RANDOM WORDS, QUANTUM STATISTICS, CENTRAL LIMITS, RANDOM MATRICES*

GREG KUPERBERG†

Abstract. Recently Tracy and Widom conjectured [29] and Johansson proved [17] that the expected shape λ of the semi-standard tableau produced by a random word in k letters is asymptotically the spectrum of a random traceless $k \times k$ GUE matrix. In this article we give two arguments for this fact. In the first argument, we realize the random matrix itself as a quantum random variable on the space of random words, if this space is viewed as a quantum state space. In the second argument, we show that the distribution of λ is asymptotically given by the usual local limit theorem, but the resulting Gaussian is disguised by an extra polynomial weight and by reflecting walls. Both arguments more generally apply to an arbitrary finite-dimensional representation V of an arbitrary simple Lie algebra $\mathfrak{g}$. In the original question, V is the defining representation of $\mathfrak{g} = \mathfrak{su}(k)$.

What is the longest weakly increasing subsequence of a long, random string of letters? In the previous sentence, one such longest subsequence is "AEEEEEEEFLNNOSTTT". In randomly chosen English text, the longest subsequences are dominated by the letter 'E', since this letter is the most common one. This implies that the length of the longest subsequence has a Gaussian distribution. But if the letters in the string are independent with the uniform distribution, a longest subsequence will use all of them roughly equally. In this case Tracy and Widom established a non-Gaussian distribution for the length of a longest subsequence [29, 16]. Their result was motivated by recent progress in the study of the longest increasing subsequence of a random permutation, in particular the relations among longest subsequences, random matrices, and representation theory [1, 3, 2, 30, 9, 8, 4, 25].

Tracy and Widom conjectured a generalization which was proved by Johansson [17, Th. 1.6]:

THEOREM 1 (Johansson). *The distribution of the shape of a random word as given by the Robinson-Schensted-Knuth (RSK) algorithm converges locally to the distribution of the spectrum of a random traceless $k \times k$ GUE matrix.*

It is a generalization because the first row of the RSK shape is the length of the longest weakly increasing subsequence. "Traceless GUE" refers to the traceless Gaussian unitary ensemble, defined up to normalization as the Gaussian measure on traceless $k \times k$ Hermitian matrices which is invariant under conjugation by unitary matrices.

In this article we give two arguments for Theorem 1. The first argument (Section 1) is based on quantum statistics: it identifies the random matrix itself as a quantum random variable on the space of random words viewed as a quantum state space. The GUE ensemble then appears in the limit by a quantum central limit theorem. The second argument (Section 2) is based on classical statistics: it identifies the density formula

$$(1) \qquad C \prod_{a \leq b} (\lambda_a - \lambda_b)^2 e^{-\sum_a \lambda_a^2} d\lambda$$

for the distribution of the spectrum λ of a GUE matrix [23] as a disguised classical central limit. (Here C is a constant that depends on k but not λ.) The classical argument is rigorous and it establishes a precise estimate. The quantum argument can be read rigorously or non-rigorously, depending on whether the reader accepts Conjecture 2; either way it is less precise than the statement of Theorem 1. Non-rigorously, it establishes convergence in distribution. Rigorously it establishes convergence of certain moments, but not enough moments to imply convergence in distribution. Nonetheless we prefer the quantum argument since it is less traditional. (But see Biane [6, 5, 7] for closely related results.)

In both arguments, it is important to identify the vector space of traceless Hermitian matrices with the Lie algebra $\mathfrak{su}(k)$ and an alphabet with k letters with the standard basis of the defining representation $V = \mathbb{C}^k$. Both arguments then generalize to an arbitrary finite-dimensional unitary representation of V of a compact simple Lie algebra $\mathfrak{g}$. The conclusion is a relation between random words in a weight basis of V and a natural Gaussian measure on $\mathfrak{g}^*$, the vector space dual of $\mathfrak{g}$.

1. Quantum statistics. In this section we will express certain classical random variables in terms of simpler quantum random variables. The main object in our argument was also considered from the converse view by Biane [5, 6].

We refer the reader to Sakurai [27, §3] for basic notions of quantum statistics, in particular *mixed states*, which are also commonly called density matrices or density operators. In the context of operator algebras, mixed states are called *states* [18] or *normal states* [19], depending on the desired strength of the formalism.

The *RSK algorithm* is (in one version) a function that takes as input a word of length N in the alphabet $[k] = \{1, ..., k\}$ and produces as output a pair of tableaux (P, Q) of shape λ, where

$$\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k) \vdash N$$

is a partition of N into non-increasing, non-negative integers [28, §7.11]. The partition is considered synonymous with its *Young diagram*, meaning its horizontal histogram. The tableau P is semi-standard and is called the *insertion tableau*, while the tableau Q is standard and is called the *recording tableau*. Given the uniform distribution on the set of words $[k]^N$, we can view the shape λ as a random variable λ_{RSK}. Finally, given a partition λ, it will sometimes be convenient to subtract the mean from each part to form a "partition of 0":

$$\widehat{\lambda} = (\lambda_1 - \frac{N}{k}, \lambda_2 - \frac{N}{k}, \ldots, \lambda_k - \frac{N}{k}).$$

We do not need the precise definition of the RSK algorithm in this section, merely one of its important properties: It is a combinatorial model for the direct sum decomposition of the representation $V^{\otimes N}$ of the Lie algebra $\mathfrak{u}(k)$ (or the Lie group $\mathrm{U}(k)$ or $\mathrm{GL}(k, \mathbb{C})$), where $V = \mathbb{C}^k$ is the defining representation [28, §A2]. This representation decomposes as

$$(2) \qquad\qquad V^{\otimes N} \cong \bigoplus_{\lambda \vdash N} R_\lambda \otimes V_\lambda,$$

where $V_{\widehat{\lambda}}$ is the irreducible representation of $\mathfrak{u}(k)$ of shape λ and R_λ is the irreducible representation of the symmetric group S_N of shape λ. For any given $\lambda = \lambda_{\mathrm{RSK}}$, the set of associated insertion tableaux P indexes a basis of V_λ, while the set of recording tableaux Q indexes a basis of R_λ. In particular,

$$\dim R_\lambda \otimes V_\lambda = n_\lambda,$$

where n_λ is the number of words that have shape $\lambda = \lambda_{\mathrm{RSK}}$. Finally, as a representation of the Lie subalgebra $\mathfrak{su}(k)$, V_λ is unchanged if we add a constant to each component of λ. The convention is to call it $V_{\widehat{\lambda}}$, the representation of highest weight is $\widehat{\lambda}$.

We can view the vector space $V^{\otimes N}$ as a quantum state space $\mathcal{H}$ of some quantum system Q with $[k]^N$ as an orthonormal basis. The maximum-entropy state (or *tracial* state) ρ of Q is then realized by the uniform distribution on $[k]^N$, as well as by the uniform distribution on any other orthonormal basis. At the same time, an arbitrary orthogonal direct sum decomposition

$$\mathcal{H} \cong \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \ldots \oplus \mathcal{H}_t$$

of $\mathcal{H}$ can be interpreted as a random variable taking values in the set of summands. Relative to the state ρ, the probability of a given summand $\mathcal{H}_i$ is the ratio $(\dim \mathcal{H}_i)/(\dim \mathcal{H})$. In particular, the direct sum decomposition in equation (2) expresses a random variable $\lambda_{\mathrm{QM}} = \widehat{\lambda}$. The previous paragraph tells us that $\lambda_{\mathrm{QM}} \overset{\mathrm{d}}{=} \widehat{\lambda}_{\mathrm{RSK}}$, meaning that they have the same distribution.

