Random Matrix Theory over Integers of Local Fields

Roger Van Peski

Advisor: Professor Ju-Lee Kim

Submitted in partial fulfillment
of the requirements for the degree of
Bachelor of Arts
Department of Mathematics
Princeton University
May 7, 2018

This thesis represents my own work in accordance with University regulations.

Roger Van Peski

I authorize Princeton University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize Princeton University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

Roger Van Peski

Abstract

This thesis is meant to bring together several different strands of existing work on random matrix theory over the ring of integers R of a non-Archimedean local field and over finite fields. It attempts to give some insight into the relation of these two theories to one another and to classical random matrix theory, while also presenting a few new results. We discuss the number-theoretic background motivating the study of random matrices over $\mathbb{Z}_p$ via the Cohen-Lenstra heuristics, as well as the analogies between eigenvalues of random matrices over $\mathbb{C}$ and cokernels of random matrices over R, which include Horn's (ex-)conjecture for Hermitian matrices and an analogous result for cokernels. We prove that the distribution of the cokernel of a product of Haar-distributed random matrices over R is described by products of Hall-Littlewood polynomials, in analogy with the known relation between the distribution of eigenvalues of a sum of random Hermitian matrices and multivariate Bessel functions. On the finite field side, we present a modified version of Fulman's construction of random partitions from random matrices over $\mathbb{F}_q$. Reinterpreting these partitions in terms of cokernels of random matrices over $\mathbb{F}_q[T]$, we prove that the same random partitions may also be obtained using a different distribution on $M_n(\mathbb{F}_q[T])$ which is more natural from the point of view of random matrices over rings of integers of global fields, of which the case of $\mathbb{Z}$ is well-studied. Finally, we define a plane partition associated to any matrix over R, which includes the information of both the matrix's cokernel and the partitions studied by Fulman; related to the distribution of this object is an interesting Markov chain on filtered vector spaces.

Acknowledgements

Roughly a year ago I walked into Professor Ozsváth's office in a mild panic over finding a thesis advisor, and left with the name 'Ju-Lee Kim'. I could not have imagined then how much I would benefit from her mathematical mentorship, and it is to her first and foremost that I owe a debt of gratitude for this thesis and for her advice and encouragement in the rest of my mathematical life during my senior year. She has been an extremely helpful and caring mentor, and I look forward to the opportunity (rare for a thesis advisor) of her being in my department for the coming years as well.

I would also like to thank the whole Princeton math department for being a both enjoyable and serious mathematical environment over the past few years. Professor Peter Ozsváth and Professor Sophie Morel in particular both gave generously of their time and expertise in teaching, reading courses and (in the latter case) my junior paper. My perspective on random matrix theory and enjoyment in playing with matrices have both benefited from many conversations with Professor Adam Marcus, and indeed the main result of Section 2.2 is a p-adic version of a result I learned through my work with him. Outside of Princeton I was fortunate to work under Professor Steven J. Miller of Williams and Professors Ken Ono and John Duncan of Emory during the past two summers, all of whom have continued to be supportive mentors long after their official time was up.

Several people deserve a more specialized thanks for specific contributions to this thesis. My second reader Professor Chris Skinner was a significant mathematical help in pursuing a direction related to the structure theory of modules over the Iwasawa algebra, and though this was a dead end, I was able to reach it and turn back faster through his guidance. Professor Jason Fulman, on whose work much of this thesis is based, was very helpful in answering some early questions about his papers. I would also like to thank Professor Alexei Borodin for helpful remarks during the MIT open house and for forwarding a recent paper of Fulman and Kaplan.

Nikita Lvov kindly explained his work on a different family of Markov chains coming from p-adic random matrix theory.

For many shared long nights and problem sets, good mathematical (and otherwise) conversations, and good meals I thank my Princeton math friends Chris Zhang, Josh Wang, Alex De Faveri, Rodrigo Angelo, and others. Sameera Vemulapalli has been a wonderful partner this year in grad school app-editing and in life, and I am grateful for both. I thank Nhi Truong for continuing to be a part of my mathematical life, and close friend, from the other side of the US. Lastly, I would like to thank my parents for their continued support, and my grandmother Ouma for making many things possible.

Contents

Introduction

The motivation for this thesis is to better understand random matrix theory over $\mathbb{Z}_p$ and other rings of integers of (non-Archimedean) local fields, and its relation to random matrix theory over finite fields and over $\mathbb{C}$ (the classical case). Each one is interesting and has been studied substantially in its own right, but most of the literature focuses on one specific case, while it is often easier to appreciate both the guiding themes and the differences between cases by looking at them side by side.

Classical random matrix theory over $\mathbb{C}$, while now a full-fledged field of its own, arguably grew out of physicist Eugene Wigner's attempts to model the energy levels of heavy nuclei using eigenvalues of random matrices [46, 47, 48]. In this setting, existing atomic physics predicted that the energy levels of such nuclei were given by eigenvalues of a corresponding infinite-dimensional Hermitian Hamiltonian operator, but this proved too difficult to analyze. Wigner's approach was to model this operator by a random matrix with independent Gaussian entries subject to a symmetry restriction so that the matrix would be Hermitian, and study the eigenvalue distribution of such random matrices as a model. Depending on whether one takes real, complex or quaternion Gaussian entries, the resulting random matrix ensembles are known as the Gaussian orthogonal ensemble (GOE), Gaussian unitary ensemble (GUE), and Gaussian symplectic ensemble (GSE).[1]

There is a general philosophy behind Wigner's development of early random matrix theory, which is worth mentioning because it will appear again. Suppose one has some collection of objects (in this case heavy nuclei) to which there is some associated matrix (the Hamiltonian) which appears to behave approximately randomly, and from which one can extract statistics one cares about (the energy levels, as its eigenvalues). Then study the distribution of the eigenvalues

[1] Throughout this thesis, a *random matrix ensemble* simply refers to a probability space with underlying set $M_n(K)$ for some field K. Often we are actually restricting to a subset of $M_n(K)$, but those elements outside the subset may simply be assigned probability 0. Often, the ensemble will implicitly refer to a series of ensembles for each n, as is the case with the GOE, GUE, and GSE just mentioned.

of matrices chosen randomly in the most natural way (Gaussian entries) subject to whatever restrictions one knows the actual matrices satisfy (the Hermitian condition). The eigenvalue distributions of the random matrix ensembles studied by Wigner ended up also providing very accurate predictions to number theory, specifically for the distribution of zeros of L-functions on the critical line, but this is more mysterious–for a readable survey see [18]. For other objects in number theory, such as the class groups of quadratic imaginary number fields, heuristics similar to Wigner's–often called Cohen-Lenstra heuristics after the inventors of the first [10]–are used to predict their distributions. The need to analyze random matrices over $\mathbb{Z}_p$ for such heuristics has given rise to most existing literature on random matrices over rings of integers of local fields, and we give an overview of this history in Section 1.1.

Classical random matrix theory is concerned with the distributions of eigenvalues of random Hermitian matrices, but it turns out that even before probability is introduced to the mix, the analogies between cokernels of matrices over DVRs and eigenvalues of Hermitian matrices go much deeper than might be expected. In particular, a very classical problem is to understand, if the eigenvalues of two Hermitian matrices A and B are known, what are the possible eigenvalues of $A + B$? This is the subject of Horn's celebrated conjecture [31], now proven, which gave necessary and sufficient conditions for the eigenvalues of $A + B$. Horn's conjecture has a rich history, and so it is natural to ask the analogous question for cokernels: if the cokernels of matrices A and B over a DVR are known, what are the possible cokernels of AB? What is most surprising from the perspective of this thesis is that this question has essentially the same answer as Horn's conjecture, after translating appropriately from eigenvalues to cokernels. This is detailed in Section 2.1, largely following the treatment of [27].

Section 2.2 discusses a probabilistic version of the same question: if A and B are random with some reasonable distribution, conditional on having certain eigenvalues (resp. cokernels), what is the distribution of the eigenvalues of $A + B$ (resp. distribution of the cokernel of AB)? When A and B come from one of the classical Hermitian random matrix ensembles, the answer is classically known and the distribution is governed by multivariate Bessel functions. On the cokernel side, we prove Theorem 2.2.17, which says that the distribution of cokernels of a product of random matrices over the ring of integers of a local field is described by the structure coefficients of a product of corresponding Hall-Littlewood polynomials. The proof is a modification of the argument used in [36] to compute the structure coefficients of the Hecke ring of $\mathrm{GL}_n(\mathbb{Z}_p)$, but the probabilistic interpretation appears new, and puts it in context as a

p-adic version of 'free convolution' in random matrix theory over $\mathbb{C}$ [29, 38, 39].

A third strand of random matrix theory is the theory over finite fields, which has closer ties to finite group theory and algebraic combinatorics, and which is introduced in Section 3.1. The main problem is to study the distribution of partitions associated to conjugacy classes in $\mathrm{GL}_n(\mathbb{F}_q)$ [24] or other finite groups of Lie type [20, 21]. The probability measures coming from these random partitions are, up to a small change in parameter, the same as the Cohen-Lenstra measure observed in the context of cokernels of random matrices over $\mathbb{Z}_p$. These measures may be defined and studied independently of their respective random matrix models, and [26] defines a two-parameter family which specializes to many of the cases considered in random matrix theory over $\mathbb{F}_q$ and $\mathbb{Z}_p$. There is also interesting algebraic combinatorics associated with these measures, in which the Hall-Littlewood polynomials also make an appearance [22], and certain Markov chains coming from random matrices over $\mathrm{GL}_n(\mathbb{F}_q)$ have been used to give a probabilistic proof of the Rogers-Ramanujan identities [23].

While literature on random matrices over rings of integers of local fields usually treats $M_n(\mathbb{Z}_p)$ with additive Haar measure (perhaps conditioned to lie on some subset), because of the connection to group theory most work over finite fields has been done on $\mathrm{GL}_n(\mathbb{F}_q)$ rather than $M_n(\mathbb{F}_q)$. To bypass this mismatch we reprove essentially the same results as [24] but over $M_n(\mathbb{F}_q)$ instead of $\mathrm{GL}_n(\mathbb{F}_q)$, for which except for a few small points the proofs are the same. We reinterpret the study of random matrices in $M_n(\mathbb{F}_q)$ in terms of random matrices in $M_n(\mathbb{F}_q[T])$, and prove in Theorem 3.2.2 that one can obtain the same random partitions from random matrices in $M_n(\mathbb{F}_q[T])$ with independent random entries taken uniformly from elements of $\mathbb{F}_q[T]$ of degree d, in the limit $n, d \to \infty$. This proof uses similar methods to the analysis of random matrices over $\mathbb{Z}$ in [44], and highlights the relation between the random partitions arising from random matrices over the integers of global fields and local fields.

Given that finite fields appear as residue fields of rings of integers of local fields, and the Cohen-Lenstra measures appear in random matrix theory over each, it seems natural to look for connections between both strands of random matrix theory. Section 3.3 discusses the role Markov chains play in the existing literature on random matrices over local and finite fields, and gives a simpler proof of a result of [15] that the cokernels of random matrices over $M_n(\mathbb{Z}_p)$ may be generated by a certain Markov chain. In an attempt to connect the random partitions and Markov chains studied on the finite field and local fields sides, we define (Definition 3.3.6) certain filtered vector spaces and a plane partition associated to any nonsingular $A \in M_n(\mathbb{Z}_p)$, which

include the data of both the cokernel and the partition associated to $A \pmod p$ in finite-field random matrix theory. In Theorem 3.3.11 we give progress toward computing the analogue of the Markov chain in [15] for these filtered vector spaces.

The background material in this thesis is generally introduced as needed, rather than all in the beginning. In particular, several things (Haar measure, a certain q-series identity) are used in Section 1.1 and the relevant background deferred to later; the reader with no familiarity with local fields may wish to skip to Section 1.2 as necessary while reading the next section. A note on probability: since the subject matter is largely algebraic we have suppressed most references σ-algebras, probability spaces, and the like, and only use random variables as a concrete means to visualize the probability measures we care about as their distributions. Nearly always, distinct random variables discussed in the same context will be independent and identically distributed (iid). We use the notation $\Pr(A|B)$ for the conditional probability of event A given event B throughout.

Chapter 1

Motivation and background

1.1 A short history of the Cohen-Lenstra heuristics

Number theory provides a wealth of natural classes of objects with enough structure to have interesting distributions, but enough depth that proving the accuracy of conjectured distributions is usually very difficult. Some of the most classical such objects are quadratic imaginary number fields; the study of the distribution of class groups among these fields goes back at least to Gauss, who conjectured finite lists of quadratic imaginary fields with class numbers $1, 2$ and 3, and that the class number $h(\mathbb{Q}(\sqrt{-d})) \to \infty$ as $d \to \infty$. All of these problems were subsequently resolved (see [28] for a survey of the class number 1 problem), but finer questions about the distribution of class groups of number fields remained open.

In 1983, number theorists Cohen and Lenstra looked at large tables of class groups and conjectured that, asymptotically, abelian p-groups appear as the p-Sylow subgroup of the class group of a randomly-chosen quadratic imaginary number field with probability inversely proportional to the order of their automorphism group [10]. Precisely, they conjectured the following.

Conjecture 1.1.1. *Let p be an odd prime, H an abelian p-group, and let G_p denote the p-Sylow subgroup of a group G and S_X the set of squarefree integers $1 \leq d \leq X$. Then*

$$\lim_{X \to \infty} \frac{|\{d \in S_X | \operatorname{Cl}(\mathbb{Q}(\sqrt{-d}))_p \cong H\}|}{|S_X|} = \frac{\prod_{i=1}^{\infty}(1 - p^{-i})}{|\operatorname{Aut}(H)|}.$$

The $\prod_{i=1}^{\infty}(1 - p^{-i})$ factor is simply a normalizing constant so that the probabilities sum to 1. This led to the *Cohen-Lenstra probability measure* on the set of abelian p-groups, assigning

to each one probability

$$\Pr(G) = \frac{\prod_{i=1}^{\infty}(1-p^{-i})}{|\operatorname{Aut}(G)|} \tag{1.1}$$

In Definition 1.3.10 we will define a one-parameter family of Cohen-Lenstra measures generalizing this, and show in Lemma 3.1.9 that they are probability measures, which is not at all clear from the above formula. Conjecture 1.1.1 is justified by a so-called Cohen-Lenstra heuristic, which reasons as follows.[1] Let K be a number field, $\mathcal{O}_K$ its ring of integers and I_K the group of fractional ideals of K (under multiplication). The image of the map $(\cdot) : K \to I_K$ given by $\alpha \mapsto (\alpha)$ is the set of all principal ideals, hence the cokernel of this map is exactly the class group $\mathrm{Cl}(K)$. Since $(\cdot)$ is in general a linear map between infinitely-generated $\mathbb{Z}$-modules, it would be easier to understand if we could restrict to maps of finite-dimensional modules. To do this, let S be any finite set of prime ideals such that the image of S in $\mathrm{Cl}(K)$ is a generating set, and let $I_K^S \subset I_K$ be the group they generate. Let $\mathcal{O}_S^*$ be the group of S-units, i.e. elements α such that the factorization of (α) contains only ideals in S. Then $(\cdot)$ restricts to a map $\mathcal{O}_S^* \to I_K^S$ of finitely-generated abelian groups, with cokernel $\mathrm{Cl}(K)$. To pick out the p-part of the class group, we tensor with $\mathbb{Z}_p$, since

$$\mathrm{Cl}(K)_p = \operatorname{coker}(\mathcal{O}_S^* \otimes \mathbb{Z}_p \to I_K^S \otimes \mathbb{Z}_p).$$

Hence, in general, $\mathrm{Cl}(K)_p$ is the cokernel of some $n \times n$ matrix with entries in $\mathbb{Z}_p$.

Here it is natural to wonder, in the same spirit as Wigner, how a 'randomly chosen' class group would behave if these matrices were randomly distributed in the most naive natural way possible. In 1987 Friedman and Washington [19] showed the following result.

Theorem 1.1.2. *Let p be any prime and μ_n be the additive Haar measure on $M_n(\mathbb{Z}_p)$ normalized to have total mass 1 (so it is a probability measure). Let H be any abelian p-group of rank $r = \dim_{\mathbb{F}_p}(H/pH)$, and define $\mu_n(H) := \mu_n(\{A \in M_n(\mathbb{Z}_p) : \operatorname{coker}(A) \cong H\})$. Then*

$$\mu_n(H) = \frac{\left(\prod_{i=1}^{n}(1-p^{-i})\right)\cdot\left(\prod_{i=n-r+1}^{n}(1-p^{-i})\right)}{|\operatorname{Aut}(H)|}. \tag{1.2}$$

In particular,

$$\lim_{n\to\infty}\mu_n(H) = \frac{\prod_{i=1}^{\infty}(1-p^{-i})}{|\operatorname{Aut}(H)|}, \tag{1.3}$$

which is the probability predicted for class groups in Conjecture 1.1.1.

[19] uses this as part of a heuristic justification for an analogous Cohen-Lenstra conjecture for Picard groups of hyperelliptic curves over finite fields, though it does not discuss the class

[1] This exposition draws heavily from [50].

group heuristic given above. In any case, Proposition 1.1.2 predicts that if the p-adic matrices $\mathcal{O}_S^* \otimes \mathbb{Z}_p \to I_K^S \otimes \mathbb{Z}_p$ are approximately distributed according to the Haar probability measure as one varies through quadratic imaginary fields, then Conjecture 1.1.1 should hold. We will not prove Theorem 1.1.2 yet, but it will follow by Theorem 3.3.4.

Heuristics based on similar reasoning abound. The Tate-Shafarevich group of an elliptic curve over $\mathbb{Q}$ is a group Ш which, when finite (and conjecturally this is always the case), has a nondegenerate alternating pairing $\beta : Ш \times Ш \to \mathbb{Q}/\mathbb{Z}$ [40]. Applying $\otimes_{\mathbb{Z}}\mathbb{Z}_p$ as before, there is a nondegenerate alternating pairing $\beta_p : Ш_p \times Ш_p \to \mathbb{Q}_p/\mathbb{Z}_p$. Abelian groups with such a pairing are referred to as *symplectic abelian groups* in [4] or *groups of type S* in [13].

Suppose instead that one starts with such a pairing $\langle \cdot, \cdot \rangle : \mathbb{Z}_p^n \times \mathbb{Z}_p^n \to \mathbb{Z}_p$, which is represented by a matrix A such that $A = -A^T$. Then after checking some details (see [4, p. 7-8]), $\langle \cdot, \cdot \rangle$ induces a nondegenerate alternating pairing

$$\langle \cdot, \cdot \rangle_A : \operatorname{coker}(A) \times \operatorname{coker}(A) \to \mathbb{Q}_p/\mathbb{Z}_p. \tag{1.4}$$

Hence one natural way to produce a random abelian p-group H with an alternating pairing $H \times H \to \mathbb{Q}_p/\mathbb{Z}_p$ is to take the cokernel of a random alternating matrix with respect to the Haar measure, with the alternating pairing above. The resulting distribution on symplectic abelian p-groups converges to a fixed distribution as the matrix size $n \to \infty$; letting $\mu_{s,p}$ be the limiting distribution, [4] conjectured the following[2] :

Conjecture 1.1.3. *Varying through the set of elliptic curves over $\mathbb{Q}$ ordered by height, the distribution of $Ш_p$ is asymptotically given by $\mu_{s,p}$.*

This heuristic is different from the last, and perhaps less intuitively satisfying, because the authors do not give a corresponding construction of Ш as a cokernel of an alternating matrix analogously to the construction of $\mathrm{Cl}(K)$ as a cokernel of a general matrix. Using very similar reasoning to the proof of Theorem 1.1.2 in [19], [4] explicitly computes $\mu_{s,p}$. For a symplectic abelian p-group H with alternating form $\langle \cdot, \cdot \rangle$, let $\operatorname{Aut}_S(H)$ be the set of automorphisms ϕ of H which preserve the form ($\langle \phi(g), \phi(g') \rangle = \langle g, g' \rangle$ for any $g, g' \in H$).

Theorem 1.1.4. *For p prime, H a symplectic abelian p-group, and $\mu_{n,alt}$ the Haar probability measure on the set of alternating matrices in $M_n(\mathbb{Z}_p)$, then*

$$\lim_{\substack{n \to \infty \\ n \text{ even}}} \mu_{n,alt}(\{A \in M_n(\mathbb{Z}_p)_{alt} : \operatorname{coker}(A) \cong H\}) = \prod_{i=1}^{\infty}(1 - p^{1-2i}) \cdot \frac{|H|}{|\operatorname{Aut}_S(H)|}. \tag{1.5}$$

[2] Actually, they give a stronger conjecture on the distribution of a certain short exact sequence involving the p-parts of the group of rational points and the Selmer group as well as Ш.

As with Theorem 1.1.2, this formula involves division by the order of an automorphism group, for the appropriate notion of automorphism. There are other, simpler heuristic arguments for why many objects occurring 'in nature' should appear with frequency inversely proportional to the order of their automorphism groups, of which we give one simple example from [51]. Consider a graph G on n labeled vertices, where S_n acts on the set of such graphs by permuting the labels. Two graphs are said to be isomorphic if there is a bijection between their vertex sets such that two vertices are connected by an edge if and only if their images are, so it is clear that action by a permutation always gives an isomorphic graph. Some permutations will fix G, and there are $|\operatorname{Aut}(G)|$ of these. Hence by the orbit-stabilizer theorem, the number of graphs on n labeled vertices which are isomorphic to G is $\frac{n!}{|\operatorname{Aut}(G)|}$.

