Representation Theory of Big Groups and Probability.

Lecture Notes. Draft

Leonid Petrov

Fall 2012

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Contents

1 Introduction **4**
1.1 LLN and CLT . . . 4
1.2 Characteristic functions . . . 5
1.3 De Moivre-Laplace theorem . . . 7
1.4 Other distributions . . . 8
1.4.1 Normal . . . 8
1.4.2 Poisson distribution . . . 9
1.4.3 Geometric, exponential, ... 10
1.5 Conclusions . . . 10
1.6 Where now? . . . 10
1.7 Some nice pictures on slides . . . 11
1.8 Interlacing arrays and representations . . . 11
1.8.1 Unitary groups and their representations. Signatures . . . 11
1.8.2 Branching rule . . . 11
1.8.3 Branching rule and Gelfand–Tsetlin basis in π_N^λ . . . 12
1.8.4 Gelfand–Tsetlin schemes. Interlacing arrays . . . 13
1.9 References . . . 14

2 De Finetti's Theorem **15**
2.1 Formulation . . . 15
2.2 Remarks on de Finetti's theorem . . . 16
2.2.1 Bernoulli sequences . . . 16
2.2.2 Ergodicity . . . 17
2.2.3 Hausdorff moment problem . . . 17
2.3 Pascal triangle . . . 18
2.3.1 Definition . . . 18
2.3.2 Harmonic functions . . . 19
2.3.3 Coherent systems of measures. Boundary . . . 19
2.3.4 Approximation method . . . 20
2.3.5 Proof of de Finetti's theorem (= description of the boundary of the Pascal triangle) . . . 21
2.4 Algebraic structure under de Finetti's theorem . . . 23
2.5 References . . . 26

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

3 Irreducible characters of finite and compact groups **27**
3.1 Basic definitions . . . 27
3.2 Characters . . . 28
3.3 Examples . . . 30
3.3.1 Characters of finite symmetric and finite-dimensional unitary groups . . . 30
3.3.2 Big groups . . . 31
3.4 References . . . 32

4 Symmetric polynomials and symmetric functions. Extreme characters of $U(\infty)$ and $\mathfrak{S}(\infty)$ **33**
4.1 Symmetric polynomials, Laurent-Schur polynomials and characters of unitary groups . . . 33
4.1.1 Laurent symmetric polynomials . . . 33
4.1.2 Anti-symmetric polynomials . . . 34
4.1.3 Schur polynomials . . . 35
4.1.4 Orthogonality of Schur polynomials . . . 36
4.1.5 Characters of the unitary groups. Radial part of the Haar measure . . . 37
4.1.6 Branching of Schur polynomials and restriction of characters 40
4.1.7 Combinatorial formula for Schur polynomials . . . 42
4.1.8 Semistandard Young tableaux . . . 42
4.1.9 Elementary symmetric polynomials . . . 43
4.1.10 Two examples of extreme characters of $U(\infty)$ . . . 44
4.1.11 Extreme characters of $U(\infty)$ . . . 46
4.1.12 Remark: Dimensions of irreducible representations of $U(N)$ 50
4.1.13 Proof of the Edrei-Voiculescu theorem . . . 52
4.2 Symmetric functions. Properties. Characters of symmetric groups 54
4.2.1 Symmetric polynomials. Stability. Algebra of symmetric functions . . . 54
4.2.2 Multiplicative bases in *Sym*: e- and h-functions . . . 56
4.2.3 Power sums . . . 58
4.2.4 Cauchy identity . . . 59
4.2.5 Jacobi-Trudi identity . . . 60
4.2.6 Pieri formula. Standard Young tableaux . . . 60
4.2.7 Frobenius characteristic map. Characters of the symmetric grous $\mathfrak{S}(n)$ . . . 61
4.2.8 Branching of irreducible characters of symmetric groups . 62
4.2.9 Young graph . . . 63
4.2.10 Classification theorem for extreme characters of $\mathfrak{S}(\infty)$ . . 64
4.2.11 Specizations of the algebra of symmetric functions. 'Ring theorem' . . . 65
4.3 References . . . 67

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

5 Totally nonnegative sequences, planar networks, Geseel-Viennot determinants 69
5.1 Unified picture of extreme characters of $U(\infty)$ and $\mathfrak{S}(\infty)$ and connection to totally nonnegative sequences . . . 69
5.1.1 Totally nonnegative matrices: definition and examples . . 69
5.1.2 Unitary group . . . 70
5.1.3 Symmetric group . . . 74
5.1.4 Conclusion . . . 76
5.2 Gessel-Viennot determinants . . . 76
5.3 Example with Catalan-Hankel determinants . . . 79
5.4 Totally nonnegative Toeplitz matrices via Gessel-Viennot theory 81
5.5 Schur polynomials via Gessel-Viennot theory . . . 85
5.6 References . . . 88