1.1. The case $k = 2$ and spin 1/2 particles. As a concrete example, we consider the physically realizable case $k = 2$. In this case V is the familiar state space of a spin $1/2$ particle, and the action of SU(2) is the projective action of the spatial rotation group SO(3). We will use the alphabet $\{\uparrow, \downarrow\}$ rather than $\{1, 2\}$ as a basis of V. The space V admits angular momentum operators J_x, J_y, and J_z which satisfy the commutation relations

$$(3) \qquad [J_x, J_y] = iJ_z \qquad [J_y, J_z] = iJ_x \qquad [J_z, J_x] = iJ_y.$$

The operators J_x, J_y, and J_z are a basis of $i \cdot \mathfrak{su}(2)$, by which we mean the image of $\mathfrak{su}(2)$ in $\mathfrak{sl}(2, \mathbb{C})$ under multiplication by i. Thus these are just the usual commutation relations in the Lie algebra $\mathfrak{su}(2)$ up to a factor of i. The tracial state on V is the mixture

$$(4) \qquad \rho = \frac{|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|}{2}.$$

Note that probabilistic mixtures of states should not be confused with quantum superpositions. A superposition of $|\uparrow\rangle$ and $|\downarrow\rangle$ is another vector in V and cannot be invariant under rotations. By contrast the mixed state ρ is SU(2)-invariant.

The vector space $V^{\otimes N}$ is then the state space of N such particles. The tth particle has angular momentum operators $J_x^{(t)}$, $J_y^{(t)}$, and $J_z^{(t)}$. By equation (4) and the isotropy of ρ, each of these operators is a centered Bernoulli random variable with equally likely values $\frac{1}{2}$ and $-\frac{1}{2}$. Since the three operators for any fixed t do not commute, the corresponding random variables cannot be simultaneously observed. The sums of these operators form the total angular momentum,

$$(5) \qquad J_\alpha = J_\alpha^{(1)} + J_\alpha^{(2)} + \ldots + J_\alpha^{(N)},$$

for each $\alpha \in \{x, y, z\}$. Each operator J_α is a centered binomial random variable because the terms are independent commuting Bernoulli variables. The three operators J_x, J_y, and J_z do not commute either, but rather satisfy the same commutation relations in equation (3), since they express the natural (diagonal) action of $\mathfrak{su}(2)$ on $V^{\otimes N}$. Finally, the total angular momentum

$$J^2 = J_x^2 + J_y^2 + J_z^2$$

is diagonalized by the direct sum decomposition in equation (2). In a summand with weight $\widehat{\lambda}$, its eigenvalue is

$$J^2|\psi\rangle = \widehat{\lambda}_1(\widehat{\lambda}_1 + 1)|\psi\rangle.$$

Thus if $L = (L_1, L_2)$ is the shape-valued operator that measures λ_{QM}, it is related to J^2 by

$$J^2 = L_1(L_1 + 1).$$

If we define the scaled angular momentum operators

$$\widetilde{J}_\alpha = \frac{J_\alpha}{\sqrt{N}}$$

with $\alpha \in \{x, y, z\}$, then by the above reasoning, these three operators become commuting Gaussian random variables in the limit $N \to \infty$. By rotational symmetry they are also independent and identically distributed (i.i.d.). This behavior of the total angular momentum of independent random spins can be witnessed in nuclear magnetic resonance experiments, among other places. Using these operators we may form a matrix

$$\widetilde{M} = \begin{pmatrix} \widetilde{J}_z & \widetilde{J}_x + i\widetilde{J}_y \\ \widetilde{J}_x - i\widetilde{J}_y & -\widetilde{J}_z \end{pmatrix}.$$

In the limit $N \to \infty$, $\widetilde{M}$ becomes a traceless GUE matrix! (The normalization is also consistent with Mehta [23].) For finite N, the determinant of $\widetilde{M}$ must be interpreted carefully because its entries do not commute. If we define it by averaging over orderings of the entries,

$$\det \widetilde{M} = \frac{1}{2}(\widetilde{M}_{11}\widetilde{M}_{22} + \widetilde{M}_{22}\widetilde{M}_{11} - \widetilde{M}_{12}\widetilde{M}_{21} - \widetilde{M}_{21}\widetilde{M}_{12}),$$

then it turns out that

$$\det \widetilde{M} = -\widetilde{J}^2.$$

It follows that

$$\lim_{N \to \infty} \frac{\lambda_{\mathrm{QM}}}{\sqrt{N}} \overset{\mathrm{d}}{=} \lambda_{\mathrm{GUE}},$$

where λ_{GUE} is a random variable representing the spectrum λ of a traceless GUE matrix. This is precisely Theorem 1 for $k = 2$.

1.2. The general case. The argument in Section 1.1 generalizes with only natural changes to all values of k. The defining representation V of $\mathfrak{su}(k)$ has a basis of states

$$|1\rangle, |2\rangle, \ldots, |k\rangle.$$

The elements of $i \cdot \mathfrak{su}(k)$ may be viewed as generalized angular momentum operators. We define two matrices of operators A and B whose entries linearly span $i \cdot \mathfrak{su}(k)$. (Note that the diagonal entries are not linearly independent.) Let $E_{ab} \in \mathfrak{sl}(k, \mathbb{C})$ be the elementary matrix whose non-zero entry is

$$(E_{ab})_{ab} = 1.$$

Then the entries of A and B are

$$A_{ab} = \frac{1}{2}(E_{ab} + E_{ba}) \qquad\qquad B_{ab} = \frac{i}{2}(E_{ba} - E_{ab}) \qquad (a \neq b)$$

$$A_{aa} = E_{aa} - \frac{1}{k}I \qquad\qquad B_{aa} = 0.$$

Let M be the matrix of operators

$$M_{ab} = A_{ab} + iB_{ab}.$$

Each entry A_{ab} and B_{ab} is a real-valued measurement operator. Relative to the tracial state on V, A_{ab} for $a \neq b$ takes each of the values $1/2$ and $-1/2$ with probability $1/k$ and otherwise has the value 0. The same is true of B_{ab}. The measurement A_{aa} takes the value $(k-1)/k$ with probability $1/k$ and otherwise takes the value $-1/k$. The operator M_{ab} may appear to be a complex-valued measurement whose real and imaginary parts have these distributions, but this is not quite true. An operator can only be interpreted as a complex-valued measurement if it is a normal operator, defined as an operator that commutes with its adjoint, or equivalently an operator whose self-adjoint and anti-self-adjoint parts commute. When $a \neq b$, M_{ab} is not a normal operator; it represents a complex random variable whose real and imaginary parts are not simultaneously observable.

For words in $[k]^N$, we consider the standard (additive) action of $\mathfrak{su}(k)$ on $V^{\otimes N}$. Equation (2) gives us a shape-valued operator L which has the eigenvalue $\lambda_{\mathrm{QM}} = \widehat{\lambda}$ on the summand $V_\lambda \otimes R_\lambda$. The operator L can be realized algebraically using the characteristic polynomial of M, thought of as a polynomial-valued operator:

$$
\begin{aligned}
C(x) &= \det(xI - M) \\
&= (x - L_1)(x - L_2)\ldots(x - L_k) + c(x, L),
\end{aligned}
$$

(6)

where $c(x, L)$ is a polynomial of total degree at most $k - 1$. As before, each term of the determinant is defined by averaging over the $k!$ orderings of its factors. The left side of equation (6) is a disguised version of the composition $\psi \circ \pi$ in the proof of Harish-Chandra's Theorem given by Humphreys [15, §23.3], while the leading term on the right side is a disguised version of the map θ. Here we are applying both maps to the coefficients of the ordinary characteristic polynomial of an element in $\mathfrak{su}(k)$; in the context of Humphreys, each coefficient is a particular SU(k)-invariant polynomial on $\mathfrak{su}(k)$. As Humphreys explains, the maps $\psi \circ \pi$ and θ agree in the top degree, which is exactly what equation (6) asserts. (See Okounkov and Olshanski [26] for an analysis of the correction term $c(x, L)$.)