When sizes of automorphism groups appear during this thesis, it is essentially always through some application of the orbit-stabilizer theorem, which is worth keeping in mind. A general formulation of probability measures coming from orbit-stabilizer counting comes from the framework of groupoid cardinality, and includes the Poisson distribution on $\mathbb{Z}_{\geq 0}$ as well–see [32] for a good discussion of this, and [52] for a more general exposition of groupoid probability.

However, our interest in Cohen-Lenstra heuristics is mainly in how they motivate the study of various distributions on p-adic random matrices. Before moving entirely to random matrices, it is worth mentioning some other related and recent work.

- [50] studies the $n \to \infty$ limiting distribution of cokernels of Haar-distributed random matrices in $M_{n\times(n+u)}(\mathbb{Z}_p)$, which for $u = 0$ is the Cohen-Lenstra distribution and for $u = 1$ conjecturally describes the distribution of p-parts of class groups of real quadratic fields [10].

- [37] and [50] show that random matrices with iid entries taken from distributions very different from the additive Haar measure still have cokernels distributed according to the Cohen-Lenstra measure, provided the distributions are not too concentrated around any particular value in a precise sense. Studying cokernels of random symmetric matrices for applications to random graphs, [49] shows similar universality results when the distributions of entries are not independent. The analogous phenomenon of universality of eigenvalue distributions in classical random matrix theory is extremely well-studied [12].

- Cokernels of random integer matrices are closely related, and have received attention in their own right from [44] and [9] among others (though the latter uses the language of

random lattices in $\mathbb{Z}^d$).

- Because actually proving that any class of abelian p-groups occuring in number theory is distributed as expected is generally very difficult, it may be easier to study restricted information, such as the distribution of the p-groups' ranks rather than the groups themselves. Clearly the rank of the cokernel of a nonsingular matrix in $M_n(\mathbb{Z}_p)$ is the same as the dimension of the kernel of its reduction modulo p. [1] uses this to model the p-torsion of Jacobians of certain families of curves over finite fields by the ranks of random matrices in $\mathrm{Sp}_{2n}(\mathbb{F}_p)$.

Contrary to what the last point might suggest, the majority of the work on random matrices over finite fields was done on its own without an eye toward applications to number theory. The study of random matrices over finite fields is not merely a toy case for random matrices over local fields, but has been pushed in different and equally interesting directions, which are detailed in Section 3.1.

It is worth noting that studying eigenvalues of the classical ensembles and cokernels of $M_n(\mathbb{Z}_p)$ with additive Haar measure are both specific cases of the following kind of question:

Let G be a group which acts on a space S with a probability measure preserved by the action of G. What is the resulting probability measure on the set of G-orbits in S?

It is time for an explicit definition of one of the classical ensembles, for which a standard reference is [43]. The other classical ensembles, the Gaussian unitary ensemble and Gaussian symplectic ensemble, may be defined analogously.

Definition 1.1.5. The $N \times N$ *Gaussian orthogonal ensemble (GOE)* is the random matrix ensemble given by

$$\begin{pmatrix} c_{0,0} & c_{0,1} & \cdots & c_{0,N-1} \\ c_{1,0} & c_{1,1} & \cdots & c_{1,N-1} \\ \vdots & \vdots & \ddots & \vdots \\ c_{N-1,0} & c_{N-1,1} & \cdots & c_{N-1,N-1} \end{pmatrix}, \tag{1.6}$$

with the $c_{\ell,j}$ are random variables defined as follows. For $\ell < j$ $c_{\ell,j}$ are iid Gaussians with mean 0 and variance 1, and $c_{j,\ell} = c_{\ell,j}$ (i.e. the ensemble is symmetric). Furthermore, $c_{j,j}$ are iid Gaussians with mean 0 and variance 2.

This is a good practical definition, but the real reason the GOE is studied is that the probability measure defined above is *invariant under conjugation by the orthogonal group* $O(n)$. Putting the GOE in the framework above, S is the set of real-symmetric matrices, and G is the orthogonal group acting by conjugation. By the spectral theorem any symmetric matrix is diagonalizable by orthogonal matrices, and hence orbits are exactly determined by n-tuples of real eigenvalues and convenient representatives of our orbits are given by diagonal matrices. Thus any question about the distribution of eigenvalues in the GOE is a question about G-orbits.

In the local field case $S = M_n(\mathbb{Z}_p)$ with additive Haar measure, which is invariant by left- and right-multiplication by $\mathrm{GL}_n(\mathbb{Z}_p)$. Hence the group $G = \mathrm{GL}_n(\mathbb{Z}_p) \times \mathrm{GL}_n(\mathbb{Z}_p)^{op}$ acts on S, where the left factor multiplies on the left and the right factor multiplies on the right (which is why we need to take the opposite group to make this a group action). Analogously to the spectral theorem, we have:

Proposition 1.1.6 (Smith normal form)**.** *Let T be a principal ideal ring (not necessarily a domain), $m, n \in \mathbb{N}$, and $A \in M_{m\times n}(T)$. Then there exist matrices $P \in M_m(T)$ and $Q \in M_n(T)$ such that PAQ is diagonal, and its ii^{th} entry divides its jj^{th} entry for all $1 \leq j \leq i \leq \min(m, n)$. This diagonal form is unique up to unit multiplication.*

This may be proved via elementary row operations, using Bezout's lemma to extract the GCDs of the matrix entries. In the case where our principal ideal ring is actually a DVR, the diagonal entries of PAQ are (up to a unit) just powers of the uniformizer. Strictly speaking one has singular matrices in $M_n(\mathbb{Z}_p)$ with Smith normal forms having 0's on the diagonal as well, but because these only occur with probability 0 we may ignore them. Excepting these, the G-orbits in S are parametrized by partitions of length at most n which give the powers of p occuring on the diagonal. The astute reader may recognize that this situation is actually more similar to studying singular values rather than eigenvalues, since the former parametrize double cosets in $U(n)\backslash M_n(\mathbb{C})/U(n)$, but the principle of studying measures of orbits is the same.

1.2 Local field background

In Section 1.1, we followed the majority of the Cohen-Lenstra literature in restricting discussion of general rings of integers of local fields to $\mathbb{Z}_p$ only. However, the arguments are largely the same for any non-Archimedean local field, and the main results will be proved in this setting, so it is worthwhile to start arguing in general. We will only be concerned with the additive

structure of these rings of integers and nothing especially subtle about local fields is needed; the material of this section may be found in any standard reference such as [8].

Definition 1.2.1. An *norm* on a field K is a map $|\cdot| : K \to \mathbb{R}_{\geq 0}$ such that for all $x, y \in K$,

1. $|x| = 0 \iff x = 0$,
2. $|xy| = |x| \cdot |y|$,
3. $|x + y| \leq |x| + |y|$.

$|\cdot|$ is

- *non-Archimedean* if the last condition is replaced by the stronger *ultrametric inequality* $|x + y| \leq \max(|x|, |y|)$.
- *discrete* if $\{\log |x| : x \in K^\times\}$ is a discrete subgroup of the additive group $\mathbb{R}$.
- *trivial* if $|x| = 1$ for all $x \in K^\times$, and *nontrivial* otherwise.

A norm defines a metric $d(x, y) = |x - y|$ on K, and hence a topology. K is complete with respect to this metric if all Cauchy sequences in the metric converge to an element of K.

Definition 1.2.2. A *non-Archimedean local field* is a field K which is locally compact and complete with respect to the topology induced by a nontrivial non-Archimedean discrete norm $|\cdot|$.

The proviso that $|\cdot|$ is nontrivial is necessary since otherwise a finite field with the trivial norm would be a local field. From here onward, local field will refer only to non-Archimedean local fields for brevity, and K will always denote a local field.

Definition 1.2.3. The *ring of integers* of a local field K is $R := \{x \in K : |x| \leq 1\}$.

Example 1.2.4. The canonical example of a local field is $\mathbb{Q}_p$ with norm $|x| = p^{-\operatorname{val}_p(x)}$, where $\operatorname{val}_p(x)$ is the largest power of p such that there exists $x' \in \mathbb{Z}_p$ for which $p^{\operatorname{val}_p(x)} x' = x$. Another equally fundamental example is the power series field $\mathbb{F}_q[\![t]\!]$ over a finite field $\mathbb{F}_q$ with norm defined similarly using valuation with respect to t instead of p.

Lemma 1.2.5. *For a local field K, $\mathfrak{p} := \{x \in K : |x| < 1\}$ is the unique maximal ideal of R, and is principal. All other nontrivial ideals are given by powers of $\mathfrak{p}$.*

Proof. It follows immediately from the definition of $|\cdot|$ (remembering the ultrametric inequality) that $\mathfrak{p}$ is an ideal, as it does that R is indeed a ring. Since $|\cdot|$ is discrete, there exists an element $\omega \in \mathfrak{p}$ of maximal norm. For any other $x \in \mathfrak{p}$, we have that $|x/\omega| = |x|/|\omega| \leq 1$ by maximality, hence there exists $u \in R$ such that $x = u\omega$, so ω is a uniformizer for the ideal $\mathfrak{p}$.

By the same argument any other ideal $\mathfrak{a}$ is generated by some $a \in R$. If $|a| = 1$ then $|1/a| = 1$ so $a \in R^\times$ and hence $\mathfrak{a} = R$; otherwise, $\mathfrak{a} \subset \mathfrak{p}$.

Let n be the lowest integer such that $|\omega|^n \leq |a|$. If $|\omega|^n < |a|$, then $|\omega^n/a| < 1$ and $|\omega^{n-1}/a| > 1$, hence $|\omega^n/a| > |\omega|$ and $\omega^n/a \in R$, contradicting maximality. Hence $|\omega^n| = |a|$ so $a = u\omega^n$ for some $u \in R^\times$ and $\mathfrak{a} = (\omega^n)$. □

Definition 1.2.6. We refer to a generator ω of $\mathfrak{p}$ as a *uniformizer* of K. Given a uniformizer ω, by Lemma 1.2.5 every element of $K^\times$ is of the form $u\omega^n$ for u a unit in R and $n \in \mathbb{Z}$. We define the *valuation* on K by $\mathrm{val}(u\omega^n) = n$. By convention $\mathrm{val}(0) = \infty$.

It may be checked that the valuation is independent of the choice of uniformizer. It is easy to see that the valuation and the norm are related by $|a| = |\omega|^{\mathrm{val}(a)}$, and the valuation may also be defined as $\mathrm{val}(a) = \log_{|\omega|}(|a|)$. The following nontrivial theorem shows that all local fields come from the two given in the previous example.

Theorem 1.2.7. *Every non-Archimedean local field of characteristic* 0 *is a finite algebraic extension of* $\mathbb{Q}_p$ *for some* p*, and every one of characteristic* p *is a finite extension of a field of Laurent series* $\mathbb{F}_q((t))$ *for* q *a power of* p*.*

It is common to define $\mathbb{Z}_p$ as the ring of 'power series in p', i.e. expansions of the form $\sum_{i=0}^{\infty} a_i p^i$ where $a \in \{0, \ldots, p-1\}$. This representation is the most useful way to concretely manipulate p-adic integers for the later results of this thesis, and the lemma below shows that it extends to general local fields. We use here that the residue field $R/\mathfrak{p}$ is compact and discrete, hence finite and isomorphic to $\mathbb{F}_q$ for some q.

Lemma 1.2.8. *Let* K *be a local field,* ω *a uniformizer for* p*, and* $S \subset R$ *be a (finite) set of representatives for* $R/\mathfrak{p}$*. Then every* $x \in R$ *may be written as*

$$x = \sum_{i=0}^{\infty} a_i \omega^i \tag{1.7}$$

for a unique sequence $(a_i)_{i \geq 0}$ *of elements of* S*.*

Proof. We induct on the decomposition $x = a_0 + a_1\omega + \ldots + a_n\omega^n + b\omega^{n+1}$ for $b \in R$ and $a_i \in S$. For the base case, choose a_0 to be the representative of $x \pmod{\mathfrak{p}}$; then $x = a_0 + b_1\omega$ for some $b_1 \in R$. If we have a decomposition $x = a_0 + a_1\omega + \ldots + a_n\omega^n + b\omega^{n+1}$, then $x \equiv a_0 + a_1\omega + \ldots + a_n\omega^n + a_{n+1}\omega^{n+1} \pmod{\mathfrak{p}^{n+2}}$ for a unique a_{n+1}. Hence there is a uniquely determined sequence $(a_i)_{i\geq 0}$ as required. □

These power series expansions mesh well with the additive Haar measure on the additive group R.

Definition 1.2.9. The Haar measure μ on a compact topological group is the unique (up to scaling) measure such that, for any set $S \subset G$ measurable with respect to the Borel σ-algebra generated by open sets in R,

$$\mu(\{g \cdot s : s \in S\}) = \mu(\{s \cdot g : s \in S\}) = \mu(S) \tag{1.8}$$

for all $g \in G$. We denote the first two sets of the previous equation as $g \cdot S$ and $S \cdot g$ respectively.

A proof of its existence and uniqueness for compact topological groups may be found in [30]. Given any open set S of R containing 0, we have $R = \bigcup_{a\in R}(S + a)$; R is compact because it is a closed and bounded set in a locally compact metric space, hence a finite collection of the sets $S + a$ suffices to cover R. Since each set has the same measure by translation invariance, R has finite Haar measure. Hence we may scale the Haar measure so that $\mu(R) = 1$ and it defines a probability measure. The following lemma gives a very useful interpretation of this measure in terms of power series, and is the way the Haar probability measure should be thought of in all subsequent results.

Lemma 1.2.10. *Let $S \subset R$ be measurable with respect to μ the normalized Haar probability measure on R. Choose a random element x of R by choosing each coefficient in the power series representation of Lemma 1.2.8 uniformly randomly from the given set of representatives of $R/\omega \cong \mathbb{F}_q$. Then $\mu(S)$ is equal to the probability that x lies in S.*

Proof. Let $\mu'(S) = \Pr(x \in S)$ for x chosen randomly as above. μ' is the measure induced by the random variable x, so it is a measure as well. Given any $a \in R$, we see that $x + a$ has the same distribution as x for x chosen as above, therefore $\mu'(S + a) = \mu'(S)$ for all S and a. Since both μ and μ' have total mass 1, by uniqueness of the Haar measure $\mu' = \mu$. □

Corollary 1.2.11. *For any n and x chosen randomly as above, $x \pmod{\omega^n}$ has uniform distribution in R/ω^n.*

Notation 1.2.12. As a recap, we fix the following conventions for the rest of this thesis:

1. R denotes the ring of integers of a non-Archimedean local field.
2. ω is a generator of the maximal ideal of R.
3. $R/\omega \cong \mathbb{F}_q$ is the residue field of R.

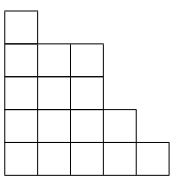
The remainder of the results of this thesis will be stated for general rings of integers of local fields, but the reader looking for concreteness may mentally substitute $\mathbb{Z}_p$ for R and p for ω.

1.3 Integer partitions and R-modules

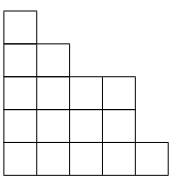
We now consider finitely-generated torsion modules over R. When $R = \mathbb{Z}_p$, this reduces to studying the group automorphisms of an abelian p-group as in Section 1.1. The Cohen-Lenstra measure can be defined in completely the same way in this more general setting of R-modules. The following proposition shows that if R is the ring of integers of a local field, then for our purposes the only feature of R which matters is the size of its residue field. We define our conventions regarding integer partitions.

Definition 1.3.1 (Integer partitions)**.** An integer partition is a nonincreasing tuple $\lambda = (\lambda_1, \ldots, \lambda_n, 0, \ldots, 0)$ of nonnegative integers, which are referred to as its *parts*. We view two partitions as equivalent if they have the same nonzero parts. A partition λ with n nonzero parts as above is said to have length n, and we denote the length of a partition by $\ell(\lambda)$. λ is *a partition of m* if $|\lambda| := \sum_{i=1}^{n} \lambda_i$ is equal to m. The *conjugate partition* λ' of λ is the partition with parts $\lambda'_i = |\{j : \lambda_j \geq i\}|$. $\mathcal{P}$ will refer to the set of all integer partitions, and () will refer to the empty partition.

Example 1.3.2. Visually, when representing a partition by its Ferrers diagram (in French notation), the conjugate partition is the flipped diagram of the partition.



(a) $\lambda = (5, 4, 4, 2, 1)$



(b) $\lambda' = (5, 4, 3, 3, 1)$

Notation 1.3.3. For a partition λ of length n, M_λ will denote the module $\bigoplus_{i=1}^n R/\omega^{\lambda_i}$ when R is clear from context.

The following result is standard and may be found, for instance, in [36]. The proof below is adapted from a proof of the special case $R = \mathbb{Z}_p$ appearing in [35].

Proposition 1.3.4. *Let R be the ring of integers of a local field with residue field $R/\omega \cong \mathbb{F}_q$. Then $|\operatorname{Aut}_R(M_\lambda)| = q^{\sum_{i\geq 1}\lambda_i'^2}\prod_{i\geq 1}(q^{-1})_{m_i(\lambda)}$, and in particular depends only on q and λ, not on R.*

Proof. Write $\lambda = (t_1[e_1], \ldots, t_r[e_r])$ where t_i ranges through the distinct parts of λ, $e_i = m_{t_i}(\lambda)$, and $t_i[e_i]$ denotes e_i copies of t_i. Then the endomorphisms of M_λ are given by block matrices

$$\begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,\ell(\lambda)} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,\ell(\lambda)} \\ \vdots & \vdots & \ddots & \vdots \\ A_{\ell(\lambda),1} & \cdots & A_{\ell(\lambda),\ell(\lambda)-1} & A_{\ell(\lambda),\ell(\lambda)} \end{pmatrix} \tag{1.9}$$

where

$$A_{i,j} \in \begin{cases} M_{e_i\times e_j}(\omega^{t_i-t_j}R/\omega^{t_j}) & i < j \\ M_{e_i\times e_j}(R/\omega^{t_j}) & i \geq j \end{cases}. \tag{1.10}$$

It is easy to check that the set of such matrices is closed under multiplication. Because M_λ is finite, an endomorphism is an automorphism if and only if it is surjective, and by right-exactness of the tensor product this is true if and only if the corresonding map $M_\lambda \otimes_R R/\omega \to M_\lambda \otimes_R R/\omega$ is surjective. This is just a map of vector spaces, so it suffices to check the vanishing of its determinant. By the above description its determinant is just $\prod_{i=1}^r \det(A_{i,i}) \pmod{\omega}$. Hence the automorphisms are the endomorphisms for which $A_{i,i}$ is invertible for each $1 \leq i \leq t$.

Since $|R/\omega| = q$, $|\omega^i R/\omega^j| = q^{j-i}$ for $i \leq j$. Hence for $i = j$, the number of $A_{i,i}$ is $(q^{-1})_{e_i}\cdot q^{e_i^2 t_i}$, and for $i \neq j$, the number of $A_{i,j}$ is $q^{e_i e_j \min(t_i,t_j)}$. Hence

$$|\operatorname{Aut}_R(M_\lambda)| = q^{\sum_{i=1}^r e_i^2 t_i}\prod_{i=1}^r (q^{-1})_{e_i}\left(\prod_{i\neq j} q^{e_i e_j \min(t_i,t_j)}\right). \tag{1.11}$$

This is equal to $q^{\sum_{i\geq 1}\lambda_i'^2}\prod_{i\geq 1}(q^{-1})_{m_i(\lambda)}$. □

Notation 1.3.5. In view of Proposition 1.3.4, for any prime power q we define $|\operatorname{Aut}_q(\lambda)| := q^{\sum_{i\geq 1}\lambda_i'^2}\prod_{i\geq 1}(q^{-1})_{m_i(\lambda)}$. We may formally extend $|\operatorname{Aut}_q(\lambda)|$ to values of q which are not prime powers.

It is thus reasonable to abstract away from the specific choice of R and consider probability measures living on integer partitions and given in terms of the size q of the residue field.

Notation 1.3.6. For all real $q > 1$, we define $(q^{-1})_n := \prod_{i=1}^{n}(1-q^{-i})$, and $(q^{-1})_\infty := \prod_{i\geq 0}(1-q^{-i})$. Furthermore, for $q > 1$ and $0 \leq u < q$ let $(u;q^{-1})_n := \prod_{i=0}^{n}(1-uq^{-i})$ and define $(u;q^{-1})_\infty$ similarly. We note that the index i starts from 0 rather than 1 in $(u;q^{-1})_n$, so $(q^{-1})_n = (q^{-1};q^{-1})_{n-1}$.

Remark 1.3.7. The most useful interpretation of $(q^{-1})_n$, for us, is that it is equal to $\frac{|\mathrm{GL}_n(\mathbb{F}_q)|}{|M_n(\mathbb{F}_q)|}$. This is easy to see as follows: to choose an element of $\mathrm{GL}_n(\mathbb{F}_q)$, there are $q^n - 1$ choices for the first column (we cannot choose all zeros), then $q^n - q$ choices for the second since it has to be linearly independent, and in general $q^n - q^{i-1}$ for the i^{th} column. Dividing by $|M_n(\mathbb{F}_q)| = q^{n^2}$ yields $(q^{-1})_n$.