6 Determinantal point processes 89
6.1 Generalities [Bor11a] . . . 89
6.1.1 Point processes . . . 89
6.1.2 Correlation measures and correlation functions . . . 89
6.1.3 Determinantal processes . . . 90
6.1.4 Remarks . . . 91
6.1.5 Examples [Bor11b] . . . 92
6.2 Biorthogonal and orthogonal polynomial ensembles . . . 93
6.2.1 Biorthogonal ensembles [Bor11b] . . . 93
6.2.2 Orthogonal polynomial ensembles [Bor11b] . . . 95
6.3 Examples: CUE, GUE, extreme characters of $U(\infty)$ . . . 98
6.3.1 CUE . . . 98
6.3.2 GUE . . . 98
6.3.3 Extreme coherent systems on the Gelfand-Tsetlin graph . 99
6.4 References . . . 102

7 Asymptotics of determinantal point processes 103
7.1 Difference/differential operators . . . 103
7.1.1 Gamma-measures and formula for the kernel . . . 103
7.1.2 Rodrigue's formula . . . 104
7.1.3 Charlier polynomials . . . 106
7.1.4 Plancherel measures . . . 107
7.1.5 Plancherel measures as limits of $M_N^{(\gamma)}$ . . . 108
7.1.6 Discrete Bessel kernel . . . 110
7.1.7 Limit behavior of the Plancherel measures . . . 111
7.1.8 Ulam's problem . . . 114
7.1.9 More general limit shapes: measures $M_N^{(\gamma)}$ . . . 114
7.1.10 GUE; bulk and edge asymptotics (and the semicircle law): differential operators approach . . . 117
7.2 Double contour integral kernels and their asymptotic analysis. Saddle points method . . . 117
7.2.1 Asymptotics of the Plancherel measure . . . 117

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Chapter 1

Introduction

In this introductory lecture we try to give an impression on what the course is about. We start with a well-known probability problem, namely, the Central Limit Theorem, and then discuss more complicated models. We aim to explain that problems/models formulated in probabilistic terms may contain some underlying structure which helps to deduce interesting properties of these problems.

1.1 LLN and CLT

We begin with recalling the well-known theorems in Probability

Theorem 1.1.1 (Law of large numbers)**.** *If $X_1, X_2, \ldots$ are iid (independent identically distributed) random variables, and $\mathbb{E}\, X_1 = \mu$ (finite first moment), then*

$$\frac{X_1 + \ldots + X_N}{N} \to \mu$$

in probability and almost surely.[1]

The words "in probability" mean that for any fixed $\varepsilon > 0$, we have

$$\mathbb{P}\left\{ \left| \frac{X_1 + \ldots + X_N}{N} - \mu \right| > \varepsilon \right\} \to 0, \qquad N \to \infty.$$

[1] This "almost surely" claim is the so-called Strong Law of large numbers, and we will not discuss it here. It may obtained from the ordinary Law of large numbers via the Borel-Cantelli Lemma.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Theorem 1.1.2 (Central limit theorem)**.** *If $X_1, X_2, \ldots$ are iid (independent identically distributed) random variables, $\mathbb{E}\, X_1 = \mu$, and $\operatorname{Var} X_1 = \sigma^2 > 0$ (two finite moments), then*

$$\frac{X_1 + \ldots + X_N - N\mu}{\sqrt{N\sigma^2}} \to \mathcal{N}(0,1),$$

in distribution (=weakly), where $\mathcal{N}(0,1)$ abbreviates the standard normal distribution.

This convergence in distribution can be formulated in a number of ways; for us it is enough to say that for any fixed $x \in \mathbb{R}$,

$$\mathbb{P}\left\{\frac{X_1 + \ldots + X_N - N\mu}{\sqrt{N\sigma^2}} \le x\right\} \to \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz.$$

The integrand in the right-hand side here is the bell-shaped standard normal probability density (Fig. 1.1).