Each coefficient of $C(x)$ lies in the center of $U(\mathfrak{su}(k))$ and is a natural generalization of the Casimir operator J^2 in the case $k = 2$. The coefficients are sometimes called elementary generalized Casimir operators.

Assuming the tracial state on $V^{\otimes N}$, each measurement A_{ab} and B_{ab} is a sum of bounded, centered i.i.d. random variables. If we define

$$\widetilde{M} = \widetilde{A} + i\widetilde{B} = \sqrt{\frac{k}{2N}} M,$$

then the entries commute in the limit $N \to \infty$ and $\widetilde{M}$ becomes a traceless GUE matrix with standard normalization. The term $c(x, L)$ in equation (6) also disappears in this limit because its degree is too low. The equation thus tells us that

$$\lim_{N \to \infty} \sqrt{\frac{k}{2N}} \lambda_{\mathrm{QM}} \overset{\mathrm{d}}{=} \lambda_{\mathrm{GUE}}.$$

1.3. What did we prove?. One important step in the argument of this section is not completely rigorous. Unquestionably each scaled angular momentum operator $\widetilde{J}_\alpha$, $\widetilde{A}_{ab}$, or $\widetilde{B}_{ab}$ converges to a Gaussian random variable by the classical central limit theorem. But we do not know that a polynomial in these variables, for example $\widetilde{J}^2$ or $\widetilde{J}_x\widetilde{J}_y\widetilde{J}_z$, converges in distribution to the corresponding polynomial of Gaussian variables. We cannot appeal to the classical multivariate central limit theorem for non-commuting variables, even if they do commute in the limit. There are also several quantum limit theorems in the literature; one of the most general ones is due to Goderis, Verbeure, and Vets [11]. But these results are apparently not sufficiently strong either.

As a stop-gap we will conjecture the quantum central limit theorem that we need, and we will prove a weak version of the conjecture. The conjecture is naturally stated in terms of C^*-algebras and von Neumann algebras [18, 19], which provide a rigorous language for quantum statistics. In this language, a non-commutative probability space is defined as a von Neumann algebra $\mathcal{M}$ with a normal state ρ. A *state* is defined as a dual vector on $\mathcal{M}$, continuous in the norm topology, with the interpretation that for a self-adjoint element A, $\rho(A)$ is the expected value of the random variable given by A; the state is *normal* if it is continuous in the weak topology as well. The reader who is uninterested in operator algebras can take

$$\mathcal{M} = \mathcal{M}_k = \mathfrak{gl}(k, \mathbb{C}),$$

the vector space of $k \times k$ matrices. A normal state ρ on $\mathcal{M}_k$ is any dual vector whose matrix is Hermitian and positive semi-definite and has trace 1; this is exactly the definition in physics of a finite density matrix [27, §3]. In particular, the tracial state is defined by

$$\rho(A) = \frac{1}{k} \operatorname{Tr} A.$$

Given two quantum systems with von Neumann algebras $\mathcal{M}$ and $\mathcal{N}$, the joint system has the algebra $\mathcal{M} \otimes \mathcal{N}$, using the tensor product in the category of von Neumann algebras. Two normal states ρ and ω on $\mathcal{M}$ and $\mathcal{N}$ form an independent joint state $\rho \otimes \omega$. Given self-adjoint operators $A \in \mathcal{M}$ and $B \in \mathcal{N}$, the operators $A \otimes I$ and $I \otimes B$ represent independent measurements in the joint system. If $A \in \mathcal{M}$ is self-adjoint, let

$$A^{(t)} = I^{\otimes t-1} \otimes A \otimes I^{\otimes N-t-1} \in \mathcal{M}^{\otimes N},$$

and let

$$\widetilde{A} = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} A^{(t)}.$$

Thus $\widetilde{A}$ expresses the scaled sum of N i.i.d. random variables.

In formulating a multivariate quantum central limit theorem, three issues arise because of non-commutativity. First, the theorem must be a statement about the distribution of non-commutative polynomials $p \in \mathbb{C}\langle A_1, \ldots, A_k \rangle$, but a Gaussian central limit would describe the distribution of commuting variables. Second, a general polynomial expression $p(A_1, \ldots, A_k)$ need not be a self-adjoint operator, even if the variables are self-adjoint. Thus we will assume that p is a *self-adjoint polynomial*, meaning that it is invariant under the anti-linear anti-involution

$$* : \mathbb{C}\langle A_1, \ldots, A_k \rangle \to \mathbb{C}\langle A_1, \ldots, A_k \rangle$$

that conjugates each coefficient and reverses the order of each term. Third, if we define the covariance matrix of the variables $A_1, \ldots, A_k$ as

$$\kappa_{ab} = \rho(A_a A_b),$$

it may not be symmetric. When this happens the behavior of the central limit is genuinely different from the classical case; it has been studied in Reference [11]. But in the case that we need (namely, when ρ is tracial), the covariance matrix is symmetric. In this case we expect that the limiting distribution of p only depends on its commutative image $\widehat{p} \in \mathbb{C}[A_1, \ldots, A_k]$, thereby resolving the first issue.

CONJECTURE 2. *Let $(\mathcal{M}, \rho)$ be a quantum probability space, and let $A_1, \ldots, A_k$ be self-adjoint elements with mean 0 and a symmetric covariance matrix. Let $p \in \mathbb{C}\langle A_1, \ldots, A_k \rangle$ be a self-adjoint non-commutative polynomial in k variables. Then*

$$\lim_{N \to \infty} p(\widetilde{A}_1, \ldots, \widetilde{A}_k) \stackrel{\mathrm{d}}{=} \widehat{p}(X_1, \ldots, X_k),$$

where $X_1, \ldots, X_k$ are classical Gaussian random variables with covariance matrix

$$E[X_a X_b] = \rho(A_a A_b).$$

If we let p be a coefficient of the polynomial $C(x)$ from Section 1, Conjecture 2 then implies Theorem 1. It may be possible to be reverse this reasoning and use Theorem 1 to prove Conjecture 2 for arbitrary p, at least when $\mathcal{M}$ is a matrix algebra and ρ is the tracial state. But a satisfactory proof would hold for arbitrary quantum probability spaces.

Our Theorem 3 below establishes convergence of moments in the context of Conjecture 2. Since this theorem is almost entirely algebraic, we do not need the full structure of a von Neumann algebra. Rather we let $\mathcal{M}$ be an arbitrary $*$-algebra, meaning a unital ring over the complex numbers with an anti-linear anti-involution $*$. A state ρ on a $*$-algebra is a $*$-invariant dual vector such that $\rho(I) = 1$ and $\rho(A^2) \geq 0$ for every self-adjoint $A \in \mathcal{M}$. Finally the nth moment

$$\gamma_n(A) = \rho(A^n)$$

is defined whether or not A is self-adjoint. (Recall that a non-self-adjoint operator may be written as $A + iB$, where A and B are self-adjoint. Consequently it may be interpreted as a complex-valued quantum random variable whose real and imaginary parts are not simultaneously observable. Our definition of moments is consistent with this interpretation.)

THEOREM 3. *Suppose that $\mathcal{M}$ a $*$-algebra with a state ρ. Let $A_1, \ldots, A_k$ be self-adjoint elements with mean 0. Suppose that for all a and b,*

$$\rho(A_a A_b) = \rho(A_b A_a).$$

If $p \in \mathbb{C}\langle A_1, \ldots, A_k \rangle$ is a non-commutative polynomial, then

$$\lim_{N \to \infty} \gamma_n(p(\widetilde{A}_1, \ldots, \widetilde{A}_k)) = \gamma_n(\widehat{p}(X_1, \ldots, X_k)),$$

where $X_1, \ldots, X_k$ are classical centered Gaussian random variables with covariance matrix

$$E[X_a X_b] = \rho(A_a A_b),$$

and $\gamma_n(A)$ is the nth moment of A.