Remark 1.3.8. Note that in much of the random matrices over finite fields literature such as [24, 25], what we denote by $(u;q^{-1})_n$ is instead denoted $(u;q)_n$. However, in classical partition theory the notation above is typically used, e.g. in [3].

In keeping with the shift in focus from abelian groups and R-modules to integer partitions, we define the following.

Definition 1.3.9. Let S be a DVR with uniformizer ω and $A \in M_n(S)$ be nonsingular. Then we define the *cokernel partition* $\lambda(A)$ *of* A to be the partition such that $\mathrm{coker}(A) \cong \bigoplus_{i\geq 1} S/\omega^{\lambda(A)_i}$.

Definition 1.3.10. For $q > 1$, we define the *Cohen-Lenstra measure* μ_q on the set of integer partitions $\mathcal{P}$ by

$$\mu_q(\lambda) = \frac{(q^{-1})_\infty}{|\operatorname{Aut}_q(\lambda)|}. \tag{1.12}$$

For $q = p$ this is the Cohen-Lenstra measure discussed in Section 1.1 (viewed as a measure on cokernel partitions rather than the p-groups arising as cokernels themselves), which was the one originally studied by Cohen and Lenstra. It is not obvious that it is indeed a probability measure, i.e. that $\sum_{\lambda\in\mathcal{P}} \frac{1}{|\operatorname{Aut}_q(\lambda)|} = \frac{1}{(q^{-1})_\infty}$, but this will follow from showing that a more general measure is a probability measure in Lemma 3.1.9. μ_q is not always referred to as the Cohen-Lenstra measure in the literature, for instance in [24], since it was discovered in the context of random matrices over finite fields independently of number-theoretic motivation. A discussion of these parallel histories may be found in [25] or [35].

Chapter 2

Random matrices over integers of local fields and analogies with classical random matrix theory

We examine the analogy between eigenvalues of random matrices over $\mathbb{R}$ or $\mathbb{C}$ and cokernels of random matrices over the ring of integers of a local field, beginning with a very classical question about eigenvalues which predates random matrix theory. The following section is mostly condensed from [27].

2.1 Existence questions and Horn's conjecture

The following are two fundamental questions on how eigenvalues (resp. cokernels) behave with respect to matrix addition (resp. multiplication).

Question 2.1.1. *Given $n \times n$ Hermitian matrices A and B with eigenvalues $\alpha_1 > \alpha_2 > \ldots > \alpha_n$ and $\beta_1 > \beta_2 > \ldots > \beta_n$ respectively, what are the possible n-tuples of eigenvalues $\gamma_1 > \gamma_2 > \ldots > \gamma_n$ of the sum $A + B$?*

Question 2.1.2. *Given a DVR S and nonsingular matrices $A, B \in M_n(S)$ with cokernel partitions λ, μ respectively, what are the possible cokernel partitions ν of the product AB?*

We begin with Question 2.1.1. The most obvious condition $\gamma = (\gamma_1, \ldots, \gamma_n)$ must satisfy is $\sum_{i=1}^n \gamma_i = \sum_{i=1}^n \alpha_i + \sum_{i=1}^n \beta_i$, since the trace is the sum of the eigenvalues. Weyl [45] gave more

necessary conditions on γ in the form of inequalities which now bear his name, the most famous being the Cauchy interlacing property: If A, B are Hermitian and B has rank 1, then $\alpha_i \geq \gamma_{i+1}$ and $\gamma_i \geq \alpha_{i+1}$ for all $1 \leq i \leq n-1$.

After many more partial results, Horn conjectured that the trivial condition $\sum_i \alpha_i + \sum_i \beta_i = \sum_i \gamma_i$, along with a certain list of inequalities on the α_i, β_i and γ_i which he defined inductively, yielded necessary and sufficient conditions for them to be the eigenvalues of Hermitian matrices A, B and $C = A + B$ respectively [31]. Each inequality had the form

$$\sum_{k \in K} \gamma_k \leq \sum_{i \in I} \alpha_i + \sum_{j \in J} \beta_j \tag{2.1}$$

for triples of sets $I, J, K \in \{1, \ldots, n\}$ all of the same cardinality r. For the exact definition of which such triples Horn's list consisted of, see [27]; we shall just refer to the set of these triples as H_n.

The answer to Question 2.1.2, surprisingly, is almost exactly the same as the answer to Question 2.1.1.

Theorem 2.1.3. *Let S be a DVR. Then λ, μ, ν occur as the cokernel partitions of some $A, B, C \in M_n(S)$ with $A \cdot B = C$ if and only if $|\nu| = |\lambda| + |\mu|$ and the inequality $\sum_{k \in K} \nu_k \leq \sum_{i \in I} \lambda_i + \sum_{j \in J} \mu_j$ holds for every $(I, J, K) \in H_n$.*

In addition to Horn's inductive description of H_n, there are several nontrivial alternative characterizations. First recall the bijection between sets $I = \{i_1, \ldots, i_r\}$ of cardinality r and partitions of length at most r given by

$$I = \{i_1, \ldots, i_r\} \leftrightarrow \lambda_I := (i_r - r, i_{r-1} - (r-1), \ldots, i_1 - 1) \tag{2.2}$$

where we choose labels so $i_r > i_{r-1} > \ldots, > i_1$.

Theorem 2.1.4. *For a triple (I, J, K) as above, the following are equivalent:*

1. *$(I, J, K) \in H_n$.*
2. *There exist Hermitian matrices A, B, C such that λ_I, λ_J and λ_K are the eigenvalues of A, B and C respectively, and $A + B = C$.*

This result is intriguing because the inequalities given by triples $(I, J, K) \in H_n$ serve to characterize the real eigenvalues of a sum of Hermitian matrices, but the inequalities themselves

come from triples of Hermitian matrices with *integer* eigenvalues, which are of course highly constrained. To actually find every triple in H_n via the second condition of Theorem 2.1.4 would still be difficult, however.

2.2 Distributions of eigenvalues and cokernel partitions

The previous discussion concerned the *extreme case* of the possible cokernel partitions of a product of matrices, but it is just as natural to ask the probabilistic version of this question. Here we return to the local field setting.

Question 2.2.1. *Let λ, μ be two partitions of length at most n. If A and B are chosen from $M_n(R)$ with respect to the additive Haar measure, conditional on* $\mathrm{coker}(A)$ *and* $\mathrm{coker}(B)$ *having type λ and μ respectively, what is the probability distribution of* $\mathrm{coker}(AB)$*?*

Thankfully, as with the extremal question, guidance comes from the answer in classical random matrix theory to the analogous question.

Question 2.2.2. *Let $\alpha_1 > \alpha_2 > \ldots > \alpha_n$ and $\beta_1 > \beta_2 > \ldots > \beta_n$ be real numbers. If A and B are chosen from the $n \times n$ GOE, GUE or GSE conditional on having α and β as their respective eigenvalues, what is the distribution of the eigenvalues of $A + B$?*

If random matrices A, B have eigenvalue distributions μ, ν, the eigenvalue distribution of $A + B$ is sometimes referred to as the additive convolution of μ and ν, and for AB there is a similar notion of multiplicative convolution. When studying the limit as matrix size goes to ∞, these are closely related to free probability, see e.g. [39]. For finite matrices, there is a related theory of finite free probability recently developed in [38]. For the classical ensembles of Question 2.2.2, the answer is classically known, and the following exposition of it largely follows [29].

We now return to Questions 2.2.1 and 2.2.2. At this point it is necessary to introduce certain symmetric functions associated to partitions, which will be useful later as well as in this section. The reader is warned in advance that the definitions of Jack and multivariate Bessel functions below are not at all enlightening without more context; they are introduced here for completeness and so as to highlight the similarity of the answers to the aforementioned questions, which both come from Macdonald polynomials.

Notation 2.2.3. The *Macdonald polynomial* $P_\lambda(x_1, \ldots, x_n; q, t)$ is a symmetric polynomial in the n variables $x_1, \ldots, x_n$ associated to a partition λ of length $\leq n$. Often we will write $P_\lambda(x; q, t)$ with x denoting the n variables.

The standard reference for these and other symmetric polynomials is [36]. Macdonald polynomials are usually defined as eigenfunctions of a certain differential operator on the algebra $\mathbb{Q}(q,t)[x_1, \ldots, x_n]^{S_n}$ of symmetric polynomials, and span the $\mathbb{Q}(q,t)$-vector space of such polynomials. Because of this last property, any product $P_\lambda P_\mu$ is expressible as a linear combination of Macdonald polynomials P_ν with certain structure coefficients in $\mathbb{Q}(q,t)$,

$$P_\lambda(x; q, t) P_\mu(x; q, t) = \sum_{\nu \in \mathcal{P}} m^\nu_{\lambda,\mu}(q,t) P_\nu(x; q, t). \tag{2.3}$$

In particular, this is true with all x_i set to 0. Hence, defining a normalized Macdonald polynomial similarly to Definition 2.2.6 by

$$\hat{P}_\lambda(x; q, t) = \frac{P_\lambda(x; q, t)}{P_\lambda(0; q, t)}, \tag{2.4}$$

we have that

$$\hat{P}_\lambda(x; q, t) \hat{P}_\mu(x; q, t) = \frac{1}{P_\lambda(0; q, t) P_\mu(0; q, t)} \sum_{\nu \in \mathcal{P}} \left(m^\nu_{\lambda,\mu}(q,t) P_\nu(0; q, t) \right) \hat{P}_\nu(x; q, t), \tag{2.5}$$

i.e. products of the $\hat{P}$ satisfy a similar relation with structure coefficients

$$\hat{m}^\nu_{\lambda,\mu}(q,t) := \frac{P_\nu(0; q, t)}{P_\lambda(0; q, t) P_\mu(0; q, t)} m^\nu_{\lambda,\mu}(q,t). \tag{2.6}$$

Remark 2.2.4. Given a discrete set $\Omega = \{\omega_1, \ldots\}$, probability measures on Ω correspond to formal sums $\sum_{\omega \in \Omega} a_i \omega_i$ with all a_i nonnegative and $\sum_i a_i = 1$. By (2.3) with all x_i set to 0 it follows that $\sum_{\nu \in \mathcal{P}} \hat{m}^\nu_{\lambda,\mu} = 1$. Hence, if for given real values of q and t we have that $\hat{m}^\nu_{\lambda,\mu}$ is nonnegative[1] for each ν, then the expansion $\sum_{\nu \in \mathcal{P}} \hat{m}^\nu_{\lambda,\mu} \hat{P}_\nu$ defines a probability measure on $\mathcal{P}$ given by $\Pr(\nu) = \hat{m}^\nu_{\lambda,\mu}$. In such a situation, we say that $\sum_{\nu \in \mathcal{P}} \hat{m}^\nu_{\lambda,\mu} \hat{P}_\nu$ *represents* the probability measure. We may similarly represent any probability measure on the set of partitions of length at most n by a linear combination of $\hat{P}_\nu$ with nonnegative coefficients summing to 1. A single polynomial $\hat{P}_\lambda$ represents a point mass on that partition, and the product of two polynomials corresponds to a kind of convolution of the measures. Some similarity with Question 2.2.1 and Question 2.2.2 is now apparent, as in these cases one is in some sense looking at how matrix multiplication or addition mixes the cokernel partitions/eigenvalues of a pair of random matrices

[1] The nonnegativity of the coefficients $\hat{m}^\nu_{\lambda,\mu}$ for $0 < q, t < 1$ is an open conjecture, and a short discussion of current work on it may be found in [29].

constrained so that the distributions of their cokernel partitions/eigenvalues are point masses on a specified partition or tuple of real numbers.

To make this analogy concrete in the case of Question 2.2.2 requires working with certain limiting degenerations of Macdonald polynomials.

Definition 2.2.5. For $\theta > 0$, the Jack polynomial J_λ is given by

$$J_\lambda(x_1, \ldots, x_n; \theta) := \lim_{q \to 1} P_\lambda(x_1, \ldots, x_n; q, q^\theta). \tag{2.7}$$

The functions actually needed to answer Question 2.2.2 are then derived from these.

Definition 2.2.6. Let $\alpha = (\alpha_1, \ldots, \alpha_n)$ be a decreasing n-tuple of distinct real numbers, and for every $\epsilon > 0$ let $\lambda(\epsilon) = (\lfloor \frac{\alpha_1}{\epsilon} \rfloor, \ldots, \lfloor \frac{\alpha_n}{\epsilon} \rfloor)$ and $x_i(\epsilon) = e^{\epsilon z_i}$. Then the *multivariate Bessel function* B_α is a function of n arguments $z_1, \ldots, z_n$ and one parameter θ given by

$$B_\alpha(z_1, \ldots, z_n; \theta) := \lim_{\epsilon \to 0} \epsilon^{\theta \frac{N(N-1)}{2}} J_{\lambda(\epsilon)}(x_1(\epsilon), \ldots, x_n(\epsilon); \theta). \tag{2.8}$$

We further define the normalized multivariate Bessel function

$$\hat{B}_\alpha(z; \theta) := \frac{B_\alpha(z; \theta)}{B_\alpha(0, \ldots, 0; \theta)} \tag{2.9}$$

We refer to [29] for details on these functions. The multivariate Bessel functions satisfy a version of (2.3) with the sum replaced by an integral: For any α, β as in Definition 2.2.6, we have

$$\hat{B}_\alpha(z; \theta) \cdot \hat{B}_\beta(z; \theta) = \int_\gamma b^\gamma_{\alpha,\beta}(\theta) \hat{B}_\gamma(z; \theta) d\gamma \tag{2.10}$$

for some function $b^\gamma_{\alpha,\beta}(\theta)$ of γ and θ, where the integral is over the space in $\mathbb{R}^n$ of all n-tuples γ as in Definition 2.2.6. As in the Macdonald case, the normalization of $\hat{B}$ ensures that $b^\gamma_{\alpha,\beta}(\theta)$ integrates to 1, so (again assuming nonnegativity, see previous footnote) $b^\gamma_{\alpha,\beta}(\theta) d\gamma$ defines a probability measure on the set of decreasing n-tuples γ.

Proposition 2.2.7. *Let α, β be decreasing n-tuples of distinct real numbers. If A, B are drawn randomly from the GOE conditional on having eigenvalues α and β, then the distribution of the eigenvalues γ of $A + B$ is given by the probability measure $d\mu = b^\gamma_{\alpha,\beta}(\theta) d\gamma$ with $\theta = \frac{1}{2}$, where $d\gamma$ is the Lebesgue measure on $\mathbb{R}^n$. The same claim holds for the GUE and GSE with $\theta = 1$ and $\theta = 2$ respectively.*

This answers Question 2.2.2; the above and related results may be found in [29]. The answer to Question 2.2.1 has two main parts: the first is expressing the desired probabilities in terms of counting submodules via Lemma 2.2.9. The second is relating these counting problems to the Hall-Littlewood polynomials, another specialization of Macdonald polynomials, at which point the similarity to Proposition 2.2.7 becomes apparent.

Notation 2.2.8. Let λ be a partition of length at most n. We denote by ω^λ the $n \times n$ diagonal matrix with ii^{th} entry ω^{λ_i} (where λ is padded with zeros as necessary to obtain a decreasing n-tuple).

Lemma 2.2.9. *Let λ, μ be two partitions of length $\leq n$ and R, ω, q as in the previous section. For any partition α, let N_α be the number of distinct R-submodules $N \subset R^n$ such that $R^n/N \cong M_\alpha$. Finally, let $g^\nu_{\lambda\mu}$ be the number of submodules $N \subset M_\nu$ such that $N \cong M_\mu$ and $M_\nu/N \cong M_\lambda$. Then if A and B are chosen from $M_n(R)$ with additive Haar measure, we have*

$$\Pr(\operatorname{coker}(AB) \cong M_\nu | \operatorname{coker}(A) \cong M_\lambda, \operatorname{coker}(B) \cong M_\mu) = \frac{N_\nu}{N_\lambda \cdot N_\mu} g^\nu_{\lambda\mu} \tag{2.11}$$

for all partitions ν of length $\leq n$.

Proof. Let $G = \mathrm{GL}_n(R)$. By Smith normal form (Proposition 1.1.6), cokernels are in bijection with double cosets of $G \backslash M_n(R)/G$:

$$\{A \in M_n(R) : \operatorname{coker}(A) \cong M_\lambda\} = G\omega^\lambda G. \tag{2.12}$$

Since multiplication by G preserves the additive Haar measure, integrating a function of $A \in G\omega^\mu G$ with the respect to the additive Haar measure, conditioned to live on $G\omega^\mu G$ and normalized to 1, is the same as integrating the same function of $g\omega^\mu h$ over $(g, h) \in G \times G$ with the (multiplicative) Haar probability measure on G. Letting $c_\nu(A)$ be the indicator function $\mathbb{1}_{A \in G\omega^\nu G}$ (note it is invariant under left- and right-multiplication of the argument by G), we have

$$\Pr(\lambda(AB) = \nu | \lambda(A) = \lambda, \lambda(B) = \mu) = \int_{(g,h,g',h') \in G^4} c_\nu(g\omega^\lambda h g' \omega^\mu h'). \tag{2.13}$$

Choose representatives $x_i \in M_n(R)$ so that $G\omega^\lambda G = \bigsqcup_i x_i G$, and representatives y_j so $G\omega^\mu G = \bigsqcup_j y_j G$. These representatives x_i are in bijection with the submodules $V \subset R^n$ such that $R^n/V \cong M_\lambda$ given by their image as a map $R^n \to R^n$, so there are N_μ of them and similarly

are N_λ of the representatives y_j. Thus

$$\int_{(g,h,g',h')\in G^4} c_\nu(g\omega^\lambda hg'\omega^\mu h') = \frac{1}{N_\lambda}\sum_i \int_{(k,g',h')\in G^3} c_\nu(x_i k g'\omega^\mu h') \tag{2.14}$$

$$= \frac{1}{N_\lambda}\sum_i \int_{(k',h')\in G^2} c_\nu(x_i k'\omega^\mu h') \tag{2.15}$$

$$= \frac{1}{N_\lambda N_\mu}\sum_{i,j} \int_{h''\in G} c_\nu(x_i y_j h'') \tag{2.16}$$

$$= \frac{1}{N_\lambda N_\mu}\sum_{i,j} c_\nu(x_i y_j) \tag{2.17}$$

here using the facts that the product of Haar-distributed elements is Haar-distributed, and that c_ν is a function on double cosets.

Hence it suffices to show that

$$|\{(i,j) : R^n/x_i y_j R^n \cong M_\nu\}| = N_\nu \cdot g^\nu_{\lambda\mu} \tag{2.18}$$

Fix a submodule $V_\nu \subset R^n$ with $R^n/V_\nu \cong M_\nu$. The submodules of R^n/V_ν of type μ with quotient of type λ are in bijection with the submodules $V_{\lambda,\mu} \subset R^n$ such that (1) $V_{\lambda,\mu} \supset V_\nu$, (2) $V_{\lambda,\mu}/V_\nu \cong M_\mu$, and (3) $R^n/V_{\lambda,\mu} \cong M_\lambda$. Each such module $V_{\lambda,\mu}$ is the image $x_i R^n$ of a unique x_i, since these parametrize submodules of R^n with quotient M_λ. Furthermore, since $x_i R^n/V_\nu \cong M_\mu$, then $x_i^{-1}V_\nu \subset R^n$ and $R^n/x_i^{-1}V_\nu \cong M_\mu$. Hence $x_i^{-1}V_\nu = y_j R^n$ for a unique x_i. It follows that $V_\nu = x_i y_j R^n$.

Hence for each submodule V_ν, the submodules $V_{\lambda,\mu}$ defined above are in bijection with pairs (i,j) such that $x_i y_j R^n = V_\nu$. The claim follows immediately. □

Definition 2.2.10. The Hall-Littlewood polynomial is a symmetric polynomial in n variables associated to a partition λ of length $\leq n$ and given by $H_\lambda(x_1,\ldots,x_n;t) = P_\lambda(x_1,\ldots,x_n;t,q=0)$. Alternatively, it is given by the explicit formula

$$H_\lambda = \frac{1}{v_\lambda(t)}\sum_{\sigma\in S_n} \sigma\left(x_1^{\lambda_1}\cdots x_n^{\lambda_n}\prod_{1\leq i<j\leq n}\frac{x_i - tx_j}{x_i - x_j}\right) \tag{2.19}$$

where σ acts by permuting the variables and $v_\lambda(t) = \prod_{i\geq 0}\prod_{j=1}^{m_i(\lambda)}\frac{1-t^j}{1-t}$.

It is not immediately clear that (2.19) actually yields a polynomial, but this is true since the Vandermonde determinant divides any alternating polynomial. The division by $v_\lambda(t)$ has the effect of clearing common factors of the coefficients of monomial terms in the expansion of H_λ, so in fact the GCD of these coefficients is 1.

For reasons that will soon become clear, it makes more sense to work not in the ring of symmetric polynomials but in the ring of symmetric functions.