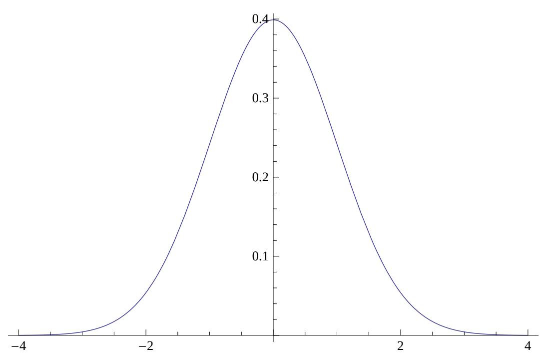


Figure 1.1: Normal density

Of course there exist more precise statements than Theorems 1.1.1 and 1.1.2, but let us here discuss the basic ones.

1.2 Characteristic functions

LLN (Theorem 1.1.1) may be proved using the Chebyshev inequality which is a quite straightforward way. The CLT (Theorem 1.1.2) may be proved using the so-called method of moments.

Here let us briefly outline the way of proof of Theorems 1.1.1 and 1.1.2 using *characteristic functions*. This is "modern technique" which is late XIX and early XX century.

Definition 1.2.1. If X is a real random variable, its *characteristic function* is defined by

$$\varphi_X(t) = \mathbb{E}\, e^{\mathrm{i}tX}, \qquad t \in \mathbb{R},$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

where $\mathrm{i} = \sqrt{-1}$.

Characteristic function always exists (EX),[2] and, moreover, defines the random variable uniquely. Furthermore,

Theorem 1.2.2 (Continuity theorem)**.** *If for a sequence of random variables Y_N, the characteristic functions φ_{Y_N} converge (pointwise) to a characteristic function of, say, random variable Y, then the random variables Y_N themselves converge to Y in distribution.*

Remark 1.2.3. Usually, the claim of the above theorem is relaxed by saying that the functions $\varphi_{Y_N}(t)$ must converge pointwise and the limit function must be continuous at $t = 0$. Then it is in fact a characteristic function itself.

We also use the fact that

$$\mathbb{E}(XY) = \mathbb{E}\,X \cdot \mathbb{E}\,Y$$

for independent random variables.

Proof of LLN (Theorem 1.1.1). We write:

$$\varphi_{(X_1+\ldots+X_N)/N}(t) = [\varphi_{X_1}(t/N)]^N.$$

For the random variable X_1 having the first moment μ, we have (EX)

$$\varphi_{X_1}(t) = 1 + \mathrm{i}t\mu + o(t), \qquad t \to 0;$$

so we have for any fixed t:

$$\varphi_{(X_1+\ldots+X_N)/N}(t) = [1 + \mathrm{i}t\mu/N + o(1/N)]^N \to e^{\mathrm{i}t\mu}.$$

The right-hand side is a characteristic function of the constant, and so we establish the convergence in distribution to the constant (using the continuity theorem). In fact, convergence in distribution to a constant is the same as convenience in probability. □

Proof of CLT (Theorem 1.1.2). Assume for simplicity that $\mu = 0$, this does not restrict the generality.

We do the same as for LLN, and write

$$\varphi_{(X_1+\ldots+X_N)/\sqrt{N\sigma^2}}(t) = \left[\varphi_{X_1}(t/\sqrt{N\sigma^2})\right]^N.$$

For random variable X_1 with finite two moments, we have for fixed t:

$$\varphi_{X_1}(t/\sqrt{N\sigma^2}) = 1 - \sigma^2/2(t/\sqrt{N\sigma^2})^2 + o(1/N^2) = 1 - t^2/(2N) + o(1/N^2).$$

Thus,

$$\varphi_{(X_1+\ldots+X_N)/\sqrt{N\sigma^2}}(t) = [\varphi_{X_1}(t/\sqrt{N\sigma^2})]^N = [1 - t^2/(2N) + o(1/N^2)]^N \to e^{-t^2/2}$$

This right-hand side is the characteristic function of the standard normal random variable (EX), and in this way we obtain the desired convergence. □

[2]Here and below (EX) stands for an exercise for the reader. For the MATH 7382 course I'll compose separate graded problem sets.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

1.3 De Moivre-Laplace theorem

Historically, the LLN and especially CLT were not that general. Indeed, observe that the proofs we gave are based on rather "modern" technical statements. But first, CLT appeared in the first half of XVIII century in the form of *de Moivre-Laplace theorem*. This theorem deals with the binomial distribution, a very special case. This special case corresponds to X_1 in Theorem 1.1.2 taking only values 0 and 1.

Definition 1.3.1 (Binomial distribution)**.** Let $0 < p < 1$, $q = 1 - p$. Consider the binomial (Bernoulli) scheme: one tosses the coin which has p as the probability of Heads, N times, and records the total number of Heads. Denote this random variable — total number of Heads — by S_N. This random variable is said to have the *binomial distribution*, and one can readily show (an easy combinatorial exercise) that

$$\mathbb{P}(S_N = k) = \binom{N}{k} p^K q^{N-K}$$

for $k = 0, \ldots, N$.