Proof. Since the assertion is claimed for every polynomial, it suffices to prove that the expectation

$$\rho^{\otimes N}(p(\widetilde{A}_1, \ldots, \widetilde{A}_k))$$

converges. To show convergence of expectation we may let p be a monomial. Indeed the monomial

$$p(A_1, \ldots, A_k) = A_1 A_2 \ldots A_k$$

will do, since some of the factors may be equal.

Expanding the expression

$$\gamma = \rho^{\otimes N}(\widetilde{A}_1, \ldots, \widetilde{A}_k)$$

using the definition of $\widetilde{A}_a$, it has a term for each function ϕ from $[k]$ to $[N]$:

$$\gamma = N^{-k/2} \sum_{\phi:[k] \to [N]} \rho^{\otimes N}(\prod_a A_a^{(\phi(a))}).$$

Let

$$\gamma_\phi = N^{-k/2} \rho^{\otimes N}(\prod_a A_a^{(\phi(a))})$$

be an individual term in this expansion. Since we are computing the expectation with respect to the product state, we can arrange the factors with respect to $[N]$ rather than $[k]$:

$$\gamma_\phi = N^{-k/2} \prod_{t \in [N]} \rho(\prod_{\phi(a)=t} A_a^{(t)}).$$

In this form it is clear that

$$\gamma_\phi = 0$$

if there is a t such that $\phi^{-1}(t)$ has one element. At the same time, if S is the set of those functions ϕ whose images have fewer than $k/2$ elements, then

$$\lim_{N\to\infty} \sum_{\phi\in S} \gamma_\phi = 0$$

because

$$|S| = o(N^{k/2}).$$

In other words, there are sufficiently few such functions ϕ that they are negligible in the limit. What is left is the set of functions that are exactly 2-to-1, which only exist when k is even. Thus if M is the set of perfect matchings of $[k]$, then

$$\gamma = N^{-k/2} \binom{N}{k/2} (k-1)!!(k/2)! \sum_{m\in M} \prod_{(a,b)\in m} \rho(A_a A_b)$$
$$+ o(1)$$
$$= (k-1)!! \sum_{m\in M} \prod_{(a,b)\in m} \rho(A_a A_b) + o(1),$$

where

$$(k-1)!! = (k-1)(k-3)\cdots 5\cdot 3\cdot 1$$

is the odd factorial function. In the limit γ exactly matches the corresponding expectation

$$E[X_1 \ldots X_k] = (k-1)!! \sum_{m\in M} \prod_{(a,b)\in m} E[X_a X_b]$$

of the classical variables $X_1, \ldots, X_k$. □

In relation to Theorem 1, Theorem 3 says that if

$$C_\lambda(x) = (x-\lambda_1)(x-\lambda_2)\ldots(x-\lambda_k)$$

for a partition λ, then the moments of $C_{\widehat{\lambda}_{\mathrm{RSK}}}(x)$ converge to the moments of $C_{\lambda_{\mathrm{GUE}}}(x)$. In other words, the "moments of the moments" of λ_{RSK} converge after scaling to those of λ_{GUE}. Unfortunately, when $k > 2$ the tail of the distribution of $C_{\lambda_{\mathrm{GUE}}}(x)$ is too thick for convergence of moments to imply convergence in distribution.

2. Local limits. The argument of this section and its generalization below (Theorem 6) are very similar to a result of Biane [7]. It was also found by Grinstead [14] in the case $k = 2$, and it is related to some results of Grabiner [12].

The idea of the argument is that, if we name the dimensions appearing in equation (2),

$$f_\lambda = \dim R_\lambda \qquad d_\lambda = \dim V_\lambda \qquad n_\lambda = f_\lambda d_\lambda,$$

the quantity f_λ can be considered in the context of the *k-ballot problem*. Suppose N voters vote sequentially for an ordered list of k candidates. In how many ways can they cast their votes so that the ath candidate is never ahead of the bth candidate for $a > b$, and at the end the ath candidate has λ_a votes for every a? Such a sequence of votes is a *ballot sequence* of shape λ and there are f_λ of them [28, Prop. 7.10.3]. In this context of the RSK algorithm, f_λ is the number of standard tableaux of shape λ. If the tth entry of such a tableau is in row a, we can say that the tth voter votes for candidate a. This establishes a bijection between standard tableaux and ballot sequences. That f_λ is the number of standard tableaux of shape λ can also be seen directly from the representation theory of $\mathfrak{gl}(k)$: The generalized Clebsch-Gordan rule states that

$$V \otimes V_\lambda = \bigoplus_{\lambda' = \lambda + \square} V_{\lambda'},$$

where the sum is over shapes λ' that are obtained from λ by adding a single box. Thus the multiplicity of V_λ in $V^{\otimes N}$ is the number of increasing chains of partitions from the empty partition to the partition λ. Such a chain is equivalent to a standard tableau of λ by assigning t to a box if it appears at step t.

In this formulation, the number f_λ may be computed by the *reflection principle* [10], which is also a disguised version of the Weyl character formula [13]. Recognizing the set of partitions as a subset of $\mathbb{Z}^k$, there is an action of the symmetric group S_k on $\mathbb{Z}^k$ given by permuting coordinates. A chain of partitions is then a lattice path in $\mathbb{Z}^k$ that happens to stay in the cone of partitions. Here a valid lattice path is one which increases one coordinate by one at each step. Consider the partition

$$\delta = (k-1, k-2, k-3 \ldots, 2, 1, 0)$$

and let m_λ for any $\lambda \in \mathbb{Z}^k$ be the number of lattice paths from the origin to λ. The reflection principle shows that

$$(7) \qquad\qquad f_\lambda = \sum_{\sigma \in S_k} (-1)^\sigma m_{\lambda + \delta - \sigma(\delta)}$$

Equation (7) says that the number of ballot sequences from 0 to λ is the alternating sum of unrestricted lattice paths from a set of image points of the form $\sigma(\delta) - \delta$ to λ. Figure 1 shows an example of the principle when $k = 3$; in the figure, partitions λ are replaced by $\widehat{\lambda}$ to obtain walks in a 2-dimensional lattice.

Since the numbers m_λ are defined by walks with the same k possible steps at each time t, they can be approximated by the local central limit theorem:

$$(8) \qquad\qquad m_\lambda \sim C e^{-k\widehat{\lambda}^2/2N}$$

Here and below we assume that C is a constant depending only on N and k, and we use the notation

$$\lambda^2 = \sum_{a=1}^{k} \lambda_a^2.$$

If the approximation (8) were robust with respect to local finite differences, then by the reflection principle it would give us an estimate for f_λ. If it were also robust under amplification by a polynomial in λ, it would give us an estimate for

$$n_\lambda = f_\lambda d_\lambda,$$
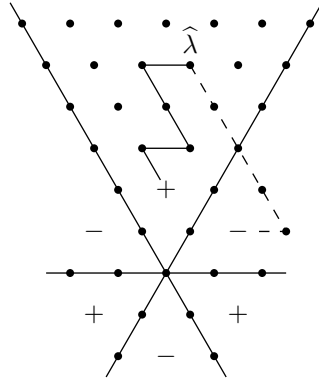
FIGURE 1. *Paths from $\widehat{\lambda}$ to 0 and an image point in the weight lattice of $\mathfrak{su}(3)$.*

since the Weyl dimension formula [15, §24.3] says that

$$(9) \qquad d_\lambda = \prod_{a>b} \frac{\lambda_a - \lambda_b + a - b}{a - b}$$

is a polynomial in λ. Both of these refinements of the local central limit theorem are true for arbitrary bounded lattice walks. To state the theorem, we introduce a few definitions.