Definition 2.2.11. For every $n \geq 1$, let $\Lambda_n = \mathbb{Z}[x_1, \ldots, x_n]^{S_n}$ be the ring of symmetric polynomials in n variables with integer coefficients, and Λ_n^k be the sub-$\mathbb{Z}$-module of n^{th} degree polynomials. Then for each $m > n$, we have a map $\rho_{m,n} : \Lambda_m^k \to \Lambda_n^k$ given by setting $x_{n+1}, \ldots, x_m$ to 0. Let Λ^k be the inverse limit of Λ_n^k with respect to the maps $\rho_{m,n}$. The *ring of symmetric functions* is

$$\Lambda = \bigoplus_{k \geq 0} \Lambda^k \tag{2.20}$$

with the natural ring structure.

Remark 2.2.12. The reason Λ is constructed by taking inverse limits for the module of symmetric polynomials of each degree and then taking a direct sum is to disallow 'symmetric functions' with monomials of arbitrarily high degree, such as the geometric series expansion of $\prod_{i \geq 1} \frac{1}{1-x_i}$, which would be part of Λ if the inverse limit were defined using the rings Λ_n.

For a given λ of length $\leq n$, one can check from the explicit formula of Definition 2.2.10 that $H_\lambda(x_1, \ldots, x_n, 0; t) = H_\lambda(x_1, \ldots, x_n; t)$. Hence there exists an element of Λ, denoted $H_\lambda(x_1, \ldots; t)$, which specializes to $H_\lambda(x_1, \ldots, x_n; t)$ upon setting $x_{n+1}, x_{n+2}, \ldots$ to 0. We refer to $H_\lambda(x_1, \ldots; t)$ as the *Hall-Littlewood function*. The following useful property of Hall-Littlewood symmetric functions, which connects them to the Cohen-Lenstra measure, comes from [36, III.2, Example 2].

Proposition 2.2.13. *For any $\lambda \in \mathcal{P}$ and any $q > 1$, specializing $t = q^{-1}$ and $x_i = q^{-i}$ yields*

$$H_\lambda(q^{-1}, q^{-2}, \ldots; q^{-1}) = \frac{q^{n(\lambda)}}{|\operatorname{Aut}_q(\lambda)|} \tag{2.21}$$

where M_λ is an R-module of type λ.

Because the above specialization involves setting infinitely many variables to nonzero values, it requires working with Hall-Littlewood functions rather than Hall-Littlewood polynomials.

Definition 2.2.14. Given R, q as in the previous section, the associated *Hall algebra* $H(R)$ is the $\mathbb{Z}$-algebra whose underlying $\mathbb{Z}$-module is free with generators u_λ for every partition λ, and whose multiplication is given by

$$u_\lambda \cdot u_\mu = \sum_{\nu \in \mathcal{P}} g_{\lambda\mu}^\nu u_\nu \tag{2.22}$$

where $g_{\lambda\mu}^\nu$ is defined as in Lemma 2.2.9.

The identity of the Hall algebra is $u_{()}$. For associativity of its multiplication it is easy to check that the coefficient of u_α in the expansion of $u_\lambda u_\mu u_\nu$ (multiplied in either order) is the number of chains of R-modules $0 \subset N_1 \subset N_2 \subset M_\alpha$ with $N_1 \cong M_\lambda, N_2/N_1 \cong M_\mu$, and $M_\alpha/N_2 \cong u_\nu$.

One of the historical motivations for the theory of Hall-Littlewood functions was that their multiplicative structure mimics that of the Hall algebra.

Definition 2.2.15. For any partition λ, we let $n(\lambda) = \sum_{i\geq 1}(i-1)\lambda_i = \sum_{i\geq 1}\binom{\lambda_i'}{2}$ (it is a straightforward calculation that the latter two are the same).

Proposition 2.2.16. *There is a ring isomorphism $\psi : H(R) \otimes_{\mathbb{Z}} \mathbb{Q} \to \Lambda$ defined by $\psi(u_\lambda) = q^{-n(\lambda)}P_\lambda(x; q^{-1})$.*

Finally we are in a position to answer Question 2.2.1. Similarly to our normalizations for the Macdonald and multivariate Bessel functions, we define the *normalized Hall-Littlewood function* by $\hat{H}_\lambda(x_1, \ldots; q^{-1}) = \frac{H_\lambda(x_1,\ldots;q^{-1})}{H_\lambda(q^{-1},q^{-2},\ldots;q^{-1})}$. We define the structure coefficients $\hat{c}^\nu_{\lambda\mu}$ as before by

$$\hat{H}_\lambda \cdot \hat{H}_\mu = \sum_{\nu\in\mathcal{P}} \hat{c}^\nu_{\lambda\mu}\hat{H}_\nu. \tag{2.23}$$

Note that for the denominator in the definition of $\hat{H}$ we have not set all variables to 0 as before, but to powers of q; as mentioned in the discussion on Macdonald polynomials, provided that nonnegativity of structure coefficients holds, any choice of values for the variables will yield a probability measure. This nonnegativity is assured by Proposition 2.2.16, which relates the non-normalized Hall-Littlewood polynomials to the Hall algebra (which has nonnegative structure coefficients), and Proposition 2.2.13, which ensures that the normalizing factors $H_\lambda(q^{-1}, q^{-2}, \ldots; q^{-1})$ are positive.

Theorem 2.2.17. *Let R be the ring of integers of a local field, with residue field of size q, $n \in \mathbb{N}$ and λ, μ, ν partitions of length $\leq n$. Then if A and B are independent random elements of $M_n(R)$ distributed according to the additive Haar measure, we have*

$$\Pr(\lambda(AB) = \nu | \lambda(A) = \lambda, \lambda(B) = \mu) = \frac{(q^{-1})_{n-\ell(\lambda)}(q^{-1})_{n-\ell(\mu)}}{(q^{-1})_{n-\ell(\nu)}(q^{-1})_n}\hat{c}^\nu_{\lambda\mu}. \tag{2.24}$$

In particular, let $\Pr_n$ be the probability measure on partitions assigning to a partition ν the probability given by the RHS of the above equation. Then $\lim_{n\to\infty}\Pr_n(\nu) = \hat{c}^\nu_{\lambda\mu}$, so $\Pr_n$ converges pointwise to the measure which is represented by the product $\hat{H}_\lambda(x; q^{-1}) \cdot \hat{H}_\mu(x; q^{-1})$ in the sense of Remark 2.2.4.

Proof. It suffices by Lemma 2.2.9 to show that

$$\frac{N_\nu}{N_\lambda \cdot N_\mu} g^\nu_{\lambda\mu} = \frac{(q^{-1})_{n-\ell(\lambda)}(q^{-1})_{n-\ell(\mu)}}{(q^{-1})_{n-\ell(\nu)}(q^{-1})_n} \hat{c}^\nu_{\lambda\mu}. \tag{2.25}$$

It follows from Proposition 2.2.16 and the definition of the normalized Hall-Littlewood functions that

$$g^\nu_{\lambda\mu} = q^{n(\nu)-n(\lambda)-n(\mu)} \frac{H_\lambda(q^{-1}, q^{-2}, \ldots; q^{-1}) H_\mu(q^{-1}, q^{-2}, \ldots; q^{-1})}{H_\nu(q^{-1}, q^{-2}, \ldots; q^{-1})} \hat{c}^\nu_{\lambda\mu}. \tag{2.26}$$

By Proposition 2.2.13 this is equal to $\frac{|\operatorname{Aut}_q(\nu)|}{|\operatorname{Aut}_q(\lambda)| \cdot |\operatorname{Aut}_q(\mu)|} \hat{c}^\nu_{\lambda\mu}$. Hence

$$\Pr(\operatorname{coker}(AB) \cong M_\nu) = \frac{N_\nu \cdot |\operatorname{Aut}_q(\nu)|}{(N_\lambda \cdot |\operatorname{Aut}_q(\lambda)|) \cdot (N_\mu \cdot |\operatorname{Aut}_q(\mu)|} \hat{c}^\nu_{\lambda\mu} \tag{2.27}$$

We compute N_α for general partitions α; the result was known going back at least to [10], but it is worth seeing the computation in detail.

Any surjection $R^n \to M_\alpha$ has kernel a submodule of R^n with quotient isomorphic to M_α, and any two such surjections are related by composition with an automorphism of M_α. Hence $N_\alpha = \frac{|\{\phi : R^n \twoheadrightarrow M_\alpha\}|}{|\operatorname{Aut}(M_\alpha)|}$, so we compute the number of surjections. A map $(R/\omega)^n \to M_\alpha/\omega M_\alpha$ is determined by where it sends the standard basis vectors, so there are $|M_\alpha|^n = q^{n|\alpha|}$ such maps. Each one is surjective if and only if the corresponding map of vector spaces $(R/\omega)^n \to M_\alpha/\omega M_\alpha$ is surjective: clearly surjectivity of the former implies the latter, and viewing elements of R as power series in ω by Lemma 1.2.8, the reverse implication follows. Hence the proportion of the maps which are surjective is exactly the proportion of elements of $M_{\ell(\alpha)\times n}(\mathbb{F}_q)$ with rank $\ell(\alpha)$. Each full-rank element is just an $\ell(\alpha)$-tuple of linearly independent elements of $\mathbb{F}_q^n$. To choose such a tuple, there are $q^n - 1$ choices for the first vector (since it must be nonzero), then $q^n - q$ choices for the second since it cannot be a multiple of the first, etc., yielding

$$N_\alpha \cdot |\operatorname{Aut}(M_\alpha)| = |\{\phi : R^n \twoheadrightarrow M_\alpha\}| = q^{n\cdot|\alpha|} \frac{1}{q^{n\cdot\ell(\alpha)}} (q^n - 1) \cdots (q^n - q^{\ell(\alpha)-1}) = q^{n\cdot|\alpha|} \frac{(q^{-1})_n}{(q^{-1})_{\ell(\alpha)}}. \tag{2.28}$$

If $g^\nu_{\lambda\mu} = 0$ then $\hat{c}^\nu_{\lambda\mu} = 0$ as well by Proposition 2.2.16, so assume this is not the case. Then there is some $N \cong M_\mu$ with $M_\nu/N \cong M_\lambda$, so $q^{|\nu|} = |M_\nu| = |M_\lambda| \cdot |M_\mu| = q^{|\lambda|+|\mu|}$, so $|\lambda| + |\mu| = |\nu|$. Hence the $q^{n\cdot|\alpha|}$ factors (for $\alpha = \lambda, \mu, \nu$ appearing when (2.28) is substituted into (2.27) cancel one another, and we are left with $\frac{(q^{-1})_{n-\ell(\lambda)}(q^{-1})_{n-\ell(\mu)}}{(q^{-1})_{n-\ell(\nu)}(q^{-1})_n} \hat{c}^\nu_{\lambda\mu}$, showing (2.24).

The second part of the theorem follows since $\frac{(q^{-1})_{n-\ell(\lambda)}(q^{-1})_{n-\ell(\mu)}}{(q^{-1})_{n-\ell(\nu)}(q^{-1})_n} \to 1$ as $n \to \infty$. □

This shows that, in the limit, the way that multiplying two random matrices mixes their cokernels is described by the multiplication of corresponding Hall-Littlewood polynomials. One

interesting contrast between Theorem 2.2.17 and Proposition 2.2.7 is that in the latter, the measure given by the eigenvalues of $A+B$ is exactly described by products of multivariate Bessel functions, while in Theorem 2.2.17 the corresponding statement is true only asymptotically.

Remark 2.2.18. It is very easy to show by the above argument that if A is a random matrix with Haar distribution conditioned on having cokernel partition λ, then $\mathrm{coker}(A^2)$ has the same distribution as the one computed in Theorem 2.2.17. A very natural question is to compute the distribution of $\mathrm{coker}(A^n)$, or the joint distribution of $\mathrm{coker}(A), \mathrm{coker}(A^2), \ldots, \mathrm{coker}(A^n)$. However, it is no longer true that $\mathrm{coker}(A^n)$ has the same distribution as $\mathrm{coker}(A_1 \cdot A_2 \cdots A_n)$ with the A_i iid with the same distribution as A. Consequently, the simple method used to compute the distribution in Theorem 2.2.17 seems unusable without serious modification, but the approach of Section 3.3 is somewhat motivated by this problem.

Chapter 3

Random matrices over finite fields and integers of local fields

3.1 Random matrices over finite fields

As with p-adic random matrix theory, random matrix theory over finite fields came later than classical random matrix theory and has some analogies with the latter. After Wigner's original Gaussian unitary, orthogonal, and symplectic ensembles, Dyson [14] shifted focus to what he called the circular unitary ensemble (CUE), the local eigenvalue statistics of which reflect those of the GUE. This CUE is none other than the unitary group $U(n)$ with Haar measure, which is finite because $U(n)$ is compact, and hence may be normalized to a probability measure. This ensemble, and the related ones obtained by putting a Haar probability measure on any compact classical group, are interesting partly for their importance in number theory alluded to in the Introduction: conjecturally, the low-lying zeros of L-functions chosen randomly from any reasonably-defined family have the same distribution as the eigenvalues near 1 of matrices drawn Haar-randomly from one of the compact classical groups. [33] and [11] are good surveys, and a more detailed version of the previous discussion can be found in [7].

Because every matrix in $U(n)$ is conjugate by a matrix in $U(n)$ to a diagonal matrix, the study of the eigenvalues of a random matrix in $U(n)$ is the same as the study of the conjugacy class of a random element in this group. Random matrix theory over finite fields is, analogously, largely concerned with data coming from the conjugacy class of matrices chosen uniformly randomly from a group of Lie type over $\mathbb{F}_q$. It thus draws much more from finite group theory, and

some interest in studying it comes from computational group theory and analysis of certain algorithms, see [24, 2.2]. One of its most surprising connections is a probabilistic proof and interpretation of the Rogers-Ramanujan identities [23]. What concerns the present work most directly, however, is that the Cohen-Lenstra distribution appears in this context as well, a fact only noticed long after the Cohen-Lenstra heuristics were formulated. In this section we will show how to extract random partitions from the rational canonical form of a random matrix over a finite field, and that the induced probability measure on such partitions is given by a generalization of the Cohen-Lenstra measure. The definitions, results and proofs of this section are generally modifications of the ones in [24]; the alteration which was least obvious to the author is that n must be chosen differently in Definition 3.1.6 than in [24], but once this change is made the rest of the arguments follow with some minor differences.

The first task is to parametrize orbits by means of rational canonical form, which associates integer partitions to irreducible factors of the characteristic polynomial of a matrix.

Recall that the *companion matrix* of a polynomial $f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_0$ is

$$C(f) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{pmatrix}.$$

The characteristic polynomial of this matrix is $f(x)$.

Proposition 3.1.1 (Rational canonical form)**.** *Let K be a field and $A \in M_n(K)$, with characteristic polynomial $f(x) = f_1^{e_1} \cdots f_r^{e_r}$ where the f_i are distinct irreducible factors of degrees d_i. Then there is a matrix $P \in \mathrm{GL}_n(K)$ such that*

$$PAP^{-1} = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & A_r \end{pmatrix} \tag{3.1}$$

where

$$A_i = \begin{pmatrix} C(f_i^{\lambda_1^{(f_i)}}) & 0 & \cdots & 0 \\ 0 & C(f_i^{\lambda_2^{(f_i)}}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & C(f_i^{\lambda_{j_i}^{(f_i)}}) \end{pmatrix} \tag{3.2}$$

for a partition $\lambda^{(f_i)} = (\lambda_1^{(f_i)}, \ldots, \lambda_{j_i}^{(f_i)})$ *of* e_i*. Furthermore,* A *is conjugate to a unique matrix of the above form.*

Proof. Apply the structure theorem to the PID $K[x]$ acting on the module K^n where x acts by A, then choose an appropriate basis. □

Notation 3.1.2. Given a matrix $A \in M_n(\mathbb{F}_q)$ and a monic irreducible nonconstant polynomial f dividing its characteristic polynomial, we will refer to the partition given by the blocks corresponding to f in the rational canonical of A as in Proposition 3.1.1 by $\lambda^{(f)}(A)$. We will sometimes write $\lambda^{(f)}$ when A is clear.

There are several different versions of rational canonical form listed in the literature corresponding to different choices of basis for the summands in the structure theorem; the above convention is used in [24] but not in the original [42] or [34]. In any case, this yields that the conjugacy class of a matrix is determined by a finite set of irreducible polynomials with a partition associated to each one, subject to the condition

$$\sum_i \deg(f_i)|\lambda^{(f_i)}| = n \tag{3.3}$$

where n is the matrix size and the sum is over irreducible polynomials. Conversely, any such set of irreducible polynomials over $\mathbb{F}$ with associated partitions determines a unique conjugacy class. Fixing a polynomial $f \in K[x]$, we may thus associate to any matrix in $M_n(K)$ the corresponding corresponding partition λ (possibly the empty partition) from Proposition 3.1.1. Hence given a random matrix ensemble, we obtain a measure on the set of integer partitions–which, if preferred, may be thought of as a partition-valued random variable.

If K is a field such as $\mathbb{R}$ or $\mathbb{Q}_p$, then any individual polynomial f will occur as a factor of the characteristic polynomial of a generic matrix with probability 0, so we obtain an extremely uninteresting measure on partitions which assigns probability 1 to the empty partition. However, if K is a finite field, then for any reasonable random matrix ensemble over K each irreducible polynomial occurs often enough as a factor of the characteristic polynomial that we obtain a

nontrivial measure on partitions. In the case of $M_n(\mathbb{F}_q)$, this measure turns out to be the Cohen-Lenstra measure.

The main piece of machinery in the proof is that of cycle indices and their generating functions, used in [42] and [24]. The following definition is general, but for the moment we will only use the case $S = M_n(\mathbb{F}_q)$.

Definition 3.1.3. Given a subset $S \subseteq M_n(\mathbb{F}_q)$, define its *cycle index* to be

$$Z_S := \frac{1}{|S|} \sum_{A \in S} \prod_{f : |\lambda^{(f)}(A)| > 0} x_{f,\lambda^{(f)}}, \tag{3.4}$$

where the $x_{f,\lambda^{(f)}}$ are formal variables associated to each pair of an irreducible monic nonconstant polynomial and a partition.

Common choices of S in the literature are $\mathrm{GL}_n(\mathbb{F}_q)$ [24], or another group of Lie type over $\mathbb{F}_q$ such as $U_n(\mathbb{F}_q), O_n(\mathbb{F}_q)$, or $\mathrm{Sp}_{2n}(\mathbb{F}_q)$ [25]. As mentioned at the beginning of this section, we are interested in the case of $M_n(\mathbb{F}_q)$.

Remark 3.1.4. Cycle indices provide an algebraic way to extract a measure on partitions from a measure on the set S. Recall from Remark 2.2.4 that given a discrete set $\Omega = \{\omega_1, \ldots\}$, probability measures on Ω correspond to formal sums $\sum_{\omega \in \Omega} a_i \omega_i$ with all a_i nonnegative and $\sum_i a_i = 1$. We have already used this correspondence in Section 2.2 to represent measures on $\mathcal{P}$, with the normalized Hall-Littlewood polynomial $\hat{H}_\lambda$ playing the role of the formal symbol corresponding to λ.

Since $\prod_{f : |\lambda^{(f)}(A)| > 0} x_{f,\lambda^{(f)}}$ depends only on the conjugacy class of A, the cycle index $Z_S = \frac{1}{|S|} \sum_{A \in S} \prod_{f : |\lambda^{(f)}| > 0} x_{f,\lambda^{(f)}}$ thus represents the measure on conjugacy classes of G which weights each one according to its size, where we identify a product of $x_{f,\lambda^{(f)}}$ with a conjugacy class via Proposition 3.1.1. Fix n and an irreducible monic nonconstant polynomial f, and set all $x_{h,\lambda^{(h)}}$ to 1 for all polynomials $h \neq f$. Then Z_G becomes an element of $\bigoplus_{\lambda \in \mathcal{P}} \mathbb{R} x_{f,\lambda}$ with all coefficients nonnegative and summing to 1, defining a probability measure on $\mathcal{P}$ which is easily seen to be the measure on partitions described after Proposition 3.1.1. It is clear that if the coefficient of $x_{f,\lambda}$ converges to a fixed value for each $x_{f,\lambda}$, this is equivalent to the corresponding measures converging pointwise[1] to the measure defined by the limits of each coefficient for each $x_{f,\lambda}$. In general, these limits of coefficients may not define a measure (e.g. they may all be 0), but they will define probability measures in the cases considered here.

[1] By a sequence of measures $(\mu_n)_{n \geq 1}$ on $\mathcal{P}$ converging pointwise to μ, we mean that $\mu_n(\{\lambda\}) \to \mu(\{\lambda\})$ for each $\lambda \in \mathcal{P}$.

Remark 3.1.5. The normalization of our cycle indices differs from that in [24] or [42], which always divide by $\frac{1}{|\mathrm{GL}_n(\mathbb{F}_q)|}$ instead of $\frac{1}{|S|}$, where in our case $S = M_n(\mathbb{F}_q)$. The former has the advantage that the factor divides out in orbit-stabilizer arguments (since we care about the action of $\mathrm{GL}_n(\mathbb{F}_q)$, regardless of S), but our convention has the advantage that Z_S represents a probability measure on the set of orbits, in the sense of the previous remark.

It is also possible to represent more complicated measures through series of cycle indices.

Definition 3.1.6. Fix a monic irreducible nonconstant polynomial f, and let $0 < u < 1$. We define $\mu_{u,q,f}$ to be the probability measure on $\mathcal{P}$ giving the distribution of partitions chosen by the following procedure:

1. Choose a nonnegative integer n with probability $(u;q^{-1})_\infty \frac{u^n}{(q^{-1})_n}$, where $(u;q^{-1})_\infty$ is as in Notation 1.3.6.

