **2001g:**11059

[BKW] Blömer, J., Karp, R., and Welzl, E., The rank of sparse random matrices over finite fields, *Random Structures Algorithms* **10** (1997) 407-419. MR **99b:**15028

[Bo] Bollobas, B., Random graphs, Academic Press, London, 1985. MR **87f:**05152

[B] Borodin, A., The law of large numbers and the central limit theorem for the Jordan normal form of large triangular matrices over a finite field, Zap. Nauchn. Sem. LOMI, Vol. 240, 1997, pg. 18-43 (Russian); English translation in *J. Math. Sci. (New York)* **96** (1997), 3455-3471. MR **2000f:**60044

[BOOl] Borodin, A., Okounkov, A., and Olshanki, G., Asymptotics of Plancherel measures for symmetric groups, *J. Amer. Math. Soc.* **13** (2000), 481-515. MR **2001g:**05103

[BOl] Borodin, A. and Ohlshanski, G., Harmonic functions on multiplicitive graphs and interpolation polynomials, *Electron. J. Combin.* **7** (2000), 39 pp. MR **2001f:**05160

[CeLg] Celler, F. and Leedham-Green, C.R., A constructive recognition algorithm for the special linear group, in *The atlas of finite groups: ten years on (Birmingham, 1995)*, London Math. Soc. Lecture Note Ser., 249, Cambridge Univ. Press (1998), 11-26. MR **99g:**20001

[CeLgMuNiOb] Celler, F., Leedham-Green, C.R., Murray, S.H., Niemeyer, A.C., and O'Brien, E.A., Generating random elements of a finite group, *Commun. in Algebra* **23** (1995), 4931-4948. MR **96h:**20115

[ChReRo] Charlap, L., Rees, H., and Robbins, D., The asymptotic probability that a random biased matrix is invertible, *Discrete Math.* **82** (1990), 153-163. MR **91g:**05009

[CTY] Chigira, N., Takegahara, Y., and Yoshida, T., On the number of homomorphisms from a finite group to a general linear group, *J. Algebra* **232** (2000), 236-254. MR **2001h:**20069

[DiaGr] Diaconis, P. and Graham, R., An affine walk on the hypercube, *J. Comput. Appl. Math.* **41** (1992), 215-235.

[DiaMcPi] Diaconis, P., McGrath, M., and Pitman, J., Riffle shuffle, cycles, and descents, *Combinatorica* **15** (1995), 11-29. MR **96g:**05009

[Dij] Dijkgraaf, R., Mirror symmetry and elliptic curves, in *The Moduli Space of Curves*, Prog. in Math. **129** (1995). MR **96m:**14072

[Ed] Edelman, A., *Eigenvalues and condition numbers of random matrices*, Ph.D. Thesis, MIT, 1989.

[El] Elkies, N., On finite sequences satisfying linear recursions. Available at http://xxx.lanl.gov/abs/math.CO/0105007.

[ErT] Erdös, P. and Turan, P., On some problems of statistical group theory. I, *Zeit. Wahr. Verw. Gabiete* **4** (1965), 175-186. MR **32:**2465

[Ew] Ewens, W.J., The sampling theory of selectively neutral alleles, *Theoret. Population Biology* **3** (1972), 87-112. MR **48:**3526

[FeigFre] Feigin, B. and Frenkel, E., Coinvariants of nilpotent subalgebras of the Virasoro algebra and partition identities. I. M. Gelfand Seminar, *Adv. Soviet Math.* **16**, Part 1, 139-148, Amer. Math. Soc., Providence, RI, 1993. MR **94g:**17054

[FineHer] Fine, N.J. and Herstein, I. N., The probability that a matrix be nilpotent, *Illinois J. Math.* **2** (1958), 499-504. MR **20:**3160

[FinkKa] Finkelstein, L. and Kantor, W. (editors), *Groups and computation II*, DIMACS Series 28, Amer. Math. Soc., Providence, RI, 1997. MR **97m:**20003

[Fey] Feynman, R., Statistical mechanics. A set of lectures. Reprint of the 1972 original. Perseus Books, Reading, MA, 1998. MR **99i:**82001

[FlJ] Fleischmann, P. and Janiszczak, I., The number of regular semisimple elements for Chevalley groups of classical type. *J. Algebra* **155** (1993), 482-528. MR **94f:**20090

[Frip1] Fripertinger, H., Cycle indices of linear, affine, and projective groups, *Lin. Alg. Appl.* **263** (1997), 133-156. MR **98h:**05179

[Frip2] Fripertinger, H., Random generation of linear codes, *Aequationes Math.* **58** (1999), 192-202. MR **2001d:**94028

[Fris] Fristedt, B., The structure of random partitions of large integers, *Trans. Amer. Math. Soc.* **337** (1993), 703-735. MR **93h:**11090

[F1] Fulman, J., *Probability in the classical groups over finite fields: symmetric functions, stochastic algorithms and cycle indices*, Ph.D. Thesis, Harvard University, 1997.