Here $\binom{N}{k}$ is the binomial coefficient, the number of k-element subsets in $\{1, \ldots, N\}$:

$$\binom{N}{k} = \frac{N!}{(N-k)!k!},$$

$N! := 1 \cdot 2 \cdot \ldots \cdot N$.

In fact,

$$S_N = X_1 + \ldots + X_N,$$

where X_i is the number of Heads in the ith trial,

$$X_i = \begin{cases} 0, & \text{with probability } 1-p; \\ 1, & \text{with probability } p. \end{cases}$$

Theorem 1.3.2 (de Moivre-Laplace theorem)**.** *Let k be in the neighborhood of Np. Then we can approximatie*

$$\binom{N}{k} p^k q^{N-k} \approx \frac{1}{\sqrt{2\pi Npq}} e^{-(k-Np)^2/(2Npq)},$$

as N is large, p is fixed and k grows in the neighborhood of Np.

Theorem 1.3.3 (Integral de Moivre-Laplace theorem)**.** *We have*

$$\mathbb{P}\left(\frac{S_N - Np}{\sqrt{Npq}} \le x\right) \approx \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz, \qquad N \to \infty.$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

One can readily obtain this integral theorem from the "local" Theorem 1.3.2.

Sketch of proof of Theorem 1.3.2. We compute everything "by hand"; using the Stirling formula for the factorial:[3]

$$n! \approx \sqrt{2\pi} \cdot n^{n+1/2} e^{-n}.$$

Let $k \sim aN$, that is, we zoom at some particular "local" part of the binomial distribution. We have

$$\binom{N}{k(N)} = \frac{N!}{k(N)!(N-k(N))!} = \frac{1}{\sqrt{2\pi N a(1-a)}} \frac{1}{(a^a(1-a)^{1-a})^N}$$

Then we also multiply this by $p^{k(N)}(1-p)^{N-k(N)}$; so

$$\binom{N}{k(N)} p^{k(N)}(1-p)^{N-k(N)} = \frac{1}{\sqrt{2\pi N a(1-a)}} \left(\frac{p^a(1-p)^{1-a}}{a^a(1-a)^{1-a}}\right)^N.$$

If $a \neq p$, this is < 1 (EX); and thus the probability exponentially tends to zero. So, we assume that up to the first order $k = Np$; and introduce

$$x = \frac{k(N) - Np}{\sqrt{Npq}}.$$

We assume that this $x \in \mathbb{R}$ is a new scaled coordinate. Then (EX; see also the Poisson distribution example below) it is possible to substitute $k = Np + x\sqrt{Npq}$, and conclude the proof. □

Using finer estimates for the factorial, it is possible to deduce more information about the behavior of the binomial probabilities; i.e., see the rate of convergence of the normalized sums to the standard normal law.

We see that the brute force computation works for this particular example, and allows to obtain the CLT (and, more precisely, the "local" CLT in the form of Theorem 1.3.2).

1.4 Other distributions

Let us discuss what may happen if one replaces the Bernoulli (zero-one) distribution of X_1 with some other known distribution (having a name...)

1.4.1 Normal

If the distribution is normal, there is no theorem: CLT is satisfied automatically; the normalized sums already have standard normal distribution.

[3] In fact, there is a whole asymptotic expansion of the factorial, but here we use only the first term.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

1.4.2 Poisson distribution

Let us consider this example more carefully. The computation is similar to the one in the binomial case, but is slightly shorter, so we may perform it in full detail.

If X_i is Poisson with parameter λ, then $S_N = X_1 + \ldots + X_N$ is also Poisson with parameter $N\lambda$ (EX; consider sum of two independent Poisson variables):

$$\mathbb{P}(S_N = k) = \frac{e^{-N\lambda}(N\lambda)^k}{k!}.$$

If we want CLT, let $k = k(N)$, consider

$$\frac{e^{-N\lambda}N^{k(N)}\lambda^{k(N)}}{k(N)!} = \frac{e^{-N\lambda+k}(N\lambda)^k}{\sqrt{2\pi}\cdot k^{k+1/2}}.$$

If $k \sim aN$ (again, we zoom around some global position), then

$$\frac{e^{-N\lambda}N^{k(N)}\lambda^{k(N)}}{k(N)!} = \frac{e^{-N\lambda+Na}(N\lambda)^{Na}}{\sqrt{2\pi}\cdot (Na)^{Na+1/2}} = \left(e^{(a-\lambda)}\frac{\lambda^a}{a^a}\right)^N \frac{1}{\sqrt{2\pi Na}}$$