A *finite difference operator* D is a linear operator on functions

$$p : \mathbb{R}^k \to \mathbb{R}$$

defined by a finite sum

$$Dp(v) = \sum_t c_t f(v + v_t)$$

for some constants $\{c_t\}$ and some vectors $\{v_t\}$. The *degree* a of D is the minimum degree of a polynomial p such that $Dp \neq 0$. If $L \subset \mathbb{R}^k$ is a lattice, the determinant

$$\det L = \operatorname{Vol} \ \mathbb{R}^k / L$$

is defined as the volume of the quotient space, or equivalently as the determinant of a positive basis for L.

THEOREM 4. *Let X be a bounded, mean 0 random variable taking values in $z + L$ for some lattice $L \subset \mathbb{R}^k$ and some vector $z \in \mathbb{R}^k$. Assume that L is the thinnest such lattice for the given X. Let*

$$X' = \sum_{t=1}^{N} X^{(t)}$$

denote the sum of N independent copies of X. Let

$$p(v) = P[X' = v]$$
$$q(v) = \frac{\det L}{(2\pi)^{k/2}\sqrt{\det \kappa}} e^{-\kappa^{-1}(v,v)/2N},$$

where κ is the covariance form of X. Then for every finite difference operator D of degree a and for every integer $b \geq 0$,

$$\lim_{N \to \infty} N^{(k-b+a)/2} |v|^b D(p-q)(v) = 0$$

uniformly for $v \in Nz + L$.

Lawler [21, Th. 1.2.1] proves a special case of Theorem 4 in which a and b are 0 or 2, X has the uniform distribution among nearest-neighbor steps in $\mathbb{Z}^k$, and D has a restricted form. However, the proof actually establishes Theorem 4 in its almost its full generality, requiring only that b be even. The conclusion for b odd follows by taking the geometric mean of the formulas for $b - 1$ and $b + 1$.

Since f_λ is given by the hook-length formula [28, Cor. 7.21.6], we can also prove Theorem 4 in this special case using Stirling's approximation [17, §4]. Such a special argument is analogous to the special argument for the Laplace-de Moivre theorem, which is the simplest case of the usual central limit theorem. But it is not enough for our later generalization, Theorem 6.

COROLLARY 5. *If $\lambda \vdash N$, then*

$$\lim_{N \to \infty} N^{k/2} \left(k^{-N} n_\lambda - C \prod_{a<b} (\lambda_a - \lambda_b)^2 e^{-k\widehat{\lambda}^2/2N} \right) = 0$$

uniformly in λ.

Proof. We apply Theorem 4. First, we change λ from a subscript to an argument in certain quantities that depend on it (and implicitly on N):

$$f(\widehat{\lambda}) = f_\lambda \qquad m(\widehat{\lambda}) = m_\lambda \qquad n(\widehat{\lambda}) = n_\lambda,$$

where dependence on N is implicit in the notation. Let $L = \Lambda$ be the set of all centered partitions $\widehat{\lambda}$ (the weight lattice). Define a finite difference operator D by

$$Dp(\widehat{\lambda}) = \sum_{\sigma \in S_k} (-1)^\sigma p(\widehat{\lambda} + \widehat{\delta} - \sigma(\widehat{\delta}))$$

so that

$$Dm(\widehat{\lambda}) = f(\widehat{\lambda})$$

by equation (7). The two important properties of the operator D are first, that it is antisymmetric under the Weyl group S_k after translation by $\widehat{\delta}$, and second, that it has degree $k(k-1)/2$. The degree of D follows from a factorization that appears in proofs of the Weyl dimension formula [15, §24.3]:

$$D = \prod_{\alpha \in \Phi_+} D_\alpha,$$

where Φ_+ is the set of positive roots of $\mathfrak{su}(k)$ and

$$D_\alpha(\widehat{\lambda}) = p(\widehat{\lambda}) - p(\widehat{\lambda} + \alpha).$$

Each D_α has degree 1 and there are $k(k-1)/2$ of them. Note that the only antisymmetric polynomial of degree $k(k-1)/2$ is

$$\Delta = \prod_{a<b} (\widehat{\lambda}_a - \widehat{\lambda}_b) = \prod_{a<b} (\lambda_a - \lambda_b),$$

When N is large,

$$De^{-k\widehat{\lambda}^2/2N} \sim C\Delta e^{-k\widehat{\lambda}^2/2N}$$

because in the limit D becomes an antisymmetric differential operator of degree $k(k-1)/2$. When applied to a symmetric Gaussian, it produces an anti-symmetric polynomial factor of degree $k(k-1)/2$. The polynomial Δ is the only choice for this factor up to scale. Thus the operator D explains one factor of Δ in the statement of the corollary. The other factor is given by d_λ, which is also proportional to Δ in the limit as $\lambda \vdash N$ goes to ∞. Theorem 4 then establishes the stated approximation for n_λ, where D and d_λ each contribute a factor of Δ. $\square$

Corollary 5 is evidently the precise statement of Theorem 1.

3. Generalizations. The first way that we can generalize Theorem 1 that is that we can replace the representation V of $\mathfrak{su}(k)$ by some other finite-dimensional representation W. In general a tensor power of such a representation decomposes as

$$(10) \qquad W^{\otimes N} \cong \bigoplus_{\lambda \vdash 0} T_{\lambda,N} \otimes V_\lambda,$$

where each $T_{\lambda,N}$ is a vector space on which $\mathfrak{su}(k)$ acts trivially. (In this generality it does not make sense to make λ a partition of N or any other particular integer, so we take it to be a highest weight, or a partition of 0.) The space $T_{\lambda,N}$ is a representation of the symmetric group S_N, but it is not usually irreducible, not even when W is. Assuming the a state on $W^{\otimes N}$ which is invariant under the action of $\mathfrak{su}(k)$, we may as before use equation (10) to define a quantum random variable λ_{QM}.

It is less trivial to define a classical counterpart λ_{RSK}, or even the space of words on which it is defined. If W is irreducible, we can model it as a summand of $V^{\otimes \ell}$ for some ℓ. More precisely, we choose a partition $\mu \vdash \ell$ such that $W \cong V_{\widehat{\mu}}$, and we choose a specific standard tableau Q_W of shape μ. Then the set $S \subset [k]^\ell$ of words with recording tableau Q_W indexes a basis of W. The set S can be interpreted as a "syllabic alphabet", in the sense that a word w of length N over the alphabet S is simultaneously a word of length ℓN over the alphabet $[k]$. Remarkably, the RSK algorithm is compatible with this dual interpretation: If we define the shape λ of $w \in S^N$ by spelling it out in $[k]^{\ell N}$ and taking the usual shape, then once again

$$\dim T_{\lambda,N} \otimes V_\lambda = n_\lambda,$$

where n_λ is the number of words w with shape λ. (One way to argue this fact is with the theory of Littelmann paths; see below. Syllabic expansion corresponds to concatenation of paths.)