[F2] Fulman, J., Cycle indices for the finite classical groups. *J. Group Theory* **2** (1999), 251-289. MR **2001d:**20045

[F3] Fulman, J., A probabilistic approach to conjugacy classes in the finite general linear and unitary groups, *J. Algebra* **212** (1999), 557-590. MR **2000c:**20072

[F4] Fulman, J., The Rogers-Ramanujan identities, the finite general linear groups, and the Hall-Littlewood polynomials, *Proc. Amer. Math. Soc.* **128** (2000), 17-25. MR **2000h:**05229

[F5] Fulman, J., The eigenvalue distribution of a random unipotent matrix in its representation on lines, *J. Algebra* **228** (2000), 497-511. MR **2001d:**20072

[F6] Fulman, J., A probabilistic approach to conjugacy classes in the finite symplectic and orthogonal groups, *J. Algebra* **234** (2000), 207-224. CMP 2001:05

[F7] Fulman, J., A probabilistic proof of the Rogers-Ramanujan identities. *Bull. London Math. Soc.* **33** (2001), 397-407.

[F8] Fulman, J., New examples of potential theory on Bratteli diagrams. Available at http://xxx.lanl.gov/abs/math.CO/9912148.

[F9] Fulman, J., Finite affine groups: cycle indices, symmetric functions, and probabilistic algorithms. To appear in *J. Algebra*.

[F10] Fulman, J., Applications of the Brauer complex: card-shuffling, permutation statistics, and dynamical systems. To appear in *J. Algebra*.

[FNP] Fulman, J., Neumann, P.M., and Praeger, C.E., A generating function approach to the enumeration of cyclic and separable matrices in the finite classical groups. Preprint.

[GarsH] Garsia, A. and Haiman, M., A random q,t-hook walk and a sum of Pieri coefficients, *J. Combin. Theory Ser. A* **82** (1998), 74-111. MR **2000b:**05133

[GoSchm] Goh, W. and Schmutz, E., A central limit theorem on $GL(n,q)$, *Random Struc. Alg.* **2** (1991), 47-53. MR **92f:**11176

[Gon] Goncharov, V., Du domaine d'analyse combinatoire, *Bull. Acad. Sci. URSS Ser. Math* **8** (1944), 3-48, translated in *Amer. Math. Soc. Transl.* **19** (1950).

[Gor] Gordon, B., A combinatorial generalization of the Rogers-Ramanujan identities, *Amer. J. Math.* **83** (1961), 393-99. MR **23:**A809

[GuLub] Guralnick, R. and Lübeck, F., The proportion of p-singular elements in simple groups of Lie type, to appear in *Groups and Computation III*.

[HSchm] Hansen, J. and Schmutz, E., How random is the characteristic polynomial of a random matrix?, *Math. Proc. Cambridge Philos. Soc.* **114** (1993), 507-515. MR **94j:**05009

[Her] Herstein, I.N., Topics in algebra. Second edition. Xerox College Publishing, Lexington, Mass.-Toronto, Ont., 1975. MR **50:**9456

[IsKanSp] Isaacs, I.M., Kantor, W.M., and Spaltenstein, N., On the probability that a group element is p-singular, *J. Algebra* **176** (1995), 139-181. MR **96f:**20035

[IsKar] Isaacs, I. and Karagueuzian, D., Conjugacy in groups of upper triangular matrices, *J. Algebra* **202** (1998), 704-711. MR **99b:**20011

[ItTWi] Its, A.R., Tracy, C.A., and Widom, H., Random words, Toeplitz determinants and integrable systems, I. Preprint math.CO/9909169 at xxx.lanl.gov.

[Jo] Johansson, K., Discrete orthogonal polynomial ensembles and the Plancherel measure, Ann. of Math. (2) **153** (2001), 259–296. CMP 2001:11

[KeaSn] Keating, J.P. and Snaith, N.C., Random matrix theory and $\zeta(1/2+it)$. *Comm. Math. Phys.* **214** (2000), 57-89.