2. If $n = 0$, take the empty partition. Otherwise, choose a uniformly random element A from $M_n(\mathbb{F}_q)$ and take the partition $\lambda^{(f)}(A)$ defined in Proposition 3.1.1.

We defer the proof that the measure on $\mathbb{Z}_{\geq 0}$ defined in Step 1 above is actually a probability measure until Lemma 3.1.9.

As per Remark 3.1.4, $\mu_{u,q,f}$ is represented by setting $x_{h,\lambda} = 1$ for all $\lambda \in \mathcal{P}, h \neq f$ in the generating function-like expression

$$(u;q^{-1})_\infty \sum_{n\geq 0} \frac{u^n}{(q^{-1})_n} Z_{M_n(\mathbb{F}_q)} \tag{3.5}$$

where by convention $Z_{M_0(\mathbb{F}_q)} = 1$.

In Theorem 3.1.10 we will use nice factorization properties of (3.5) show that $\mu_{q,u,f}$ yields a generalization of the Cohen-Lenstra measure. The following key lemma makes the connection between orbits of $M_n(\mathbb{F}_q)$ under conjugation by $\mathrm{GL}_n(\mathbb{F}_q)$ and automorphisms of R-modules, which is key to the appearance of a Cohen-Lenstra-like measure.

Lemma 3.1.7. *Let C be an orbit in $M_n(\mathbb{F}_q)$ determined by polynomials $\{f_1, \ldots, f_m\}$ and associated partitions $\lambda^{(f_i)}, 1 \leq i \leq m$ as in Proposition 3.1.1. Then the size of the stabilizer of an element A of C is*

$$\prod_{i=1}^{m} |\operatorname{Aut}_{q^{\deg(f_i)}}(\lambda^{(f_i)})| \tag{3.6}$$

Proof. By the definition of the polynomials $\{f_1, \ldots, f_m\}$ and associated partitions from Proposition 3.1.1 we have that, viewing $M := \mathbb{F}_q^n$ as an $\mathbb{F}_q[x]$-module with x acting by A, then $M \cong \bigoplus_{i=1}^m \bigoplus_{j\geq 1} \mathbb{F}_q[x]/(f_i(x)^{\lambda_j^{(f_i)}})$ as $\mathbb{F}_q[x]$-modules. Any element $g \in \mathrm{GL}_n(\mathbb{F}_q)$ acts on $\mathbb{F}_q^n$, and if g stabilizes A under conjugation then it commutes with A, hence its action on $\mathbb{F}_q^n$ commutes with that of A and it yields an $\mathbb{F}_q[x]$-module isomorphism. Clearly, distinct matrices g, g' yield distinct isomorphisms. Conversely, any $\mathbb{F}_q[x]$-module isomorphism is determined by its action on the standard basis of $\mathbb{F}_q^n$, hence given by an invertible matrix, and this matrix commutes with A. Hence

$$|\mathrm{Stab}(A)| = \left| \mathrm{Aut}_{\mathbb{F}_q[x]} \left(\bigoplus_{i=1}^m \bigoplus_{j\geq 1} \mathbb{F}_q[x]/(f_i(x)^{\lambda_j^{(f_i)}}) \right) \right|. \tag{3.7}$$

We have

$$\mathrm{Aut} \left(\bigoplus_{i=1}^m \bigoplus_{j\geq 1} \mathbb{F}_q[x]/(f_i(x)^{\lambda_j^{(f_i)}}) \right) = \prod_{i=1}^m \mathrm{Aut} \left(\bigoplus_{j\geq 1} \mathbb{F}_q[x]/(f_i(x)^{\lambda_j^{(f_i)}}) \right) \tag{3.8}$$

because the there are no nontrivial maps between the summands corresponding to each i. Given any h not a multiple of f_i, the action of $h(x)$ on $\bigoplus_{j\geq 1} \mathbb{F}_q[x]/(f_i(x)^{\lambda_j^{(f_i)}})$ is invertible. Hence $\bigoplus_{j\geq 1} \mathbb{F}_q[x]/(f_i(x)^{\lambda_j^{(f_i)}})$ naturally has the structure of a module over the localization $\mathbb{F}_q[x]_{f_i(x)}$, and the automorphisms are the same. Since $\mathbb{F}_q[x]_{f_i(x)}$ is a local ring with maximal ideal $f_i(x)$, the size of its residue field is $q^{\deg(f_i)}$, so

$$\left| \mathrm{Aut}_{\mathbb{F}_q[x]} \left(\bigoplus_{j\geq 1} \mathbb{F}_q[x]/(f_i(x)^{\lambda_j^{(f_i)}}) \right) \right| = |\operatorname{Aut}_{q^{\deg(f_i)}}(\lambda_j^{(f_i)})|. \tag{3.9}$$

Combining this with (3.7) and (3.8) completes the proof. □

We now define the generalized Cohen-Lenstra measure alluded to earlier.

Definition 3.1.8. For $0 < u \leq 1$ and $q > 1$, define the *u-measure* $\mu_{u,q}$ on the set of all partitions $\mathcal{P}$ by

$$\mu_{u,q}(\lambda) = \left(\prod_{i\geq 1} (1 - uq^{-i}) \right) \frac{u^{|\lambda|}}{|\operatorname{Aut}_q(\lambda))|}. \tag{3.10}$$

When $u = 1$, this is just the measure μ_q of Definition 1.3.10. Note that the index of the product starts at 1 rather than 0 in the case of $(u;q)_\infty$. We verify that the measures we have defined are in fact probability measures; aside from being necessary for the discussion, the algebraic fact that these measures have total mass 1 will be used in the proof of Theorem 3.1.10 to come. It is worth noting as an aside that the measure $\mu_{u,q}$ has many surprising properties and connections not touched upon in this thesis, for which a good source is [20].

Lemma 3.1.9. *For $q > 1$ and $0 < u < q$, $\mu_{u,q}$ is a probability measure on $\mathcal{P}$. Additionally, for $0 < u < 1$, q any prime power and f any monic irreducible nonconstant polynomial, $\mu_{u,q,f}$ is a probability measure.*

Proof. It is clear that the formula in Definition 3.1.8 is nonnegative, so we must show the total mass is 1, i.e.

$$\sum_{n\geq 0}\sum_{\lambda\vdash n}\frac{u^n}{|\operatorname{Aut}_q(\lambda))|}=\prod_{i\geq 1}\frac{1}{1-uq^{-i}}. \tag{3.11}$$

Since $|\operatorname{Aut}_q(\lambda))|$ is the size of the stabilizer of a nilpotent matrix with rational canonical form corresponding to λ by Lemma 3.1.7, the number of such matrices is $\frac{|\mathrm{GL}_n(\mathbb{F}_q)|}{|\operatorname{Aut}_q(\lambda))|}$ by the orbit-stabilizer theorem. By the Fine-Herstein theorem [17], the number of nilpotent matrices over in $M_n(\mathbb{F}_q)$ is $q^{n(n-1)}$. Hence for each n,

$$\sum_{\lambda\vdash n}\frac{1}{|\operatorname{Aut}_q(\lambda))|}=\frac{q^{n(n-1)}}{|\mathrm{GL}_n(\mathbb{F}_q)|}=\frac{q^{n(n-1)}}{q^{n^2}\cdot(q^{-1})_n}=\frac{1}{q^n\cdot(q^{-1})_n}. \tag{3.12}$$

Euler's identity [3, Eq. (2.2.5)] states that, for complex numbers t, z with $|t| < 1$ and $|z| < 1$,

$$1+\sum_{n=1}^{\infty}\frac{t^n}{(1-z)\cdots(1-z^n)}=\frac{1}{\prod_{i\geq 0}(1-tz^i)}. \tag{3.13}$$

Plugging in (3.12) to the LHS of (3.11) and applying Euler's identity with $t = \frac{u}{q}$ and $z = \frac{1}{q}$ yields

$$\sum_{n\geq 0}\frac{u^n}{q^n\cdot(q^{-1})_n}=\sum_{n\geq 0}\frac{\left(\frac{u}{q}\right)^n}{(1-q^{-1})\cdots(1-q^{-n})}=\frac{1}{\prod_{i\geq 0}\left(1-\frac{u}{q}\cdot\frac{1}{q^i}\right)}=\frac{1}{\prod_{i\geq 1}(1-uq^{-i})}. \tag{3.14}$$

Hence $\mu_{u,q}$ is a probability measure.

To show $\mu_{u,q,f}$ is a probability measure we need only show that the measure on $\mathbb{Z}_{\geq 0}$ used to define it is a probability measure. That is exactly the identity

$$\sum_{n\geq 0}\frac{u^n}{(q^{-1})_n}=\frac{1}{(u;q^{-1})_\infty}=\frac{1}{\prod_{i\geq 0}(1-uq^{-i})} \tag{3.15}$$

which is Euler's identity for $t = u$ and $z = q^{-1}$. □

The parameter u determines how much the measure is weighted toward small partitions, as the power $u^{|\lambda|}$ becomes smaller for large $|\lambda|$ (and shrinks faster for small u); the meaning of u will become more clear in Theorem 3.1.10, which finally shows that the limiting measure on partitions obtained from $M_n(\mathbb{F}_q)$ is the u-deformed Cohen-Lenstra measure of Definition 3.1.8.

Theorem 3.1.10. *Let $0 < u < 1$, f monic irreducible nonconstant, and $\mu_{u,q,f}$ as in Definition 3.1.6. Then $\mu_{u,q,f} = \mu_{u^{\deg(f)},q^{\deg(f)}}$. Furthermore, if A is chosen randomly as in Definition 3.1.6, the random partitions $\lambda^{(h)}(A)$ are independent of one another, i.e. for any $g_1, \dots, g_r$ distinct monic irreducible nonconstant polynomials, and sets $S_1, \dots, S_r \subset \mathcal{P}$, we have*

$$\Pr(\lambda^{(g_i)}(A) \in S_i \text{ for all } 1 \le i \le r) = \prod_{1 \le i \le r} \Pr(\lambda^{(g_i)}(A) \in S_i). \tag{3.16}$$

Proof. Let $\mathrm{Orbit}(M_n(\mathbb{F}_q))$ denote the set of orbits of $M_n(\mathbb{F}_q)$ under conjugation by $\mathrm{GL}_n(\mathbb{F}_q)$. Applying Lemma 3.1.7, Remark 1.3.7 and the orbit-stabilizer theorem, we have

$$Z_{M_n(\mathbb{F}_q)} = \frac{|\mathrm{GL}_n(\mathbb{F}_q)|}{|M_n(\mathbb{F}_q)|} \cdot \frac{1}{|\mathrm{GL}_n(\mathbb{F}_q)|} \sum_{A \in M_n(\mathbb{F}_q)} \prod_{f : |\lambda^{(f)}| > 0} x_{f,\lambda^{(f)}} \tag{3.17}$$

$$= (q^{-1})_n \sum_{C \in \mathrm{Orbit}(M_n(\mathbb{F}_q))} \frac{1}{|\operatorname{Stab}(C)|} \prod_{f : |\lambda^{(f)}| > 0} x_{f,\lambda^{(f)}} \tag{3.18}$$

$$= (q^{-1})_n \sum_{\lambda^{(1)},\dots,\lambda^{(m)},f_1,\dots,f_m} \prod_{i=1}^{m} \frac{x_{f_i,\lambda^{(f_i)}}}{|\operatorname{Aut}_{q^{\deg(f_i)}}(\lambda^{(f_i)})|} \tag{3.19}$$

where the sum is over all sets of monic irreducible nonconstant polynomials and corresponding partitions which determine an orbit. Applying this to the expression in (3.5) we have

$$(u;q^{-1})_\infty \sum_{n \ge 0} \frac{u^n}{(q^{-1})_n} Z_{M_n(\mathbb{F}_q)} = (u;q^{-1})_\infty \prod_h \left(1 + \sum_{n \ge 1} \sum_{\lambda \vdash n} x_{h,\lambda} \frac{u^{n \deg(h)}}{|\operatorname{Aut}_{q^{\deg(h)}}(\lambda)|} \right) \tag{3.20}$$

where the product is over all nonconstant monic irreducible h. This is equal to

$$(u;q^{-1})_\infty \prod_h \left(\frac{1}{\prod_{i \ge 1}(1 - u^{\deg(h)} q^{-i \deg(h)})} \sum_{\lambda \in \mathcal{P}} \mu_{u^{\deg(h)},q^{\deg(h)}}(\lambda) x_{h,\lambda} \right) \tag{3.21}$$

by Definition 3.1.8, where again the product is over all nonconstant monic irreducible h, and we take $x_{h,()}$ to be 1. Setting $x_{h,\lambda} = 1$ for $h \neq f$, (3.21) represents the measure $\mu_{u,q,f}$, which by Lemma 3.1.9 is indeed a probability measure; hence (3.21) is equal to 1 if all $x_{h,\lambda}$ are set to 1. Again by Lemma 3.1.9, $\mu_{u^{\deg(h)},q^{\deg(h)}}$ is a probability measure for each h, therefore setting each $x_{h,\lambda}$ to 1 in (3.21) yields[2]

$$(u;q^{-1})_\infty \prod_h \frac{1}{\prod_{i \ge 1}(1 - u^{\deg(h)} q^{-i \deg(h)})} = 1. \tag{3.22}$$

Applying this to (3.21) and following back our chain of equalities, we have

$$(u;q^{-1})_\infty \sum_{n \ge 0} \frac{u^n}{(q^{-1})_n} Z_{M_n(\mathbb{F}_q)} = \prod_h \left(\sum_{\lambda \in \mathcal{P}} \mu_{u^{\deg(h)},q^{\deg(h)}}(\lambda) x_{h,\lambda} \right). \tag{3.23}$$

[2] (3.22) also has a simple independent proof, which is used in the original version of this argument [24].

Setting $x_{h,\lambda} = 1$ for $h \neq f$, the LHS represents $\mu_{u^{\deg(f)},q^{\deg(f)}}$ and the RHS represents $\mu_{u,q,f}$, so the two measures are equal.

Independence of the random partitions (3.16) follows immediately from the factorization in (3.23). Taking (3.23) and setting $x_{h,\lambda} = 1$ for $h \notin \{g_1, \ldots, g_r\}$, and $x_{g_i,\lambda} = \mathbb{1}_{\{\lambda \in S_i\}}$ for $1 \leq i \leq r$ yields the LHS of (3.16). Each term $\Pr(\lambda^{(g_i)}(A) \in S_i)$ in the RHS of (3.16) is given by setting $x_{h,\lambda} = 1$ for $h \neq g_i$ and $x_{g_i,\lambda} = \mathbb{1}_{\{\lambda \in S_i\}}$ in (3.21), and clearly taking the product of these yields the same expression as the one in the previous sentence. If one of the S_i is infinite then we must justify interchanging order of summation here, but this does not pose an issue since all terms are nonnegative. □

Definition 3.1.6 defines a measure on $\bigsqcup_{n \geq 0} M_n(\mathbb{F}_q)$ (where we formally take $M_0(\mathbb{F}_q)$ to have one element) by choosing n as stated and then choosing a uniformly random matrix, but not taking any associated partitions. It is clear that while this measure is not well-defined if $u = 1$, as u approaches 1 it becomes more and more probable to choose a large matrix size n. Hence it makes sense that the limit $u \to 1$ should be similar to the large n limit. Corollary 3.1.13 shows that this limit is indeed the usual Cohen-Lenstra measure (with $u = 1$).

Remark 3.1.11. The reason that using the parameter u rather than simply taking an $n \to \infty$ limit is useful is because for finite n, the random partitions $\lambda^{(f)}$ have the restriction $|\lambda^{(f)}| \cdot \deg(f) \leq n$, which complicates their distribution; however, if the size of the matrix is also randomly distributed as in Theorem 3.1.10 these partitions may be arbitrarily large. The idea of randomizing on the dimension to obtain tractable formulas and then derandomizing through some limiting procedure occurs across many contexts: [6] and others use a technique of "poissonization" and "depoissonization," in which dimensions are chosen according to a Poisson distribution, to study asymptotic behavior of Plancherel measures for the symmetric group. Our parameter u is the exact same construction, but for distribution different from the Poisson. Another example of this auxiliary randomization is the grand canonical ensemble of statistical mechanics; this observation was brought to the author's attention by [24, p. 67], which gives references to the statistical mechanics literature for the interested reader.

Lemma 3.1.12. *Let $[u^n]S$ denote the coefficient of the u^n term in a power series $S \in \mathbb{R}[\![u]\!]$. If $f(u)$ is a function with Taylor series $f(u) = \sum_{n=0}^{\infty} a_n u^n$ converging at $u = 1$, then*

$$\lim_{n \to \infty} [u^n] \frac{f(u)}{1-u} = f(1). \tag{3.24}$$

Proof. Observe that $[u^n]\frac{f(u)}{1-u} = \sum_{i=0}^{n} a_i$. □

Corollary 3.1.13. *For each n, let A_n be a random matrix in $M_n(\mathbb{F}_q)$ with uniform distribution. If $f_1, \ldots, f_r \in \mathbb{F}_q[z]$ are monic irreducible nonconstant polynomials, the distribution of the r-tuple of random partitions $(\lambda^{(f_1)}(A_n), \ldots, \lambda^{(f_r)}(A_n))$ converges pointwise to the product measure $\prod_{1\le i\le r} \mu_{1,q^{\deg(f_i)}}$ on $\mathcal{P}^r$ as $n \to \infty$. In particular, the distribution of an individual random partition $\lambda^{(f)}(A_n)$ converges to the Cohen-Lenstra measure $\mu_{1,q^{\deg(f)}}$.*

Proof. Setting $x_{h,\lambda} = 1$ for $h \notin \{f_1, \ldots, f_r\}$ in (3.23), dividing by $\prod_{i\ge1}(1-uq^{-i})$, and expanding yields

$$(1-u)\left(\sum_{n\ge0} \frac{u^n}{(q^{-1})_n} Z_{M_n(\mathbb{F}_q)}\right) = \frac{1}{\prod_{i\ge1}(1-uq^{-i})} \sum_{(\lambda^1,\ldots,\lambda^r)\in\mathcal{P}^r} \left(\prod_{1\le i\le r} \mu_{u^{\deg(f_i)},q^{\deg(f_i)}}(\lambda^i) x_{f_i,\lambda^i}\right). \tag{3.25}$$

The RHS is an infinite sum over all $(\lambda^1, \ldots, \lambda^r) \in \mathcal{P}^r$ of terms of the form $\left(\prod_{1\le i\le r} x_{f_i,\lambda^i}\right) \cdot \phi_{\lambda^1,\ldots,\lambda^r}(u)$ where

$$\phi_{\lambda^1,\ldots,\lambda^r}(u) := \frac{1}{\prod_{i\ge1}(1-uq^{-i})} \prod_{1\le i\le r} \mu_{u^{\deg(f_i)},q^{\deg(f_i)}}(\lambda^i) \tag{3.26}$$

is a function of u (dependent on q, but since q is fixed we suppress this dependence). $\phi_{\lambda^1,\ldots,\lambda^r}(u)$ may be expressed as a power series in $\mathbb{R}[\![u]\!]$ which converges at $u = 1$, since $\frac{1}{\prod_{i\ge1}(1-uq^{-i})}$ and each $\mu_{u^{\deg(f_i)},q^{\deg(f_i)}}(\lambda^i)$ may each be expressed as a power series converging at $u = 1$, and $\phi_{\lambda^1,\ldots,\lambda^r}$ is a (finite) product of these. Setting $u = 1$ for each such coefficient yields $\phi_{\lambda^1,\ldots,\lambda^r}(1) = \frac{1}{(q^{-1})_\infty} \prod_{1\le i\le r} \mu_{1,q^{\deg(f_i)}}(\lambda^i)$.

Equating coefficients of $\prod_{1\le i\le r} x_{f_i,\lambda^i}$ on both sides of (3.25) and recalling the definition of $Z_{M_n(\mathbb{F}_q)}$, we see that for each $(\lambda^1, \ldots, \lambda^r) \in \mathcal{P}^r$,

$$\phi_{\lambda^1,\ldots,\lambda^r}(u) = (1-u)\sum_{n\ge0} \frac{u^n}{(q^{-1})_n} \Pr(\lambda^{(f_i)}(A_n) = \lambda^i \text{ for all } i). \tag{3.27}$$

Hence, applying Lemma 3.1.12 to $\phi_{\lambda^1,\ldots,\lambda^r}$ and using the computation of $\phi_{\lambda^1,\ldots,\lambda^r}(1)$ in the previous paragraph, we have that

$$\frac{1}{(q^{-1})_\infty} \lim_{n\to\infty} \Pr(\lambda^{(f_i)}(A_n) = \lambda^i \text{ for all } i) = \lim_{n\to\infty} \frac{1}{(q^{-1})_n} \Pr(\lambda^{(f_i)}(A_n) = \lambda^i \text{ for all } i) \tag{3.28}$$

$$= \lim_{n\to\infty} [u^n] \frac{\phi_{\lambda^1,\ldots,\lambda^r}(u)}{1-u} \tag{3.29}$$

$$= \phi_{\lambda^1,\ldots,\lambda^r}(1) \tag{3.30}$$

$$= \frac{1}{(q^{-1})_\infty} \prod_{1\le i\le r} \mu_{1,q^{\deg(f_i)}}(\lambda^i). \tag{3.31}$$

Hence

$$\lim_{n\to\infty} \Pr(\lambda^{(f_i)}(A_n) = \lambda^i) = \prod_{1\le i\le r} \mu_{1,q^{\deg(f_i)}}(\lambda^i), \tag{3.32}$$

finishing the proof. □

With some analytic work one can recover convergence of the measures on possibly infinite sets as well, but this is not necessary for our discussion. The fact that Theorems 1.1.2 and 3.1.10 both recover versions of the Cohen-Lenstra measure is quite surprising, and relies crucially on fact from Proposition 1.3.4 that the $|\operatorname{Aut}_R(M_\lambda)|$ depends only on the size of the residue field of R. As mentioned in Section 1.1, the Cohen-Lenstra measure typically comes from some form of orbit-stabilizer counting, which is apparent in the cycle index manipulations used for Corollary 3.1.13 and also in the original proof of Theorem 1.1.2 given in [19].