For example, if $k = 2$, then $V^{\otimes 2}$ has a summand $W = V_2$ isomorphic to the adjoint representation of $\mathfrak{su}(2)$. As it happens, this summand occurs only once. If we take left and right parentheses $\{\,)\,,\,(\,\}$ as the alphabet for the basis of V rather than $\{1, 2\}$, then the first component $\widehat{\lambda}_1$ of the centered shape of a string is half the number of unmatched parentheses. For example, $\widehat{\lambda}_1 = 1$ for the string

The alphabet S for the representation W is the set of three pairs of parentheses $\{\,((\,,\,)(\,,\,))\,\}$ other than the two that match each other. If we rename this alphabet

$\{\,\langle\,,\,|\,,\,\rangle\,\}$, then one can check that the only words that match completely are those that form nested complete "bra-kets":

$$\underline{\langle\,|\,|\,\underline{\langle\,|\,\rangle}\,|\,\rangle}$$

A general string will have a maximal substring of this form, as well as fragments consisting of unmatched "bras", unmatched "kets", and unbracketed separators:

$$\langle\,|\,|\,|\qquad|\,|\,|\,\rangle\qquad|\,|\,|\,|$$

The statistic $\widehat{\lambda}_1$ is then the number of these fragments. For example, $\widehat{\lambda}_1 = 2$ for the string

$$\underline{|\,|}\,\underline{\langle\,|\,|\,\rangle}\,\underline{\langle\,|\,|}$$

since there are four unmatched parentheses if the bra-kets are expanded into parentheses:

$$\underline{)\,\overset{\frown}{(\,)}\,(}\;\underline{\overset{\overgroup{\frown}}{(\,\overset{\frown}{(\,)}\,\overset{\frown}{(\,)}\,)}}\,,\;\underline{\overset{\frown}{(\,)}\,\overset{\frown}{(\,)}\,(}$$

More generally still, we can let W be any non-trivial, finite-dimensional, unitary representation of any compact simple Lie algebra $\mathfrak{g}$. (We say that a Lie algebra is *compact* if it integrates to a compact Lie group.) Once again there is a direct sum decomposition

$$W^{\otimes N} \cong \bigoplus_{\lambda \in \Lambda} T_{\lambda,N} \otimes V_\lambda,$$

where Λ is the weight lattice of $\mathfrak{g}$ and V_λ is the irreducible representation of highest weight λ. As before this decomposition defines a quantum random variable λ_{QM} if we assume the tracial state ρ on W.

If W is irreducible, the theory of Littelmann paths then provides a satisfactory combinatorial counterpart λ_{LP} with the same distribution as λ_{QM} [22]. If $W \cong V_\mu$, then we can apply the Littelmann lowering operators to some fixed dominant path p_μ from the origin to μ. There is a natural bijection between the resulting set of paths $P(W)$ and a basis of V_μ. Moreover, a word w in P_μ forms a longer path $\gamma(w)$ given by concatenating letters. If we apply Littelmann raising operators as many times as possible to $\gamma(w)$, the result is a highest weight λ. Assuming the uniform distribution on P_μ, this weight defines a random variable random variable $\lambda_{\mathrm{LP}} = \lambda$. Note that λ_{LP} depends on p_μ, although its distribution does not.

Although the abstract setting of Littelmann paths looks quite different from the Robinson-Schensted-Knuth algorithm, it is in fact a strict generalization [31]. Briefly, if $W = V$ is the defining representation of $\mathfrak{su}(k)$ and p_μ is a straight line segment, then every element of P_μ is a straight line segment, and these segments are naturally enumerated by the integers $1, \ldots, k$. The highest weight λ_{LP} of a word w coincides with the centered shape of the tableau produced by a *dual* RSK algorithm defined using column insertion. (The standard RSK algorithm uses row insertion.) By one of the symmetries of the RSK algorithm [28, Cor. A1.2.11], this shape is the same as the row-insertion shape of the reverse word w^*. Thus

$$\lambda_{\mathrm{LP}}(w) = \widehat{\lambda}_{\mathrm{RSK}}(w^*) \qquad \lambda_{\mathrm{LP}} \overset{\mathrm{d}}{=} \widehat{\lambda}_{\mathrm{RSK}}.$$

Finally, if W is not irreducible, then we can choose a separate alphabet for each summand in a direct-sum decomposition. For example, if

$$W \cong V_\mu \oplus V_\mu$$

for some dominant weight μ, then we can let the alphabet be the disjoint union of two copies of the same alphabet P_μ, a "red" copy and a "blue" copy. The fundamental properties of Littelmann paths imply that in all cases, $\lambda_{\mathrm{QM}} \overset{\mathrm{d}}{=} \lambda_{\mathrm{LP}}$, where

$$P[\lambda_{\mathrm{QM}} = \lambda] = \frac{\dim T_{\lambda,N} \otimes V_\lambda}{(\dim W)^N}.$$

What is the counterpart to λ_{GUE} in this context? Taking Section 1 as a guide, each element $A \in i \cdot \mathfrak{g}$ defines a real-valued random variable on the quantum space (W, ρ). These variables have the covariance form

$$\kappa_W(A, B) = \rho(AB) = \frac{\mathrm{Tr}_W(AB)}{\dim W}.$$

All together they can be taken as 1-dimensional projections of a quantum random variable $x \in i \cdot \mathfrak{g}^*$. Instead of the spectrum, we can consider the orbit of ix under the co-adjoint action of $\mathfrak{g}$ on $i \cdot \mathfrak{g}^*$. By standard representation theory, ix is conjugate to a unique weight

$$\lambda \in \mathfrak{C} \subset \mathfrak{h}^* \subset \mathfrak{g}^*,$$

where $\mathfrak{C}$ is a Weyl chamber in $\mathfrak{h}^*$, the dual space to a Cartan subalgebra of $\mathfrak{g}$. If we assume a Gaussian distribution μ_W on $i \cdot \mathfrak{g}^*$ with covariance matrix κ_W, the corresponding distribution for the weight λ is

$$e_W(\lambda)d\lambda = C \prod_{\alpha \in \Phi_+} \alpha(\lambda)^2 e^{-\kappa_W^{-1}(\lambda,\lambda)/2} d\lambda,$$

where as before Φ_+ is the set of positive roots in $\mathfrak{g}$. This distribution can be derived in the same way as equation (1). If $\mathfrak{g} = \mathfrak{sp}(2n)$ (the compact form of $\mathfrak{sp}(2n, \mathbb{C})$), it is the Gaussian symplectic ensemble (GSE). But if $\mathfrak{g} = \mathfrak{so}(n)$, it is the Gaussian antisymmetric ensemble (GAE). Since symmetric matrices do not form a Lie algebra, the Gaussian orthogonal ensemble (GOE) would require some yet more general model.

Finally we can state the general theorem.

THEOREM 6. *Let W be a non-trivial, finite-dimensional, unitary representation of a compact simple Lie algebra $\mathfrak{g}$ of rank r. Let n_λ be the dimension of the isotypic summand of $W^{\otimes N}$ of highest weight λ. Then*

$$\lim_{N \to \infty} r^{N/2} \Big(\frac{n_\lambda}{(\dim W)^N} - C e_W\big(\frac{\lambda}{\sqrt{N}}\big) \Big) = 0$$

uniformly in λ.

The arguments of Sections 1 and 2 both generalize in a straightforward way to proofs of Theorem 6. As before, Section 1 establishes a weak version of it at a rigorous level. We also comment that the tautological matrix M of Section 1 should be replaced by a certain $i \cdot \mathfrak{g}^*$-valued measurement operator

$$M \in \mathfrak{g}^* \otimes \mathfrak{g}$$

acting on $W^{\otimes N}$. As a tensor in $\mathfrak{g}^* \otimes \mathfrak{g}$, M is again tautological; it comes from the identity linear transformation from $\mathfrak{g}$ to itself.

REMARK. The Lie algebra picture of Theorem 6 suggests another interpretation which is dual to that of Section 1, and in another sense dual to that of Section 2. If G is a compact, simple Lie group with Lie algebra $\mathfrak{g}$ and W is a unitary representation of G, then the absolute value of the character χ_W of W has a local maximum at $1 \in G$. When N is large, the character χ_W^N of $W^{\otimes N}$ is approximately a Gaussian in a neighborhood of 1. If we inflate G by a factor of $\sqrt{N}$, it converges to $\mathfrak{g}$, and multiplication on G converges to addition on $\mathfrak{g}$. The character χ_W^N converges to a limit on $\mathfrak{g}$, namely the Fourier transform of the Gaussian distribution μ_W on $\mathfrak{g}^*$ defined above. This intermediate picture led the author from Section 2 to Section 1.