[Ke1] Kerov, S.V., The boundary of Young lattice and random Young tableaux, *Formal power series and algebraic combinatorics*, DIMACS Ser. Discrete Math. Theoret. Comput. Sci. **24**, Amer. Math. Soc., Providence, RI, (1996), 133-158. MR **96i:**05177

[Ke2] Kerov, S.V., Generalized Hall-Littlewood symmetric functions and orthogonal polynomials, *Adv. Sov. Math.* **9** (1992), 67-94. MR **94b:**05208

[Ke3] Kerov, S.V., Combinatorial examples in AF-algebras, Differential geometry, Lie groups and mechanics X, Zap. Nauchn. Sem. LOMI, Vol. 172, 1989, pp. 55-67 (Russian); English translation in *J. Soviet Math.* **59** (1992), 1063-1071. MR **90j:**46062

[KeOOl] Kerov, S.V., Okounkov, A., and Olshanksi, G., The boundary of Young graph with Jack edge multiplicities, *Intern. Math. Res. Notices* **4** (1998), 173-199. MR **99f:**05120

[Kin] Kingman, J.F.C., Random partitions in population genetics, *Proc. R. Soc. Lond. A* **361** (1978), 1-20. MR **58:**26167

[Kir] Kirillov, A.A., Variations on the triangular theme, *Amer. Math. Soc. Transl.* **169** (1995), 43-73. MR **97a:**20072

[Kn] Knuth, D., The art of computer programming. Vol. 2. Seminumerical algorithms. Third edition. Addison-Wesley Publishing. Reading, Mass., 1997.

[Ko] Kolchin, V., Random mappings. Optimization Software, Inc., New York, 1986. MR **88a:**60022

[Kun] Kung, J., The cycle structure of a linear transformation over a finite field, *Lin. Alg. Appl.* **36** (1981), 141-155. MR **82d:**15012

[Leh] Lehrer, G., The cohomology of the regular semisimple variety, *J. Algebra* **199** (1998), 666-689. MR **98k:**20080

[LehSe] Lehrer, G. and Segal, G.P., Homology stability for classical regular semisimple varieties, *Math. Z.* **236** (2001), 251–290. CMP 2001:09

[LiSh] Liebeck, M.W. and Shalev, A., The probability of generating a finite simple group, *Geom. Dedicata* **56** (1995), 103-113. MR **96h:**20116

[LiSh2] Liebeck, M.W. and Shalev, A., Simple groups, permutation groups, and probability. *J. of AMS* **12** (1999), 497-520. MR **99h:**20004

[LubPa] Lubotzky, A. and Pak, I., The product replacement algorithm and Kazhdan's property (T), *J. of AMS* **52** (2000), 5525-5561; *J. of AMS* **14** (2001), 347–363 (electronic).

[Mac] Macdonald, I.G., Symmetric functions and Hall polynomials, Second Edition. Clarendon Press, Oxford. 1995. MR **96h:**05207

[MaSl] MacWilliams, F. and Sloane, N., The theory of error-correcting codes, Third printing, North-Holland Publishing Co., Amsterdam-New York-Oxford, 1977. MR **57:**5408a; MR **57:**5408b

[Mar] Marsaglia, G., A current view of random number generators. Keynote Address, Sixteenth Symposium on the Interface between Computer Science and Statistics, Elsevier Press, 1984.

[MarTs] Marsaglia, G. and Tsay, L.H., Matrices and the structure of random number sequences, *Lin. Alg. Appl.* **67** (1985), 147-156. MR **86g:**65018

[Mu] Murray, S., *Conjugacy classes in maximal parabolic subgroups of the general linear group*, Ph.D. Thesis, University of Chicago, 1999.

[NP1] Neumann, P.M. and Praeger, C.E., A recognition algorithm for special linear groups, *Proc. London Math. Soc. (3)* **65** (1992), 555-603. MR **93m:**20063

[NP2] Neumann, P.M. and Praeger, C.E., Cyclic matrices over finite fields, *J. London Math. Soc. (2)* **52** (1995), 263-284. MR **96j:**15017

[NP3] Neumann, P.M. and Praeger, C.E., Derangements and eigenvalue-free elements in finite classical groups, *J. London Math. Soc. (2)* **58** (1998), 562-586. MR **2000a:**20153

[NP4] Neumann, P.M. and Praeger, C.E., Cyclic matrices and the meataxe, *Groups and Computation III*, de Gruyter, Berlin, 2001. CMP 2001:12

[NP5] Neumann, P.M. and Praeger, C.E., Cyclic matrices in classical groups over finite fields, *J. Algebra* **234** (2000), 367–418. CMP 2001:06

[NiP] Niemeyer, A. and Praeger, C.E., A recognition algorithm for classical groups over finite fields, *Proc. London Math. Soc.* **77** (1998), 117-169. MR **99k:**20002

[O1] Okounkov, A., Infinite wedge and measures on partitions. Available at http://xxx.lanl.gov/abs/math.RT/9907127.