3.2 Random partitions, primes, and random matrices over $\mathbb{F}_q[T]$

Though Theorem 3.1.10 and Corollary 3.1.13 are ostensibly about random matrices over finite fields, it is worth putting it in the context of random matrices over PIDs, of which one commonly studied case is over $\mathbb{Z}$ [44, 50]. The cokernel of a nonsingular matrix over a PID is a direct sum of p-torsion parts for each prime p in the PID, and each one is uniquely specified by a partition.

Given $A \in M_n(\mathbb{F}_q)$, $A - T \cdot \mathrm{Id}$ defines a linear map $\mathbb{F}_q[T]^n \to \mathbb{F}_q[T]^n$, the cokernel of which is a torsion $\mathbb{F}_q[T]$-module. The isomorphism type of this module carries the same data as the conjugacy class of A; in Lemma 3.1.7 we used the fact that the automorphisms of this module correspond to the matrices commuting with A, and this orbit-stabilizer argument underpins most of the previous section. Changing perspective away from orbits under conjugation, the results of this section are thus concerned with limiting the distribution of cokernels of those elements of $M_n(\mathbb{F}_q[T])$ which are given by $A - T \cdot \mathrm{Id}$ where $A \in M_n(\mathbb{F}_q)$ is chosen uniformly randomly. In Corollary 3.1.13 we showed that the random partitions corresponding to each prime in $\mathbb{F}_q[T]$, i.e. each monic irreducible nonconstant polynomial, are independent of one another and have distribution $\mu_{1,\deg(f)}$ dependent only on the degree of the polynomial.

There exists similar results for random matrices over $\mathbb{Z}$. Namely, if one chooses a random element $A \in M_n(\mathbb{Z})$ with iid entries chosen uniformly from $\{-m, -m+1, \ldots, m\}$, then one can study the limit of the distribution of $\operatorname{coker}(A)$ as $m \to \infty$, and then the limit of this as $n \to \infty$. [44] show essentially that, for fixed finite collections of rational primes p_i, this distribution is

given by a product of Cohen-Lenstra measures μ_{1,p_i}; we will not show the details but they are similar to Theorem 3.2.2 below. Hence both the integer matrix case and the finite field case discussed in this section yield independent, Cohen-Lenstra-distributed random partitions for each prime. It is worth noting that these Cohen-Lenstra measures are different for different primes: in the finite field case we just showed that $\mu_{u,q,f} = \mu_{u^{\deg f},q^{\deg f}}$ is dependent on the degree of f, and in the integral case, the measure corresponding to a rational prime p is $\mu_{1,p}$ in our notation. Though [44] studies the distribution on the cokernels themselves, as we have seen in this section it is convenient to 'localize' by restricting to studying the distribution of the partition associated to a given prime.

Additionally, Section 1.1 mentioned the universality results of [49], which show in the p-adic case that many different distributions of matrix entries still yield Cohen-Lenstra-distributed cokernels in the limit. It seems reasonable to suppose that such universality results exist for random matrices over $\mathbb{F}_q[T]$, and many different distributions on matrix entries would still yield random partitions distributed as in Corollary 3.1.13 in the limit $n \to \infty$. Theorem 3.2.2 shows that this is indeed the case for a natural choice of distribution on $M_n(\mathbb{F}_q[T])$.

Previously we used $\lambda^{(f)}(A)$ to denote the partition corresponding to the blocks associated to f in the rational canonical form of A. In light of the previous discussion this is just the partition with parts corresponding to summands $\mathbb{F}_q[T]/(f(T)^{\lambda_i})$ in the decomposition of $\operatorname{coker}(A - T \cdot \mathrm{Id})$.

Notation 3.2.1. Fix some $g(T) \in \mathbb{F}_q[T]$. For each $B \in M_n(\mathbb{F}_q[T]/g(T))$ we denote by $\nu^{(f)}(B)$ the partition with parts corresponding to summands $\mathbb{F}_q[T]/(f(T)^{\lambda_i})$ in the decomposition of $\operatorname{coker}(B)$, so that the f-torsion part of $\operatorname{coker}(B)$ is $\bigoplus_i \mathbb{F}_q[T]/(f(T)^{\nu_i^{(f)}})$. Note that with this notation, $\lambda^{(f)}(A) = \nu^{(f)}(A - T \cdot \mathrm{Id})$ for $A \in M_n(\mathbb{F}_q)$.

Theorem 3.2.2. *Let $A_{n,d}$ be a random matrix in $M_n(\mathbb{F}_q[T])$ with iid entries, each one distributed uniformly randomly among elements of $\mathbb{F}_q[T]$ of degree $\leq d$ (where we take this to include the zero polynomial). Let $f_1, \ldots, f_r \in \mathbb{F}_q[T]$ be a collection of distinct monic irreducible nonconstant polynomials. Then as $n, d \to \infty$, the distribution of the r-tuple of random partitions $(\nu^{(f_1)}(A_{n,d}), \ldots, \nu^{(f_r)}(A_{n,d}))$ converges pointwise to the product measure $\prod_{1 \leq i \leq r} \mu_{1,q^{\deg(f_i)}}$ on $\mathcal{P}^r$.*

The theorem and proof follow similar lines to [44, Theorem 3.2] for the case of integer matrices.

Proof. We must show for any $\lambda^1, \dots, \lambda^r \in \mathcal{P}$ that

$$\lim_{n,d\to\infty} \Pr(\nu^{(f_i)}(A_{n,d}) = \lambda^i \text{ for all } 1 \le i \le r) = \prod_{1\le i\le r} \mu_{1,q^{\deg(f_i)}}(\lambda^i). \tag{3.33}$$

Let $g = \prod_{1\le i\le r} f_i^{\lambda_1^i+1}$ (λ_1^i is just the first part of λ^i). Suppose $B \in M_n(\mathbb{F}_q[T])$ is any fixed matrix with $\nu^{(f_i)}(B) = \lambda^i$ for all $1 \le i \le r$, and S is diagonal matrix which is B in Smith normal form. The image B' of B in $M_n(\mathbb{F}_q[T]/(g))$ also has a unique Smith normal form, and since the image S' of S in $M_n(\mathbb{F}_q[T]/(g))$ is still in Smith normal form, S' is the Smith normal form of B'. For each i, the f_i-torsion part of $\operatorname{coker}(B)$ is $\prod_j \mathbb{F}_q[T]/(f_i^{\lambda_j^i})$, and since we chose g to be divisible by a higher power of f_i, we have that $\operatorname{coker}(B')$ has the same f_i-torsion part as $\operatorname{coker}(B)$. Hence $\nu^{(f_i)}(B) = \nu^{(f_i)}(B')$ for each i. For $d > \deg g$, we have that $A_{n,d} \pmod g$ is uniformly distributed in $M_n(\mathbb{F}_q[T]/(g))$, so

$$\lim_{n,d\to\infty} \Pr(\nu^{(f_i)}(A_{n,d}) = \lambda^i \text{ for all } 1 \le i \le r) = \lim_{n\to\infty} \Pr(\nu^{(f_i)}(B_n) = \lambda^i \text{ for all } 1 \le i \le r) \tag{3.34}$$

where B_n is taken uniformly randomly from $M_n(\mathbb{F}_q[T]/(g))$. This is a finite set, so the above is equal to

$$\lim_{n\to\infty} \frac{|\{A \in M_n(\mathbb{F}_q[T]/(g)) : \nu^{(f_i)}(A) = \lambda^i \text{ for all } 1 \le i \le r\}|}{|M_n(\mathbb{F}_q[T]/(g))|}. \tag{3.35}$$

We claim there is a bijection between

$$\{A \in M_n(\mathbb{F}_q[T]/(g)) : \nu^{(f_i)}(A) = \lambda^i \text{ for all } 1 \le i \le r\}$$

and

$$\prod_{1\le i\le r} \{A_i \in M_n(\mathbb{F}_q[T]/(f_i^{\lambda_1^i+1})) : \nu^{(f_i)}(A_i) = \lambda^i\}$$

given by $A \mapsto (A \pmod{f_1^{\lambda_1^i+1}}, \dots, A \pmod{f_r^{\lambda_1^r+1}})$. This follows upon noting that $\nu^{(f_i)}(A \pmod g) = \lambda^i$ if and only if $\nu^{(f_i)}(A \pmod{f_i^{\lambda_1^i+1}}) = \lambda^i$, by the reasoning at the beginning of this proof, and applying the Chinese remainder theorem. Since the Chinese remainder theorem also gives a bijection $M_n(\mathbb{F}_q[T]/(g)) \to \prod_{1\le i\le r} M_n(\mathbb{F}_q[T]/(f_i^{\lambda_1^i+1}))$, we have that

$$\begin{aligned}&\frac{|\{A \in M_n(\mathbb{F}_q[T]/(g)) : \nu^{(f_i)}(A) = \lambda^i \text{ for all } 1 \le i \le r\}|}{|M_n(\mathbb{F}_q[T]/(g))|}\\ &= \prod_{1\le i\le r} \frac{|\{A_i \in M_n(\mathbb{F}_q[T]/(f_i^{\lambda_1^i+1})) : \nu^{(f_i)}(A_i) = \lambda^i\}|}{|M_n(\mathbb{F}_q[T]/(f_i^{\lambda_1^i+1}))|}.\end{aligned}$$

Hence it suffices to prove

$$\lim_{n\to\infty} \frac{|\{A_i \in M_n(\mathbb{F}_q[T]/(f_i^{\lambda_1^i+1})) : \nu^{(f_i)}(A_i) = \lambda^i\}|}{|M_n(\mathbb{F}_q[T]/(f_i^{\lambda_1^i+1}))|} = \mu_{1,q^{\deg(f_i)}}(\lambda). \tag{3.36}$$

In the beginning we reduced a probability over $M_n(\mathbb{F}_q[T])$ to one over the finite set $M_n(\mathbb{F}_q[T]/(f_i^{\lambda_1^i+1}))$; we now reverse this process to get a probability over a complete DVR with finite residue field, which we know how to manipulate[3] . Specifically, if $\mathbb{F}_q[T]_{(f_i^{\lambda_1^i+1})}$ is the localization at the ideal $(f_i^{\lambda_1^i+1})$ and $\widehat{\mathbb{F}_q[T]}_{(f_i^{\lambda_1^i+1})}$ its completion, then it is a DVR with residue field isomorphic to $\mathbb{F}_q[T]/(f_i^{\lambda_1^i+1}) \cong \mathbb{F}_{q^{\deg(f_i)}}$. Since the Haar measure on $\widehat{\mathbb{F}_q[T]}_{(f_i^{\lambda_1^i+1})}$ projects to the uniform measure on finite quotients, for A random in $M_n(\widehat{\mathbb{F}_q[T]}_{(f_i^{\lambda_1^i+1})})$ with Haar measure we have

$$\frac{|\{A_i \in M_n(\mathbb{F}_q[T]/(f_i^{\lambda_1^i+1})) : \nu^{(f_i)}(A_i) = \lambda^i\}|}{|M_n(\mathbb{F}_q[T]/(f_i^{\lambda_1^i+1}))|} = \Pr(\lambda^{(f_i)}(A) = \lambda^i). \tag{3.37}$$

In Section 3.3 we will prove the analogue of Theorem 1.1.2 for rings of integers of arbitrary local fields (Theorem 3.3.4), from which it immediately follows that

$$\lim_{n\to\infty} \Pr(\lambda^{(f_i)}(A) = \lambda^i) = \mu_{1,q^{\deg(f_i)}}(\lambda^i), \tag{3.38}$$

completing the proof. □

Between Theorem 3.2.2 and Corollary 3.1.13, we now have two cases of a distribution on $M_n(\mathbb{F}_q[T])$ producing cokernels which are distributed according to a product of measures $\mu_{1,q^{\deg(f)}}$ in the limit. Given that strong universality results have already been shown in classical random matrix theory (see for instance [12]) and over $\mathbb{Z}$ and $\mathbb{Z}_p$ [37, 49, 50], this suggests the following.

Problem 3.2.3. *Prove a natural universality result for random matrices over $\mathbb{F}_q[T]$ which includes both Theorem 3.2.2 and Corollary 3.1.13 as special cases.*

The author is not aware of any substantial body of literature on random matrices over $\mathbb{F}_q[T]$, except implicitly in work on random matrices over $\mathbb{F}_q$ and over arbitrary PIDs, so this represents a gap worth filling. It remains to be seen whether the parameter u generalizes well to these settings.

3.3 Markov chains, filtered vector spaces, and a new plane partition

Recall that one of our original goals was to study the relation between random matrix theory over finite fields, which concerns distribution of orbits under GL_n, and random matrix theory over

[3] It is also possible to simply compute (3.36) directly as in [16].

local fields, which concerns distribution of cokernels/Smith normal forms. The Smith normal form of $A \in M_n(R)$ roughly gives the data of how much applying A to elements in R^n raises their valuation, while the partition $\lambda^{(x)}$ of $A \pmod{\omega}$ gives the data of how many times A must be applied to vectors in $\mathbb{F}_q^n$ to raise their valuation by at least 1 (equivalently, annihilate them modulo ω). These are very dual questions, and this parallel may be thought of as the basis for this section. This analogy is furthered by the fact that both occurrences of the Cohen-Lenstra measure discussed above are related to certain Markov chains. In the finite field case, we have the following.

Proposition 3.3.1 ([24]). *Fix u, q as before, and define a Markov chain on $\mathbb{Z}_{\geq 0} \cup \{\infty\}$ by*

$$\Pr(X_{n+1} = b | X_n = a) = \frac{u^b (q^{-1})_a (u; q^{-1})_a}{q^{b^2} (q^{-1})_{a-b} (q^{-1})_b (u; q^{-1})_b}. \tag{3.39}$$

Then the random partition $\lambda = (X_1, X_2, \ldots)'$ (note we are taking the conjugate) is distributed according to $\mu_{u,q}$.

This may be proven by explicit computation using the formula for $\mu_{u,q}$, see [24]. While this proof does not seem to give much insight into why the parts of λ' should be given by a Markov chain, it is still quite surprising that they are at all, since naively one would expect that each of $\lambda'_1, \ldots, \lambda'_r$ should affect the conditional distribution of λ'_{r+1}. This Markov chain is quite combinatorially interesting and is used to give a probabilistic proof of the Rogers-Ramanujan identities in [23], suggesting it is a natural object to attempt to generalize to the local field setting.

In this section we will define and study the distribution of the filtered vector spaces and plane partition mentioned in the Introduction, which brings together the finite field theory of the previous section and the local field theory of the previous chapters; however, before doing this, we return to cokernels. We have already seen that the limiting distribution of $\operatorname{coker}(A)$, for $A \in M_n(R)$ chosen with Haar measure, is $\mu_{1,q}$ where $q = |R/\omega|$. Proposition 3.3.1 shows in particular that the conjugate of the cokernel partition of A has distribution given by a Markov chain. This result was shown in slightly different form in [15], using a delicate inductive argument and the noncommutative q-binomial theorem. It also follows from Theorem 1.1.2 (generalized to rings of integers of arbitrary local fields) and Proposition 3.3.1, though this yields no insight. We give a new simpler and direct proof, which gives the $u = 1$ case of the formula in (3.39) a clear interpretation in terms of linear algebra over R and $\mathbb{F}_q$. This proof is a special case of the method used to prove Theorem 3.3.11, but it is instructive to see this simpler case first. The

main idea of this proof is that (1) the data of a matrix's cokernel is equivalent to that of the kernels of the reduction of the matrix modulo all powers of ω, and (2) as per Lemma 1.2.10, the Haar measure is given by choosing coefficients of the 'power series in ω' independently, and this independence yields the Markov property. The main idea may also be viewed as a probabilistic Hensel's lemma, where instead of showing that there exists a lift which is a solution, one looks at the probability that a random lift is a solution. It is worth noting that this same idea of reduction modulo various powers of the uniformizer and independence of series coefficients was used to compute the Cohen-Lenstra measure and a related symplectic version in [2], but in such a way that the Markov chain was not readily apparent.

Recall that the *q-binomial coefficient* is given by

$$\begin{bmatrix} a \\ b \end{bmatrix}_q := \frac{\prod_{i=1}^{a}(q^i-1)}{\prod_{i=1}^{a-b}(q^i-1)\cdot\prod_{i=1}^{b}(q^i-1)} \tag{3.40}$$

(we could also write this as $\frac{(q)_a}{(q)_{a-b}(q)_b}$, but the above expression reads better since all (q^i-1) terms are positive for $q>1$).

Recall the well-known result that $\begin{bmatrix} a \\ b \end{bmatrix}_q$ gives the number of b-dimensional subspaces of $\mathbb{F}_q^a$; a nice way to see this is that there are $(q^a-1)\cdots(q^a-q^{a-b-1})$ choices of an ordered set of n linearly independent vectors in $\mathbb{F}_q^a$, and for each subspace there are $|\mathrm{GL}_b(\mathbb{F}_q)|$ such choices of vectors since $\mathrm{GL}_b(\mathbb{F}_q)$ acts transitively on them, so the total is

$$\frac{(q^a-1)\cdots(q^a-q^{a-b-1})}{|\mathrm{GL}_b(\mathbb{F}_q)|} = \begin{bmatrix} a \\ b \end{bmatrix}_q. \tag{3.41}$$

Notation 3.3.2. For the remainder of this section φ_j will denote the quotient map $R^n \to (R/\omega^j)^n$, or by abuse of notation the map $(R/\omega^i)^n \to (R/\omega^j)^n$ for $i>j$. Additionally, V will denote $(R/\omega)^n \cong \mathbb{F}_q^n$. If $A \in M_n(R)$ is some matrix, then $A_r \in M_n(R/\omega^r)$ will denote the matrix with each entry of A taken modulo ω^r. We will often consider the map $A_r : (R/\omega^r)^n \to (R/\omega^r)^n$, and sometimes slightly abuse notation and consider A_r as a map $R^n \to (R/\omega^r)^n$ by implicitly precomposing with the quotient $R^n \to (R\omega^r)^n$. When speaking of $\ker A_r$ we mean as a map $(R/\omega^r)^n \to (R/\omega^r)^n$.

Notation 3.3.3. For $x=(x_1,\ldots,x_n)\in R^n$, let $\mathrm{val}(x) = \min_{1\le i\le n}\mathrm{val}(x_i)$. Equivalently, $\mathrm{val}(x)$ is the largest integer m such that there exists $x'\in R^n$ with $x=\omega^m x'$.

Theorem 3.3.4. *Let $X\in M_n(R)$ be distributed with respect to the Haar probability measure, and let $V_i(X) := \{\varphi_1(v) : \mathrm{val}(Xv)\ge \mathrm{val}(v)+i\} \subset (R/\omega)^n$; because X is random, $V_i(X)$ is a random subspace. Then $\mathbb{F}_q^n = V_0(X)\supset V_1(X)\supset\ldots$, and the distribution of the spaces $V_1(X), V_2(X),\ldots$*

is given by a Markov chain with the following transition probabilities, where $W_r, W_{r+1} \subset \mathbb{F}_q^n$ *are arbitrary subspaces and* $a = \dim W_r, b = \dim W_{r+1}$*:*

$$\Pr(V_{r+1}(X) = W_{r+1} | V_r(X) = W_r) = \begin{cases} \begin{bmatrix} a \\ b \end{bmatrix}_q \frac{|\mathrm{GL}_{a-b}(\mathbb{F}_q)|}{|M_a(\mathbb{F}_q)|} & \text{if } W_{r+1} \subset W_r \\ 0 & \text{if } W_{r+1} \not\subset W_r \end{cases}. \tag{3.42}$$

Letting $\lambda_i'(X) := \dim_{\mathbb{F}_q} V_i(X)$*, and defining the partition* $\lambda(X)$ *by defining its conjugate partition* $\lambda(X)' = (\lambda_1'(X), \lambda_2'(X) \ldots)$*, we have that* $\lambda(X)$ *is the cokernel partition of* X*. Furthermore, the integers* $\lambda_i'(X)$ *are given by a Markov chain with transition probabilities*

$$\Pr(\lambda_{r+1}(X)' = b | \lambda_r(X)' = a) = \begin{bmatrix} a \\ b \end{bmatrix}_q^2 \frac{|\mathrm{GL}_{a-b}(\mathbb{F}_q)|}{|M_a(\mathbb{F}_q)|}, \tag{3.43}$$

which (after simplifying) are the same transition probabilities as given by setting $u = 1$ *in* (3.39)*.*

Note that it follows immediately that, in the limit $n \to \infty$, the cokernel partition of X is distributed according to $\mu_{1,q}$.