3.1. Things out of reach. When $\mathfrak{g} = \mathfrak{su}(k)$, Theorem 6 can be interpreted as a limit distribution result for the shape λ_{RSK} of words with various interesting distributions. For example, if each letter of the alphabet for the representation V_2 of $\mathfrak{su}(2)$ is expanded into a pair of letters in the alphabet $\{\uparrow, \downarrow\}$, then the distribution ρ on expanded words is determined by its correlations for *digraphs* (adjacent pairs of letters). In this case the correlation between the tth and $t + 1$st letter depends on whether t is odd or even. But random words associated with representations such as $V_2 \oplus V_3$ do not exhibit such irregularities.

Especially when $k = 2$, these distributions resemble distributions given by *doubly stochastic Markov chains*. In other words, the first letter w_1 of a random word $w \in [k]^N$ has the uniform distribution. Each subsequent letter depends on the immediate predecessor (but not on earlier letters) according to a Markov matrix M:

$$P[w_{t+1} = a \mid w_t = b] = M_{ab}.$$

Here M is chosen so that every letter has the uniform distribution if the first one does.

What is the asymptotic distribution of the shape λ_{RSK} of a random word $w \in [k]^N$ generated by a doubly stochastic Markov matrix M? Non-rigorously we expect it to have the form

$$CP(\widehat{\lambda}) e^{-k\widehat{\lambda}^2/2vN}.$$

Here P is some polynomial (or at least some function which is asymptotically polynomial) and v is the *variance per letter* of w. The variance v is defined by the formula

$$v = \sum_{t=-\infty}^{\infty} \frac{kP[w_0 = w_t] - 1}{k - 1}$$

using a bi-infinite word w generated by M.

We have conducted computer experiments with different choices of M with 2- and 3-letter alphabets [20]. Figure 2 shows the distribution of $\widehat{\lambda}_1$ for 400,000 words generated by each of the following four Markov matrices M:

$$A = \frac{1}{4}\begin{pmatrix} 1 & 2 & 1 \\ 3 & 1 & 0 \\ 0 & 1 & 3 \end{pmatrix} \qquad F = \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

$$C_+ = \frac{1}{4}\begin{pmatrix} 3 & 0 & 1 \\ 1 & 3 & 0 \\ 0 & 1 & 3 \end{pmatrix} \qquad C_- = \frac{1}{4}\begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 1 & 0 & 3 \end{pmatrix}$$

The lengths of the words are 1620, 3420, 1140, and 1140 in the four respective cases. These lengths were chosen so that the four types of words would have the same total variance (ignoring boundary effects). The experiments indicate that the distribution of $\widehat{\lambda}_1$ (the centered length of the longest weakly increasing subsequence) in the asymmetric distribution A is genuinely different from the referent uniform distribution F. The lower median value of $\widehat{\lambda}_1$ in this case does not disappear as the words grow longer. It also cannot be explained as a maladjusted variance, because at the other end the tail of A eventually overtakes the tail of F. On the other hand, the distribution for the cyclic Markov chains C_+ and C_- do appear to converge to the distribution for F. Their symmetry implies that the longest weakly increasing subsequence sees the same fluctuations in the transition $1 \to 2$ as it does for the transition $2 \to 3$, which is apparently enough to produce the same distribution.
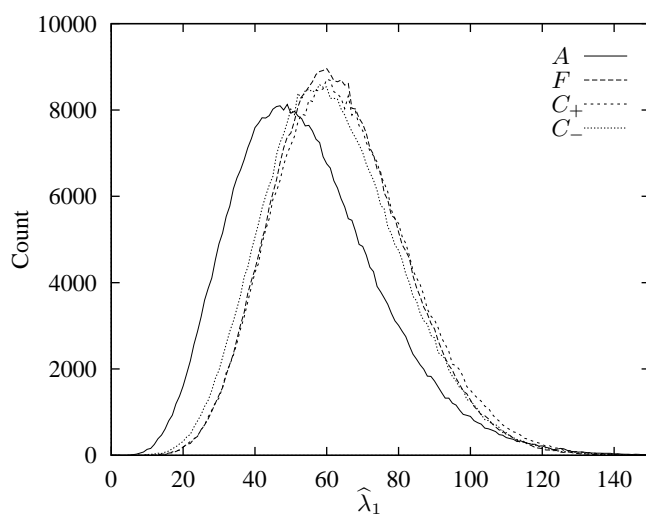

FIGURE 2. *Distribution of $\widehat{\lambda}_1$, the centered length of the longest increasing subsequence, for words generated by four different Markov chains, 400,000 trials each.*

CONJECTURE 7. *Let M be an indecomposable, doubly stochastic matrix such that*

$$M_{a,b} = M_{a+1,b+1},$$

where $k + 1 \equiv 1$. Then the distribution of the shape of a word of length N generated by M converges locally to the distribution of the spectrum of a traceless $k \times k$ GUE matrix.

Conjecture 7 can be generalized further by considering other cyclically symmetric, translation-invariant measures on words whose correlations decay sufficiently quickly.

PROBLEM 8. *Let $V = \mathbb{C}^k$ and let ρ be the state on $V^{\otimes n}$ extending a distribution on $[k]^N$ generated by a doubly stochastic Markov chain. What is the limiting distribution of λ_{QM}?*

Problem 8 is really a statistical mechanics question concerning a quantum spin chain with certain nearest-neighbor interactions. It cannot be stated in terms of λ_{RSK} because there is no reason to expect that $\widehat{\lambda}_{\mathrm{RSK}} \overset{\mathrm{d}}{=} \lambda_{\mathrm{QM}}$ in this generality. Yet more

generally, we can ask about the behavior of λ_{QM} for an arbitrary nearest-neighbor interaction that produces the tracial state when restricted to a single site.

PROBLEM 9. *What is the distribution of the longest weakly increasing circular subword of a circular word $w \in [k]^N$?*

In Problem 9, we assume that both $[k]$ and $[N]$ are circularly ordered. We do not know if there is a suitable circular generalization of the RSK algorithm.

3.2. Infinite matrices. The most interesting case of Theorem 1 to consider (indeed the case that motivated the result) is the limit $k \to \infty$. We have no firm results about this limit, but we can propose a model of it that may be important. Our model might be related to the semi-infinite wedge space model of Okounkov [24].
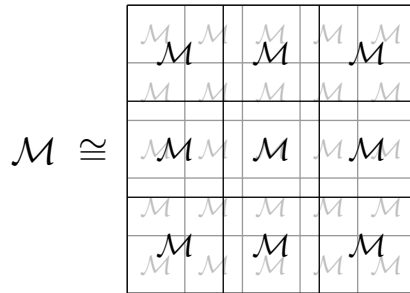

FIGURE 3. *The hyperfinite* II_1 *factor* $\mathcal{M}$ *as a matrix algebra over itself.*

Consider the Hilbert space $\mathcal{H} = L^2([0, 1])$. For every k, the matrix algebra $\mathcal{M}_k$ acts on $\mathcal{H}$ by taking

$$E_{ab}(f)(x) = f(x + \frac{b - a}{k})$$

if f is supported on $[\frac{b-1}{k}, \frac{b}{k}]$, and

$$E_{ab}(f)(x) = 0$$

if f vanishes on $[\frac{b-1}{k}, \frac{b}{k}]$. The weak-operator closure of all of these algebra actions is a von Neumann algebra $\mathcal{M}$ called the *hyperfinite* II_1 *factor* [19, §12.2]. For every k, $\mathcal{M}$ is a $k \times k$ matrix algebra over itself (Figure 3). Thus it could be generally important in random matrix theory.