[Pa] Pak, I., What do we know about the product replacement algorithm?, in *Groups and Computation III*, deGruyter, Berlin, 2001. CMP 2001:12

[PoRe] Polya, G. and Read, R.C., Combinatorial enumeration of groups, graphs, and chemical compounds. Springer-Verlag. New York, 1987. MR **89f:**05013

[Py1] Pyber, L., Asymptotic results for permutation groups, in *Groups and computation*, DIMACS Ser. 11, Amer. Math. Soc., Providence, RI (1993). MR **94g:**20003

[Py2] Pyber, L., Asymptotic results for simple groups and some applications, in *Groups and Computation II*, DIMACS Ser. 28, Amer. Math. Soc., Providence, RI (1997). MR **98a:**20030

[Py3] Pyber, L., Group enumeration and where it leads us, in *European Congress of Mathematics, Vol. II (Budapest 1996)*, Birkhäuser, 1998. MR **99i:**20037

[Ro] Rogers, L.J., Second memoir on the expansion of certain infinite products, *Proc. London Math. Soc.* **25** (1894), 318-343.

[RuShi] Rudvalis, A. and Shinoda, K., An enumeration in finite classical groups. Preprint (1988).

[Schm] Schmutz, E., The order of a typical matrix with entries in a finite field, *Israel J. Math.* **91** (1995), 349-71. MR **97e:**15011

[Sh1] Shalev, A., A theorem on random matrices and some applications, *J. Algebra* **199** (1998), 124-141. MR **99a:**20048

[Sh2] Shalev, A., Probabilistic group theory, in *Groups St. Andrews 1997*, London Math. Soc. Lecture Note Ser. 261, Cambridge Univ. Press (1999), 648-678. MR **2001b:**20117

[Sh3] Shalev, A., Asymptotic group theory, *Notices of the AMS* **48** (2001), 383-389. CMP 2001:09

[ShLl] Shepp, L.A. and Lloyd, S.P. , Ordered cycle lengths in a random permutation, *Trans. Amer. Math. Soc.* **121** (1966), 340-357. MR **33:**3320

[Shi] Shinoda, K., Identities of Euler and finite classical groups, in *Proceedings of Asian Mathematical Conference (Hong Kong, 1990)*, World Sci. Publishing (1992), 423-427. CMP 92:13

[So] Soshnikov, A., Universality at the edge of the spectrum in Wigner random matrices, *Comm. Math. Phys.* **207** (1999), 697-733. CMP 2000:05

[Sta] Stanley, R., Generalized riffle shuffles and quasisymmetric functions. Available at http://xxx.lanl.gov/abs/math.CO/9912025.

[Stei] Steinberg, R., Regular elements of semisimple algebraic groups, *R. Publ. Math. Inst. Hautes Etudes Sci.* **25** (1965), 49-80. MR **31:**4788

[St1] Stong, R., Some asymptotic results on finite vector spaces, *Adv. Appl. Math.* **9** (1988), 167-199. MR **89c:**05007

[St2] Stong, R., The average order of a matrix, *J. Combin. Theory Ser. A* **64** (1993), 337-343. MR **94j:**11094

[T] Thoma, E., Die unzerlegbaren, positiv-definiten Klassenfunktionen der abtahlbar unendlichen, symmetrischen Gruppe, *Math. Zeitschr.* **85** (1964), 40-61. MR **20:**3382

[VAr] Vera López, A. and Arregi, J., Some algorithms for the calculation of conjugacy classes in the Sylow p-subgroups of $GL(n,q)$, *J. Algebra* **177** (1995), 899-925. MR **96j:**20029

[VArV] Vera López, A., Arregi, J., and Vera López, F.J., On the number of conjugacy classes of the Sylow p-subgroups of $GL(n,q)$, *Bull. Austral. Math. Soc.* **52** (1995), 431-439. MR **96k:**20041

[VeSc] Vershik, A. and Schmidt, A., Limit measures arising in the asymptotic theory of the symmetric group, *Theory Probab. Appl.* **22** (1978), 72-88; **23** (1979), 42-54.

[W1] Wall, G.E., On conjugacy classes in the unitary, symplectic, and orthogonal groups, *J. Austr. Math. Soc.* **3** (1963), 1-63. MR **27:**212

[W2] Wall, G.E., Counting cyclic and separable matrices over a finite field, *Bull. Austral. Math. Soc.* **60** (1999), 253-284. MR **2000k:**11137

[Wi] Wieand, K., *Eigenvalue distributions of random matrices in the permutation group and compact Lie groups*, Ph.D. Thesis, Harvard University, 1998.

Department of Mathematics, University of Pittsburgh, 301 Thackeray Hall, Pittsburgh, PA 15260

E-mail address: fulman@math.pitt.edu

URL: http://www.math.pitt.edu/∼fulman/