Proof. First suppose $A \in M_n(R)$ is a fixed nonsingular matrix, and λ is its cokernel partition; we show why this is equal to the conjugate of the partition $(\dim V_1(A), \dim V_2(A), \ldots)$ defined in the statement. The parts λ_i' of the conjugate partition give the number of parts of λ of size $\geq i$, or in other words the number of summands of $\mathrm{coker}(A) \cong \bigoplus_j R/\omega^j$ with $j \geq i$. The matrix $A_r \in M_n(R/\omega^r)$ is a map $(R/\omega^r)^n \to (R/\omega^r)^n$, which now likely has nontrivial kernel. In the case $r = 1$ we just have $A_1 \in M_n(\mathbb{F}_q)$, and the kernel is some subspace of $\mathbb{F}_q^n$. If $r = 2$, then recalling the natural map $(R/\omega)^n \to \omega(R/\omega^2)^n$, it is easy to see that $\ker A_2$ includes the image of $\ker A_1$ in $\omega(R/\omega^2)^n$ under this map. However, $\ker A_2$ will also include elements of $(R/\omega^2)^n$ with valuation 0, namely those vectors v for which $\mathrm{val}(v) = 0$ and $\mathrm{val}(Av) \geq 2$. Looking at the Smith normal form of A, there is a natural choice of such a vector for each invariant factor ω^{λ_i} with $\lambda_i \geq 2$, and their R-span is the submodule $\{v \in R^n : \mathrm{val}(Av) \geq \mathrm{val}(v) + 2\}$. Projecting to $\mathbb{F}_q^n \cong (R/\omega)^n$, the dimension of the resulting vector space is exactly λ_2'. In general, recalling $V_i(A) = \{\varphi_1(v) : \mathrm{val}(Av) \geq \mathrm{val}(v) + i\} \subset (R/\omega)^n$, we have that

$$\dim_{\mathbb{F}_q} V_i(A) = \lambda_i'. \tag{3.44}$$

Furthermore, clearly $\mathbb{F}_q^n = V_0(A) \supset V_1(A) \supset V_2(A) \supset \ldots$. It follows from Smith normal form that $V_{\lambda_1+1} = 0$, so we have a filtration $\mathbb{F}_q^n = V_0(A) \supset V_1(A) \supset \ldots \supset V_{\lambda_1}(A) \supset 0$ and $\lambda' = (\lambda_1', \ldots, \lambda_{\lambda_1}')$ is just giving the dimensions of the subspaces in this filtration. The relation

$V_i(A) \supset V_{i+1}(A)$ further gives us that $\Pr(V_{r+1}(X) = W_{r+1} | V_r(X) = W_r) = 0$ if $W_{r+1} \not\subset W_r$, accounting for that case in (3.42).

Let A be a fixed matrix with $V_r(A) = W_r$, and X be a random element of $M_n(R)$, distributed according to the Haar measure conditioned on $V_r(X) = W_r$. For a subspace W_{r+1}, we will compute

$$\Pr(V_{r+1}(X) = W_{r+1} | X_r = A_r) \tag{3.45}$$

and show that this depends only on W_r and not on A, from which it will follow immediately that

$$\Pr(V_{r+1}(X) = W_{r+1} | X_r = A_r) = \Pr(V_{r+1}(X) = W_{r+1} | V_r(X) = W_r). \tag{3.46}$$

If C is a fixed matrix with $C_r = A_r$, then $C = A + \omega^r B$ for some $B \in M_n(R)$. Furthermore, Corollary 1.2.11 implies that X is equal in distribution to $A + \omega^r Y$ where A is the fixed matrix above and Y is a random matrix with Haar distribution. Hence

$$\Pr(V_{r+1}(X) = W_{r+1} | X_r = A_r) = \Pr(V_{r+1}(A + \omega^r Y) = W_{r+1}) \tag{3.47}$$

and it suffices to compute the latter.

Let B again be any fixed matrix. Choose a basis $v_1, \ldots, v_a$ for W_r and choose a lift $\tilde{v}_i \in (R/\omega^{r+1})^n$ of each basis element such that $A_{r+1}\tilde{v}_i = 0$. Let $\tilde{W}_r$ be the R/ω^{r+1}-module spanned by the $\tilde{v}_i$. Any element of $(R/\omega^{r+1})^n$ which is annihilated by $(A + \omega^r B)_{r+1}$ must also, after reducing modulo ω^r, be annihilated by $(A + \omega^r B)_r = A_r$, hence it must be of the form $x + \omega y$ for some $x \in \tilde{W}_r$ and y arbitrary. To find $V_{r+1}(A + \omega^r B)$ it suffices to find all $x \in \tilde{W}_r$ for which there exists some y such that $(A + \omega^r B)_{r+1}$ annihilates $x + \omega y$. But we have

$$(A + \omega^r B)_{r+1}(x + \omega y) = \omega A_{r+1} y + \omega^r B_{r+1} x \tag{3.48}$$

because A_{r+1} annihilates x. There exists a y for which this is 0 exactly when $\omega^r B_{r+1} x \in \omega A_{r+1}(R/\omega^{r+1})^n$, or equivalently, exactly when $\omega^r B_{r+1} x \in \omega A_{r+1}(R/\omega^{r+1})^n \cap \omega^r (R/\omega^{r+1})^n$.

This R-module

$$U_{r+1} := \omega A_{r+1}(R/\omega^{r+1})^n \cap \omega^r (R/\omega^{r+1})^n \tag{3.49}$$

naturally has the structure of an $R/\omega \cong \mathbb{F}_q$-vector space. We have that

$$V_{r+1}(A + \omega^r B) = \{\varphi_1(x) : x \in \tilde{W}_r \text{ s.t. } \omega^r B_{r+1} x \in U_{r+1}\}. \tag{3.50}$$

Note that $\omega^r B_{r+1} x$ depends only on $\varphi_1(x)$, so the map $\omega^r B_{r+1}$ on $\tilde{W}_r$ factors through a map ϕ_B on $\varphi_1(\tilde{W}_r) = W_r$. We have been considering a fixed matrix B for concreteness, but now note that

if Y is Haar-distributed, then ϕ_Y will be a uniformly random map $W_r \to \omega^r(R/\omega^{r+1})^n \cong \mathbb{F}_q^n$. It follows that the induced map $\phi'_Y : W_r \to \mathbb{F}_q^n/U_{r+1}$, given a choice of bases for W_r and $\mathbb{F}_q^n/U_{r+1}$, is uniformly randomly distributed in $M_{(n-\dim U_{r+1})\times\lambda'_r(A)}$. Hence by (3.50), $V_{r+1}(A+\omega^r Y) = \ker\phi'_Y$ is the kernel of a uniformly random element of $M_{(n-\dim U_{r+1})\times\lambda'_r(A)}$, so it suffices to compute $\dim U_{r+1}$.

Note that U_{r+1} is defined solely in terms of A. If ωA were in Smith normal form, then it is clear that for each i such that $\lambda_i \le r-1$, the corresponding invariant factor is ω^{λ_i} and consequently $(\omega A)\omega^{r-\lambda_i}e_i \in U_{r+1}$; furthermore, these elements of U_{r+1} form a basis. For A not in Smith normal form, let $D = Q(\omega A)T$ be in Smith normal form for $Q, T \in \mathrm{GL}_n(R)$; then the above argument shows that $\{T^{-1}e_i : \lambda_i \le r-1\}$ is a basis for U_{r+1}. Hence

$$\dim_{\mathbb{F}_q} U_{r+1} = n - \lambda'_r(A). \tag{3.51}$$

Hence ϕ'_Y is a uniformly random element of $M_{\lambda'_r(A)\times\lambda'_r(A)}(\mathbb{F}_q)$. Following our string of equivalent probabilities, we have

$$\Pr(V_{r+1}(X) = W_{r+1} | V_r(X) = W_r) = \Pr(\ker\phi'_Y = W_{r+1}) \tag{3.52}$$

where ϕ'_Y is a uniformly random map $W_r \to \mathbb{F}_q^a$. Any such map factors uniquely through an injection $W_r/W_{r+1} \hookrightarrow \mathbb{F}_q^a$, hence

$$\Pr(\ker\phi'_Y = W_{r+1}) = \frac{|\{\text{injections } W_r/W_{r+1} \to \mathbb{F}_q^a\}|}{|M_a(\mathbb{F}_q)|}. \tag{3.53}$$

Such an injection is uniquely specified by a pair of a $(a-b)$-dimensional subspace of $\mathbb{F}_q^a$ (representing its image) and an element of $\mathrm{GL}_{a-b}(\mathbb{F}_q)$ representing the map onto the image. Hence this is equal to

$$\begin{bmatrix} a \\ a-b \end{bmatrix}_q \frac{|\mathrm{GL}_{a-b}(\mathbb{F}_q)|}{|M_a(\mathbb{F}_q)|}. \tag{3.54}$$

Because there are $\begin{bmatrix} a \\ b \end{bmatrix}_q$ subspaces of W_r with dimension b, the probability that $V_{r+1}(X)$ is *any* subspace of dimension b, not necessarily W_{r+1}, is

$$\Pr(\lambda'_{r+1}(X) = b | \lambda'_r = a) = \begin{bmatrix} a \\ b \end{bmatrix}_q^2 \frac{|\mathrm{GL}_{a-b}(\mathbb{F}_q)|}{|M_a(\mathbb{F}_q)|}. \tag{3.55}$$

It is a quick calculation to show this is the same probability as in (3.39) when $u = 1$. □

Having gone through the details of this calculation, it is worth recapitulating heuristically where the Markov property comes from. We may view the i^{th} part of the conjugate of the

cokernel partition $\lambda(A)$ as the dimension of the space of elements of R^n of valuation zero which are brought up to valuation $\geq i$ when hit by A. If $\lambda'_r(A)$ is known, then asking the distribution of $\lambda'_{r+1}(A)$ is simply asking how many of these elements were actually brought up to valuation $\geq r+1$ when hit by A. Quite naturally, this is determined by the coefficients of ω^{r+1} in the power series representation of the entries of A. Since these coefficients are uniformly random in $R/\omega \cong \mathbb{F}_q$ by Lemma 1.2.10, independent of the coefficients of $1, \omega, \ldots, \omega^r$, the distribution of $\lambda'_{r+1}(A)$ should depend only on that of $\lambda'_r(A)$. We now generalize this idea, starting with defining the plane partition alluded to earlier.

Definition 3.3.5. A *plane partition* is an array of nonnegative integers $n_{i,j}$ indexed by pairs $(i,j) \in \mathbb{N}^2$, such that rows and columns are weakly decreasing ($n_{i,j} \geq n_{i+1,j}$ and $\geq n_{i,j+1}$), and such that only finitely many $n_{i,j}$ are nonzero.

Each row and column of a plane partition is a usual integer partition, hence the name. Plane partitions are widely-studied in combinatorics, see e.g. [41]. Probability measures on plane partitions have been studied as well, for which a good quick survey with citations to the original literature is [5, Section 6.2].

Definition 3.3.6. For any matrix $A \in M_n(R)$ and any pair $i, j \in \mathbb{N}$, we define a subspace $V_{i,j}(A) \subset \mathbb{F}_q^n$ by

$$V_{i,j}(A) := \varphi_1(\ker(A_i|_{\mathrm{Im}(A_i^{j-1})})) \tag{3.56}$$

where by convention we take A_i^0 to be the identity matrix. Equivalently, $V_{i,j}(A) = A_1^{j-1}\varphi_1(\ker(A_i^j))$.

We then define $T_{i,j}(A) = \dim_{\mathbb{F}_q} V_{i,j}(A)$, and define $T(A)$ to be the array of nonnegative integers with $(i,j)^{\text{th}}$ entry $T_{i,j}(A)$. When we speak of rows and columns of $T(A)$, we are considering the following orientation (indexed by the same convention as matrix entries)

$$\begin{array}{cccc} T_{1,1} & T_{1,2} & T_{1,3} & \\ T_{2,1} & T_{2,2} & T_{2,3} & \cdots \\ T_{3,1} & T_{3,2} & T_{3,3} & \\ \vdots & & & \ddots \end{array} \tag{3.57}$$

We often write $T_{i,j}$ or $V_{i,j}$ without A when A is clear from context.

The most important basic properties of the spaces $V_{i,j}(A)$ and the array $T(A)$ are collected below.

Lemma 3.3.7. *Let $A \in M_n(R)$ be nonsingular.*

1. *$V_{i+1,j}(A)$ and $V_{i,j+1}(A)$ are contained in $V_{i,j}(A)$ for all $i, j \in \mathbb{N}$.*

2. *$(T_{i,1}(A))_{i\geq 1} = \lambda^{(x)}(A_1)'$, where $\lambda^{(x)}(A_1)$ is the partition of rational canonical form blocks associated to $f(x) = x$ as in Proposition 3.1.1.*

3. *$(T_{1,j}(A))_{j\geq 1} = \lambda(A)'$ where $\lambda(A)$ is the cokernel partition.*

4. *$T(A)$ is a plane partition.*

Proof. (1) Since $\mathrm{Im}A_i^j \subset \mathrm{Im}A_i^{j-1}$, it follows that $V_{i,j+1} \subset V_{i,j}$. Similarly, since for any v if $A_{i+1}\varphi_{i+1}(v) = 0$ then $A_i\varphi_i(v) = 0$, it follows that $V_{i+1,j} \subset V_{i,j}$.

(2) We may assume A_1 is in rational canonical form, and consider the nilpotent blocks associated to $f(x) = x$. For each nilpotent block of size $\geq j$ there is an element of $\mathrm{Im}(A_1^{j-1})$ annihilated by A_1, and these form a basis for $\ker(A_1|_{\mathrm{Im}(A_1^{j-1})})$. Hence $T_{1,j}$ counts the number of blocks of size $\geq j$, which is $(\lambda^{(x)})'_j$.

(3) If A is in Smith normal form, then $T_{i,1}$ counts the number of zeros on the diagonal of A_i, which is the number of parts of $\lambda(A)$ greater than i, i.e. $\lambda_i(A)'$. The claim follows for A not in Smith normal form since multiplication by invertible matrices does not change $T_{i,1}(A)$.

(4) From (2) and (3) it follows that the first row and column have only finitely many nonzero entries, and by (1) all rows and columns are weakly decreasing, hence only finitely many $T_{i,j}(A)$ are nonzero. □

Definition 3.3.8. A *filtration* of a vector space V is a sequence of subspaces $V = V_1 \supset V_2 \supset \ldots$, which in this thesis we will take to stabilize at the zero vector space. A map between filtered vector spaces V and W is a linear map $\phi : V \to W$ such that $\phi(V_i) \subset W_i$ for each i. We refer to the data of a vector space and a filtration on it as a *filtered vector space*.

Remark 3.3.9. Given a filtered vector space $V = V_1 \supset V_2 \supset \ldots \supset V_n = 0$, it is a common and natural construction to take the *associated graded vector space* $\bigoplus_{i=1}^{n-1} V_i/V_{i+1}$. Taking the $V_i = V_{i,1}(A)$ (resp. $V_{1,i}(A)$) for each i, the dimension of the j^{th} component of the associated graded vector space is just the multiplicity $m_j(\lambda(A))$ (resp. $m_j(\lambda^{(x)}(A_1))$) by part (3) (resp. (2)) of Lemma 3.3.7.

Definition 3.3.10. Define $V(A)$ to be the data of every $V_{i,j}(A)$ for $i, j \in \mathbb{N}$. Denote by $V_i(A)$ the filtered vector space $V_{i,1}(A)$ with the filtration $V_{i,1}(A) \supset V_{i,2}(A) \supset \ldots$. Note that we have inclusions of filtered vector spaces (injective maps which respect the filtration in the sense of

Definition 3.3.8) $V_1(A) \hookleftarrow V_2(A) \hookleftarrow \ldots$, so $V(A)$ may be viewed as yielding a filtration of filtered vector spaces.

$V(A)$ is a vector-space analogue of $T(A)$, but carries strictly more data by specifying subspaces rather than just their dimensions. The Markov chain of Theorem 3.3.4 gave the transition dynamics of the first column of $V(A)$ (drawn in the same orientation as the plane partition in (3.57)), neglecting all other columns. In general, it is natural to ask, if one fixes the first r rows of $V(A)$, what is the distribution of the next row. For general r this seems much harder, but for the second row it is possible to obtain an explicit answer, which we do now.

Theorem 3.3.11. *Let $\Lambda \subset \mathbb{F}_q^n$ be a filtered vector space with filtration $\Lambda = \Lambda_1 \supset \Lambda_2 \supset \ldots$, and $W \subset \Lambda$ be another filtered vector space with filtration $W = W_1 \supset W_2 \supset \ldots$ such that $W_i \subset \Lambda_i$ for each i (i.e. the inclusion $W \hookrightarrow \Lambda$ is a map of filtered vector spaces). If X is a random matrix in $M_n(R)$ distributed with respect to Haar measure, then*

$$\Pr(V_{2,m}(X) = W_m | V_1(X) = \Lambda \text{ and } V_{2,j}(X) = W_j \text{ for } 1 \leq j \leq m-1) \tag{3.58}$$

depends only on W_{m-1}, W_m and Λ_m, and is given by

$$q^{-\mu_m \cdot \lambda_m} \frac{(q^{-1})_{\lambda_m}}{(q^{-1})_{\lambda_m - (d_m - \mu_m)}} \tag{3.59}$$

where $\lambda_i = \dim \Lambda_i$, $\mu_i = \dim W_i$, and $d_i = \dim W_{i-1} \cap \Lambda_i$ for each i, with d_1 taken to be λ_1. Hence

$$\Pr(V_2(X) = W | V_1(X) = \Lambda) = \prod_{m \geq 1} q^{-\mu_m \cdot \lambda_m} \frac{(q^{-1})_{\lambda_m}}{(q^{-1})_{\lambda_m - (d_m - \mu_m)}} \tag{3.60}$$

where the equalities $V_2(X) = W$ and $V_1(X) = \Lambda$ are of filtered vector spaces.

Proof. The argument is similar to the proof of Theorem 3.3.4. We first claim there must exist $A \in M_n(R)$ with $V_{1,j}(A) = W_{1,j}$ for all j and $V_{2,j}(A) = W_{2,j}$ for $1 \leq j \leq m-1$. The fact that there exists an $A_1 \in M_n(\mathbb{F}_q)$ with $V_{1,j}(A) = W_{1,j}$ follows by choosing the appropriate rational canonical form blocks; it would be possible to make a similar explicit choice of A_2 so that $V_{2,j}(A) = W_{2,j}$ for $1 \leq j \leq m-1$, but we simply induct on the theorem. The proof of Theorem 3.3.4 yields that there exists A (which in fact occurs with nonzero (Haar-)probability!) such that $V_{2,1}(A) = W_{2,1}$, which provides the base case. From here, assuming that (3.58) has been proven for $m-1$, we have that there exists A with $V_{1,j}(A) = W_{1,j}$ for all j and $V_{2,j}(A) = W_{2,j}$ for $1 \leq j \leq m-1$, since if X is random with the appropriate conditioning this occurs with nonzero probability.

Take A as in the previous paragraph, and let

$$S = \{B \in M_n(R) : B_1 = A_1 \text{ and } B_2 v = A_2 v \text{ for all } v \in \ker A_2^{m-1}\}. \tag{3.61}$$

We will compute

$$\Pr(V_{2,m}(X) = W_m | X \in S) \tag{3.62}$$

and show it depends only on Λ_m, W_{m-1}, and W_m, not on A, and hence is equal to the LHS of (3.58).

Let $B \in S$ be a fixed matrix, and let

$$U_m(A) = \ker A_1^m \cap A_1^{-1}(\varphi_1(\ker A_2^{m-1})) \tag{3.63}$$

(here A_1 may not be invertible but A_1^{-1} simply denotes the preimage under A_1). Clearly $\varphi_1(\ker B_2^m) \subset \ker B_1^m = \ker A_1^m$, and since $B_2 \ker B_2^m = \ker B_2^{m-1}$ we have

$$\varphi_1(\ker B_2^m) \subset \varphi_1(B_2^{-1}(\ker B_2^{m-1})) \subset B_1^{-1}(\varphi_1(\ker B_2^{m-1})) = A_1^{-1}(\varphi_1(\ker A_2^{m-1})). \tag{3.64}$$

Hence

$$\varphi_1(\ker B_2^m) \subset U_m(A) \tag{3.65}$$

for all $B \in S$. We now claim there exists $M \in M_n(R)$ such that (1) $\omega M_2(\ker A_2^{m-1}) = 0$ and (2) $\varphi_1(\ker(A+\omega M)_2^m) = U_m(A)$. Clearly $(A+\omega M)_1 = A_1$, and since $\omega M_2(\ker A_2^{m-1}) = 0$ we have that $(A+\omega M)_2 v = A_2 v$ for all $v \in \ker A_2^{m-1}$, hence $A + \omega M \in S$. Hence the previous argument yields $\varphi_1(\ker(A+\omega M)_2^m) \subset U_m(A)$, so it remains to show we can choose M such that the reverse inclusion holds. Fix a basis for $U_m(A) \cap \varphi_1(\ker A_2^{m-1})$, and extend by elements $v_1, \ldots, v_t \in U_m(A) \setminus \ker A_1^{m-1}$ to a basis for all of $U_m(A)$. For each $1 \le i \le t$, choose a lift $\tilde{v_i} \in (R/\omega^2)^n$ such that $\varphi_1(\tilde{v_i}) = v_i$, and consider $A_2\tilde{v_i}$. Because $v_i \in U_m(A) \subset A_1^{-1}(\varphi_1(\ker A_2^{m-1}))$, there exists some $\tilde{u_i} \in (R/\omega^2)^n$ with $\varphi_1(\tilde{u_i}) = A_1 v_i$ and $\tilde{u_i} \in \ker A_2^{m-1}$. Hence $x_i := \tilde{u_i} - A_2\tilde{v_i} \in \omega(R/\omega^2)^n$. Define M by specifying the image of each basis element: let M annihilate $U_m(A) \cap \ker A_1^{m-1}$, and let $\omega M \tilde{v_i} = x_i$. Then

$$(A+\omega M)_2 \tilde{v_i} = A_2\tilde{v_i} + x_i = \tilde{u_i} \in \ker A_2^{m-1}, \tag{3.66}$$

so $\tilde{v_i} \in \ker A_2^m$. Hence $\{v_1, \ldots, v_t\} \in \varphi_1(\ker(A+\omega M)_2^m)$. We also have that

$$U_m(A) \cap \varphi_1(\ker A_2^{m-1}) \subset \varphi_1(\ker A_2^{m-1}) = \varphi_1(\ker(A+\omega M)_2^{m-1}) \subset \varphi_1(\ker(A+\omega M)_2^m), \tag{3.67}$$

because ωM annihilates $\ker A_2^{m-1}$. Therefore $U_m(A) \subset \varphi_1(\ker(A+\omega M)_2^m)$, so the two are equal.