In this case we are interested in the Lie algebra structure of $\mathcal{M}$ (in addition to its topologies), making $\mathcal{M}$ an infinite-dimensional analogue of $\mathfrak{gl}(k, \mathbb{C})$. We are also interested in the tracial state ρ, defined as the continuous extension of the tracial state on each $\mathcal{M}_k$. It is a normal state. It is also a model of the uniform measure on the interval $[0, 1]$. Finally we define $\widehat{\mathcal{M}}$ to be the kernel of ρ, analogous to the space $\mathfrak{sl}(k, \mathbb{C})$ of traceless matrices.

As a Lie algebra, $\mathcal{M}$ acts on $\mathcal{H}^{\otimes N}$ for every N. This action commutes with the action of the symmetric group S_N on $\mathcal{H}^{\otimes N}$ given by permuting tensor factors. There is a direct-sum decomposition

$$(11) \qquad\qquad \mathcal{H}^{\otimes N} \cong \bigoplus_{\lambda \vdash N} \mathcal{H}_\lambda,$$

where $\mathcal{H}_\lambda$ is, as a representation of S_N, the isotypic summand of type R_λ. Each of these representations has a measure-theoretic dimension defined using $\rho^{\otimes N}$ on $\mathcal{H}^{\otimes N}$ (in which $\mathcal{H}$ embeds by the usual Leibniz rule for Lie algebra actions):

$$\dim \mathcal{H}^{\otimes N} = 1 \qquad \dim \mathcal{H}_\lambda = \frac{f_\lambda^2}{N!}.$$

Thus equation (11) is a quantum statistics model for the Plancherel measure on the symmetric group. For each N, it defines a quantum random variable $\lambda_{\mathrm{QM}} \vdash N$.

The state $\rho^{\otimes N}$ also expresses the uniform measure on $[0,1]^N$, *i.e.*, the process of choosing a "word" of N random points in the unit interval. The usual RSK algorithm is defined for such words. Since the letter of the word are distinct almost surely, and since the RSK algorithm depends only on the order of the letters and not their values, it defines a random variable λ_{RSK} equivalent to the shape of a random permutation. Its distribution is also the Plancherel measure.

By the quantum central limit theorem, the state $\rho^{\otimes N}$ should produce a Gaussian measure on $\widehat{\mathcal{M}}$ in the limit $N \to \infty$. So should the GUE measure on $\mathfrak{sl}(k, \mathbb{C})$ in the limit $k \to \infty$. The relation between these two limits could shed light on the Vershik-Kerov limit for Plancherel measure [32] and the Wigner semicircle for the spectrum of a random matrix [23]. The quantum central limit theorem might also predict the distribution of the deviation from a semicircle, at least to first order in N.

REFERENCES

[1] David Aldous and Persi Diaconis, *Longest increasing subsequences: from patience sorting to the Baik-Deift-Johansson theorem*, Bull. Amer. Math. Soc. (N.S.), 36 (1999), pp. 413–432.

[2] Jinho Baik, Percy Deift, and Kurt Johansson, *On the distribution of the length of the second row of a Young diagram under Plancherel measure*, arXiv:math.CO/9901118.

[3] Jinho Baik, Percy Deift, and Kurt Johansson, *On the distribution of the length of the longest increasing subsequence of random permutations*, J. Amer. Math. Soc., 12 (1999), pp. 1119–1178, arXiv:math.CO/9810105.

[4] Jinho Baik and Eric M. Rains, *Algebraic aspects of increasing subsequences*, arXiv:math.CO/9905083.

[5] Philippe Biane, *Quantum random walk on the dual of* SU(n), Probab. Theory Related Fields, 89 (1991), pp. 117–129.

[6] Philippe Biane, *Some properties of quantum Bernoulli random walks*, Quantum probability and related topics, World Sci. Publishing, 1991, pp. 193–203.

[7] Philippe Biane, *Estimation asymptotique des multiplicités dans les puissances tensorielles d'un* $\mathfrak{g}$-*module*, C. R. Acad. Sci. Paris Sér. I Math., 316 (1993), pp. 849–852.

[8] Alexei Borodin, Andrei Okounkov, and Grigori Olshanski, *Asymptotics of Plancherel measures for symmetric groups*, arXiv:math.CO/9905032.

[9] Alexei Borodin and Grigori Olshanski, *Z-measures on partitions, Robinson-Schensted-Knuth correspondence, and β=2 random matrix ensembles*, arXiv:math.CO/9905189.

[10] Ira M. Gessel and Doron Zeilberger, *Random walk in a Weyl chamber*, Proc. Amer. Math. Soc., 115 (1992), pp. 27–31.

[11] D. Goderis, A. Verbeure, and P. Vets, *Non-commutative central limits*, Probab. Theory Related Fields, 82 (1989), pp. 527–544.

[12] David J. Grabiner, *Brownian motion in a Weyl chamber, non-colliding particles, and random matrices*, Ann. Inst. H. Poincaré Probab. Statist., 35 (1999), pp. 177–204, arXiv:math.RT/9708207.

[13] David J. Grabiner and Peter Magyar, *Random walks in Weyl chambers and the decomposition of tensor powers*, J. Algebraic Combin., 2 (1993), pp. 239–260.

[14] Charles Grinstead, unpublished note.

[15] James E. Humphreys, *Introduction to Lie algebras and representation theory*, Graduate Texts in Mathematics, vol. 9, Springer-Verlag, New York-Heidelberg-Berlin, 1972.

[16] Alexander R. Its, Craig A. Tracy, and Harold Widom, *Random words, Toeplitz determinants and integrable systems. I*, arXiv:math.CO/9909169.

[17] Kurt Johansson, *Discrete orthogonal polynomial ensembles and the Plancherel measure*, arXiv:math.CO/9906120.

[18] Richard V. Kadison and John R. Ringrose, *Fundamentals of the theory of operator algebras*, vol. I, Academic Press, 1983.

[19] Richard V. Kadison and John R. Ringrose, *Fundamentals of the theory of operator algebras*, vol. II, Academic Press, 1986.

[20] Greg Kuperberg, *makewords.c*, included with the source of arXiv:math.PR/9909104.

[21] Gregory F. Lawler, *Intersections of random walks*, Birkhäuser, 1991.

[22] Peter Littelmann, *Paths and root operators in representation theory*, Ann. of Math., (2) 142 (1995), pp. 499–525.

[23] Madan Lal Mehta, *Random matrices*, 2 ed., Academic Press, 1991.

[24] Andrei Okounkov, *Infinite wedge and measures on partitions*, arXiv:math.RT/9907127.

[25] Andrei Okounkov, *Random matrices and random permutations*, arXiv:math.CO/9903176.

[26] Andrei Okounkov and Grigori Olshanski, *Shifted Schur functions*, St. Petersburg Math. J., 9 (1998), pp. 239–300, arXiv:q-alg/9605042.

[27] Jun John Sakurai, *Modern quantum mechanics*, 2 ed., Benjamin/Cummings, 1985.

[28] Richard P. Stanley, *Enumerative combinatorics*, vol. 2, Cambridge University Press, 1999.

[29] Craig A. Tracy and Harold Widom, *On the distributions of the lengths of the longest monotone subsequences in random words*, arXiv:math.CO/9904042.

[30] Craig A. Tracy and Harold Widom, *Random unitary matrices, permutations and Painlevé*, arXiv:math.CO/9811154.

[31] Marc A. A. van Leeuwen, private communication.

[32] Anatoly Vershik and Sergei Kerov, *Asymptotic behavior of the Plancherel measure of the symmetric group and the limit form of Young tableaux*, Dokl. Akad. Nauk SSSR, 233 (1977), pp. 1024–1027.