Let $\hat{A} = A + \omega M$, and let $Q = \{B \in M_n(R) : (\omega B)_2 \ker \hat{A}_2^{m-1} = 0\}$. Then if Y is distributed with Haar measure conditional on $Y \in Q$, and X is distributed with Haar measure conditional on $X \in S$, we have that X and $\hat{A} + \omega Y$ are equal in distribution. Hence

$$\Pr(V_{2,m}(X) = W_m | X \in S) = \Pr(V_{2,m}(\hat{A} + \omega Y) = W_m | Y \in Q) \tag{3.68}$$

and we may compute the latter. If $B \in Q$, then $\hat{A}+\omega B \in S$, hence by (3.65) $\varphi_1(\ker(\hat{A}+\omega B)_2^m) \subset U_m(A)$. To find $V_{2,m}(\hat{A} + \omega B)$ it suffices to discover how much the reverse inclusion holds, i.e. which $v \in U_m(A)$ have a lift $\tilde{v}$ which lies in $\ker(\hat{A} + \omega B)_2^m$.

Let $v \in U_m(A)$. Since $U_m(A) = \varphi_1(\ker \hat{A}_2^m)$, there exists a lift $\tilde{v} \in (R/\omega^2)^n$ with $\varphi_1(\tilde{v}) = v$ and $\tilde{v} \in \ker \hat{A}_2^m$. Hence any lift of v is of the form $\tilde{v} + \omega u$ for some u, so $v \in \varphi_1(\ker(\hat{A} + \omega B)_2^m)$ if and only if there exists u such that $(\hat{A} + \omega B)_2^m(\tilde{v} + \omega u) = 0$. Because $\tilde{v} \in \ker \hat{A}_2^m$ and $\omega^2 = 0$, this is equal to

$$\omega \hat{A}_2^m u + \sum_{i=0}^{m-1} \omega \hat{A}^i B \hat{A}^{m-1-i} \tilde{v}. \tag{3.69}$$

However, because $\tilde{v} \in \ker \hat{A}_2^m$, $A_2^j \tilde{v} \in \ker \hat{A}_2^{m-1}$ for all $j \geq 1$, hence since $B \in Q$, all terms of the sum but one vanish and

$$\sum_{i=0}^{m-1} \omega \hat{A}^i B \hat{A}^{m-1-i} \tilde{v} = \omega \hat{A}^{m-1} B \tilde{v}. \tag{3.70}$$

Therefore $v \in \varphi_1(\ker(\hat{A} + \omega B)_2^m)$ if and only if

$$\omega \hat{A}^{m-1} B \tilde{v} \in \mathrm{Im}(\omega \hat{A}_2^m) \tag{3.71}$$

since u is arbitrary. Note that this condition depends only on v, not on $\tilde{v}$. Finally, this condition is equivalent to

$$(\omega B)_2 \tilde{v} \in (\ker(\omega \hat{A})_2^{m-1} + \mathrm{Im}(\omega \hat{A}_2)) \cap \omega (R/\omega^2)^n. \tag{3.72}$$

Since $(\omega B)_2 \tilde{v}$ depends only on v, not on $\tilde{v}$, and $\omega(R/\omega^2)^n \cong (R/\omega)^n$, we may view $(\omega B)_2$ as a map of vector spaces. Hence $\omega(R/\omega^2)^n/((\ker(\omega \hat{A})_2^{m-1} + \mathrm{Im}(\omega \hat{A}_2)) \cap \omega(R/\omega^2)^n)$ has the structure of an $\mathbb{F}_q$-vector space, and we will denote it by H. Let $\phi_B : U_m(A) \to H$ be the map given by composing this quotient with $(\omega B)_2$. Then by (3.72), we have that for $v \in U_m(A)$, $v \in \varphi_1(\ker(\hat{A} + \omega B)_2^m)$ if and only if $\phi_B(v) = 0$.

However, recall that we do not care about $\varphi_1(\ker(\hat{A} + \omega B)_2^m)$, but rather $V_{2,m}(\hat{A} + \omega_B) = A_1^{m-1}\varphi_1(\ker(\hat{A}+\omega B)_2^m) \subset A_1^{m-1} U_m(A)$. Furthermore, $A_1^{m-1} U_m(A)$ is isomorphic to $U_m(A)/(\ker A_1^{m-1} \cap U_m(A))$ by the map A_1^{m-1}. Since $B \in Q$, $(\omega B)_2$ annihilates $\varphi_1(\ker A_2^{m-1})$, hence it annihilates $\ker A_1^{m-1} \cap U_m(A)$. Therefore ϕ_B factors through $U_m(A)/(\ker A_1^{m-1} \cap U_m(A))$, which is isomorphic to $\Lambda_m \cap W_{m-1}$.

Consider the map $\phi'_B : \Lambda_m \cap W_{m-1} \to H$ given by factoring ϕ_B through $\Lambda_m \cap W_{m-1}$ as in the previous paragraph. We then have by (3.72) and the last few paragraphs that

$$V_{2,m}(\hat{A} + \omega B) = \ker \phi'_B. \tag{3.73}$$

If Y is distributed as before, with the Haar measure conditioned on $Y \in Q$, then ϕ'_Y is uniformly distributed among maps $\Lambda_m \cap W_{m-1} \to H$. Hence to find the distribution of $V_{2,m}(\hat{A} + \omega B)$ we must compute $\dim_{\mathbb{F}_q} H$.

Let us compute $\dim_{\mathbb{F}_q}(\ker(\omega\hat{A})_2^{m-1} + \mathrm{Im}(\omega\hat{A}_2)) \cap \omega(R/\omega^2)^n$. This space is isomorphic to $\mathrm{Im}A_1 + \ker A_1^{m-1}$, so it suffices to compute the dimensions of these two spaces and their intersection. We have

$$\dim \mathrm{Im}A_1 = n - \lambda_1 \tag{3.74}$$

$$\dim \ker A_1^{m-1} = \lambda_1 + \ldots + \lambda_{m-1} \tag{3.75}$$

$$\dim \mathrm{Im}A_1 \cap \ker A_1^{m-1} = \dim \ker A_1^m - \dim \ker A_1 = (\lambda_1 + \ldots + \lambda_m) - \lambda_1, \tag{3.76}$$

so $\dim \mathrm{Im}A_1 + \ker A_1^{m-1} = n - \lambda_m$. Hence $\dim H = n - (n - \lambda_m) = \lambda_m$.

Since ϕ'_Y is uniformly distributed among maps $\Lambda_m \cap W_{m-1} \to H$, we have

$$\Pr(V_{2,m}(\hat{A} + \omega Y) = W_m | Y \in Q) = \frac{|\{\text{linear maps } \phi : \Lambda_m \cap W_{m-1} \to H \text{ with } \ker \phi = W_m\}|}{q^{d_m \cdot \lambda_m}}. \tag{3.77}$$

A map ϕ with the appropriate kernel factors uniquely through an injective map $\Lambda_m \cap W_{m-1}/W_m \to H$. Since $\dim \Lambda_m \cap W_{m-1}/W_m \leq \dim \Lambda_m = \lambda_m$, the number of such injections is

$$(q^{\lambda_m} - 1) \cdots (q^{\lambda_m} - q^{(d_m - \mu_m)-1}) = q^{(d_m - \mu_m)\cdot\lambda_m} \frac{(q^{-1})_{\lambda_m}}{(q^{-1})_{\lambda_m - (d_m - \mu_m)}}. \tag{3.78}$$

Hence we finally have

$$\Pr(V_{2,m}(\hat{A} + \omega Y) = W_m | Y \in Q) = q^{-\mu_m \cdot \lambda_m} \frac{(q^{-1})_{\lambda_m}}{(q^{-1})_{\lambda_m - (d_m - \mu_m)}}, \tag{3.79}$$

and this is equal to (3.58). From this (3.60) follows immediately by taking a product of the probabilities in (3.58). □

Setting $m = 1$, the formula of Theorem 3.3.11 is exactly the same as the one in Theorem 3.3.4, as it should be. The following example specializes Theorem 3.3.11 to what is in some sense the simplest nontrivial case.

Example 3.3.12. Consider the case where Λ_1 is one-dimensional and $\Lambda_1 = \Lambda_2 = \ldots = \Lambda_t \supset \Lambda_{t+1} = 0$. Then the possible filtered vector spaces $V_2(X)$ are completely determined by the number of nonzero vector spaces in the filtration $V_{2,1}(X) \supset V_{2,2}(X) \supset \ldots$, and we refer to this number as $N_2(X)$. Then by Theorem 3.3.4 and Theorem 3.3.11, we have

$$\Pr(N_2(X) = r | V_1(X) = \Lambda) = \begin{cases} (1-1/q)\frac{1}{q^r} & 0 \leq r \leq t-1 \\ \frac{1}{q^t} & r = t \\ 0 & r > t \end{cases}. \tag{3.80}$$

As a sanity check, the sum of these probabilities is

$$(1-1/q) + \frac{1}{q}(1-1/q) + \ldots + \frac{1}{q^{t-1}}(1-1/q) + \frac{1}{q^t} = 1. \tag{3.81}$$

This defines a Markov chain on $\mathbb{Z}_{\geq 0}$, which may be used to define a measure on partitions by generating their parts as in Proposition 3.3.1, though it is a much simpler measure.

As mentioned, it seems difficult to extend Theorem 3.3.11 to compute the distribution of $V_{r+1,m}$ for $r > 1$ given the same data (X_r and $\ker X_{r+1}^{m-1}$), and we sketch why. The same method of proof carries over, but in place of $U_m(A)$ we have

$$\varphi_1(A_r^{-1}\varphi_r(\ker A_{r+1}^{m-1})) \tag{3.82}$$

and in place of $\mathrm{Im}A_1 + \ker A_1^{m-1}$ we obtain

$$(\mathrm{Im}A_r + \ker A_r^{m-1}) \cap \omega^{r-1}(R/\omega^r)^n. \tag{3.83}$$

Computing the dimension of either such space in general is difficult because we are no longer able to reduce to a problem about vector spaces over finite fields, and though for specific A it is doable, a general computation has proven elusive. It also appears that, unless there is some significant cancellation, the answers will depend on rows other than the r^{th} row, and we may hence lose the desired Markov property. It is possible that a modified definition of $V(A)$ would overcome these difficulties, or that the Markov property on rows is not the right analogue to look for. Nonetheless, the author believes that even if the definition must be modified, the ideas and techniques used in Theorem 3.3.11 may well prove useful. We conclude with two directions for future work.

Problem 3.3.13. *Compute the limiting $n \to \infty$ distribution of the plane partition $T(X)$, and/or $V(X)$, for $X \in M_n(R)$ with Haar measure.*

Theorem 3.3.11 represents partial progress toward this goal. Abstracting away from the random matrix models, Theorem 3.3.11 suggests a natural decreasing Markov chain on filtered vector spaces which generalizes the Markov chain on $\mathbb{N}$ studied in [24, 23, 22]. Since this Markov chain has some interesting algebraic combinatorics associated with it as mentioned earlier, we suggest the following.

Problem 3.3.14. *Define a Markov chain $(X_n)_{n\geq 1}$ on the set of all filtered $\mathbb{F}_q$-vector spaces as in Theorem 3.3.11, i.e. if $W \subset \Lambda$ are two filtered vector spaces and there is an injection of filtered vector spaces, then*

$$\Pr(X_{n+1} = W | X_n = \Lambda) = \prod_{m\geq 1} q^{-\mu_m \cdot \lambda_m} \frac{(q^{-1})_{\lambda_m}}{(q^{-1})_{\lambda_m - (d_m - \mu_m)}}, \tag{3.84}$$

where the dimensions λ_m, μ_m, d_m are as in Theorem 3.3.11. Study to what extent the properties of the Markov chain in [23] generalize to this one and whether additional interesting combinatorial structure arises. See whether there is a random matrix interpretation of this Markov chain, by modifying the definition of $V(A)$ if necessary.

References

[1] J. D. Achter. The distribution of class groups of function fields. *Journal of Pure and Applied Algebra*, 204(2):316–333, 2006.

[2] M. Adam. On the distribution of eigenspaces in classical groups over finite rings. *Linear Algebra and its Applications*, 443:50–65, 2014.

[3] G. E. Andrews. *The theory of partitions*. Number 2. Cambridge university press, 1998.

[4] M. Bhargava, D. M. Kane, H. W. Lenstra, B. Poonen, and E. Rains. Modeling the distribution of ranks, Selmer groups, and Shafarevich–Tate groups of elliptic curves. *Cambridge Journal of Mathematics*, 3(3):275–321, 2015.

[5] A. Borodin and V. Gorin. Lectures on integrable probability. In *Probability and Statistical Physics in St. Petersburg, Proceedings of Symposia in Pure Mathematics*, volume 91, pages 155–214, 2012.

[6] A. Borodin, A. Okounkov, and G. Olshanski. Asymptotics of Plancherel measures for symmetric groups. *Journal of the American Mathematical Society*, 13(3):481–515, 2000.

[7] D. Bump. *Lie groups*. Springer, 2004.

[8] J. W. S. Cassels. *Local fields*, volume 3. Cambridge University Press Cambridge, 1986.

[9] G. Chinta, N. Kaplan, and S. Koplewitz. The cotype zeta function of $\mathbb{Z}^d$. *arXiv preprint arXiv:1708.08547*, 2017.

[10] H. Cohen and H. W. Lenstra. Heuristics on class groups of number fields. In *Number Theory Noordwijkerhout 1983*, pages 33–62. Springer, 1984.

[11] J. B. Conrey. L-functions and random matrices. In *Mathematics unlimitedâĂŤ2001 and Beyond*, pages 331–352. Springer, 2001.

[12] P. Deift and D. Gioev. *Random matrix theory: invariant ensembles and universality*, volume 18. American Mathematical Soc., 2009.

[13] C. Delaunay. Heuristics on Tate-Shafarevitch groups of elliptic curves defined over Q. *Experimental Mathematics*, 10(2):191–196, 2001.

[14] F. J. Dyson. Statistical theory of the energy levels of complex systems I, II and III. *Journal of Mathematical Physics*, 3(1):140–156, 157–165, 166–175, 1962.

[15] S. N. Evans. Elementary divisors and determinants of random matrices over a local field. *Stochastic processes and their applications*, 102(1):89–102, 2002.

[16] C. Feng, R. W. Nóbrega, F. R. Kschischang, and D. Silva. Communication over finite-ring matrix channels. In *Information Theory Proceedings (ISIT), 2013 IEEE International Symposium on*, pages 2890–2894. IEEE, 2013.

[17] N. Fine, I. Herstein, et al. The probability that a matrix be nilpotent. *Illinois Journal of Mathematics*, 2(4A):499–504, 1958.

[18] F. W. Firk and S. J. Miller. Nuclei, primes and the random matrix connection. *Symmetry*, 1(1):64–105, 2009.

[19] E. Friedman and L. C. Washington. On the distribution of divisor class groups of curves over a finite field. *Théorie des Nombres/Number Theory Laval*, 1987.

[20] J. Fulman. A probabilistic approach toward conjugacy classes in the finite general linear and unitary groups. *Journal of Algebra*, 212(2):557–590, 1999.

[21] J. Fulman. A probabilistic approach to conjugacy classes in the finite symplectic and orthogonal groups. *Journal of Algebra*, 234(1):207–224, 2000.

[22] J. Fulman. The Rogers-Ramanujan identities, the finite general linear groups, and the Hall-Littlewood polynomials. *Proceedings of the American Mathematical Society*, 128(1):17–25, 2000.

[23] J. Fulman. A probabilistic proof of the Rogers–Ramanujan identities. *Bulletin of the London Mathematical Society*, 33(4):397–407, 2001.

[24] J. Fulman. Random matrix theory over finite fields. *Bulletin of the American Mathematical Society*, 39(1):51–85, 2002.

[25] J. Fulman. Cohen-Lenstra heuristics and random matrix theory over finite fields. *Journal of Group Theory*, 17(4):619–648, 2014.

[26] J. Fulman and N. Kaplan. Random partitions and Cohen-Lenstra heuristics. *arXiv preprint arXiv:1803.03722*, 2018.

[27] W. Fulton. Eigenvalues, invariant factors, highest weights, and Schubert calculus. *Bulletin of the American Mathematical Society*, 37(3):209–249, 2000.

[28] D. Goldfeld. Gauss' class number problem for imaginary quadratic fields. *Bulletin of the American Mathematical Society*, 13(1):23–37, 1985.

[29] V. Gorin and A. W. Marcus. Crystallization of random matrix orbits. *arXiv preprint arXiv:1706.07393*, 2017.

[30] P. R. Halmos. *Measure theory*, volume 18. Springer, 2013.

[31] A. Horn. Eigenvalues of sums of Hermitian matrices. *Pacific Journal of Mathematics*, 12(1):225–241, 1962.

[32] Q. Y. (https://mathoverflow.net/users/290/qiaochu yuan). Cohen-Lenstra heuristics reference. MathOverflow. URL:https://mathoverflow.net/q/147039 (version: 2013-11-05).

[33] N. Katz and P. Sarnak. Zeroes of zeta functions and symmetry. *Bulletin of the American Mathematical Society*, 36(1):1–26, 1999.

[34] J. P. Kung. The cycle structure of a linear transformation over a finite field. *Linear Algebra and its Applications*, 36:141–155, 1981.

[35] J. Lengler. The Cohen-Lenstra heuristic for finite abelian groups. *Doktorarbeit, Universität des Saarlandes, Saarbrücken, Germany*, 2009.

[36] I. G. Macdonald. *Symmetric functions and Hall polynomials*. Oxford university press, 1998.

[37] K. Maples. Cokernels of random matrices satisfy the Cohen-Lenstra heuristics. *arXiv preprint arXiv:1301.1239*, 2013.

[38] A. W. Marcus. Polynomial convolutions and (finite) free probability. *Preprint, available at* `https://web.math.princeton.edu/~amarcus/papers/ff_main.pdf`.

[39] A. Nica and R. Speicher. *Lectures on the combinatorics of free probability*, volume 13. Cambridge University Press, 2006.

[40] J. H. Silverman. *The arithmetic of elliptic curves*, volume 106. Springer Science & Business Media, 2009.

[41] R. P. Stanley. Theory and application of plane partitions: Part 1. *Studies in Applied Mathematics*, 50(2):167–188, 1971.

[42] R. Stong. Some asymptotic results on finite vector spaces. *Advances in applied mathematics*, 9(2):167–199, 1988.

[43] T. Tao. *Topics in random matrix theory*, volume 132. American Mathematical Soc., 2012.

[44] Y. Wang and R. P. Stanley. The Smith normal form distribution of a random integer matrix. *SIAM Journal on Discrete Mathematics*, 31(3):2247–2268, 2017.

[45] H. Weyl. The asymptotic distribution law of the eigenvalues of linear partial differential equations (with an application to the theory of cavity radiation). *Mathematical Annals*, 71(4):441–479, 1912.

[46] E. P. Wigner. On the statistical distribution of the widths and spacings of nuclear resonance levels. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 47, pages 790–798. Cambridge University Press, 1951.

[47] E. P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *Annals of Mathematics*, pages 548–564, 1955.

[48] E. P. Wigner. *Statistical properties of real symmetric matrices with many dimensions*. Princeton University, 1957.

[49] M. Wood. The distribution of sandpile groups of random graphs. *Journal of the American Mathematical Society*, 30(4):915–958, 2017.

[50] M. M. Wood. Random integral matrices and the Cohen-Lenstra heuristics. *arXiv preprint arXiv:1504.04391*, 2015.

[51] M. M. Wood. Asymptotics for number fields and class groups. In *Directions in Number Theory*, pages 291–339. Springer, 2016.

[52] Q. Yuan. Groupoid cardinality. https://qchu.wordpress.com/2012/11/08/groupoid-cardinality/, 2012.