If $a \neq \lambda$, the base of the exponent is < 1 (EX), and all goes to zero exponentially; so let us set $a = \lambda$. Let us consider the scaling

$$k = N\lambda + x\lambda^{1/2}N^{1/2}.$$

We have then a more precise expansion

$$\begin{aligned}\frac{e^{-N\lambda}N^{k(N)}\lambda^{k(N)}}{k(N)!} &= \frac{e^{x\lambda^{1/2}N^{1/2}}}{\sqrt{2\pi(N\lambda + x\lambda^{1/2}N^{1/2})}\cdot(1+x\lambda^{-1/2}N^{-1/2})^{N\lambda+x\lambda^{1/2}N^{1/2}}}\\ &= \frac{e^{-x^2}e^{x\lambda^{1/2}N^{1/2}}}{\sqrt{2\pi N\lambda}\cdot(1+x\lambda^{-1/2}N^{-1/2})^{N\lambda}}.\end{aligned}$$

Let us deal with $(1 + x\lambda^{-1/2}N^{-1/2})^{N\lambda}$; we have

$$(1 + x\lambda^{-1/2}N^{-1/2})^{N\lambda} = e^{N\lambda\ln(1+x\lambda^{-1/2}N^{-1/2})} = e^{N\lambda(x\lambda^{-1/2}N^{-1/2}-x^2/(2N\lambda)))} = e^{N^{1/2}\lambda^{1/2}x-x^2/2}.$$

In the end we have that

$$\mathbb{P}(S_N = N\lambda + x\sqrt{N\lambda}) \sim \frac{1}{\sqrt{2\pi N\lambda}}e^{-x^2/2}.$$

This agrees with the CLT.

Remark 1.4.1. The factor $\sqrt{N\lambda}$ (as well as $\sqrt{Npq}$ in the de Moivre-Laplace theorem in the previous example) reflects the discrete to continuous scaling limit; i.e., we need to scale the Poisson distribution to get a meaningful limit. This is done according to the CLT.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

1.4.3 Geometric, exponential, ...

If X_1 has a geometric distribution (i.e., with weights $p(1-p)^k$, $k = 0, 1, 2, \ldots$), then S_N has negative binomial distribution (EX; this is a known fact but requires a computation), one can deal with this also explicitly, using the Stirling formula for the Gamma function.

If the distribution of X_1 is exponential (with density $\lambda e^{-\lambda x}$, $x > 0$), then S_N has a Gamma distribution; this example also belongs to our "explicit zoo".

There are some other examples of distributions for which the CLT can be obtained in an explicit brute force way. The computation each time has to be performed on its own. (We pretend that we don't know characteristic functions and "modern technique" of XIX century.)

1.5 Conclusions

(1) If there is no "modern" (late XIX century) technique, certain models are still acessible using direct computations; this involves combinatorics and sometimes algebra

(2) These "explicit" ("integrable") models also allow more precise results at little cost; just continue expansions. For example, for Poisson, $k = N\lambda + x\sqrt{N\lambda}$:

$$\mathbb{P}(S_N = k) \approx \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}\sqrt{N\lambda}} + \frac{e^{-\frac{x^2}{2}} x\left(x^2-3\right)}{6\sqrt{2\pi}\lambda N} + O\left(\frac{1}{N}\right)^{3/2},$$

etc.

1.6 Where now?

The LLN and CLT deal with the old and well-studied model of independent identically distributed random variables, this may be thought of as "zero-dimensional growth model". Nowadays, in probability much more complicated models are studied. Little general is known about them; and often only explicit computation techniques are available. Thus, in the realm of more complicated models, we are at the state of de Moivre-Laplace theorem. However, as the models are more complicated, we use not only combinatorial computations, but also related algebraic structures are involved.

Origins of these more complicated models some of which we are going to discuss:

(a) Random matrices (historically)

(b) Two-dimensional growth of various sorts (PNG, TASEP, etc.)

(c) Representation theory and enumerative combinatorics: provide cerain algebraic frameworks/models/insights/motivations which we will use, and on which many meaningful models are based.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

1.7 Some nice pictures on slides

1. PNG droplet
2. TASEP
3. push-block dynamics on interlacing particle arrays
4. From interlacing arrays to lozenge tilings
5. Lozenge tilings of infinite and finite regions. Features.

1.8 Interlacing arrays and representations of unitary groups

As a last topic for the first lecture, let us briefly explain what representation-theoretic background is behind the interlacing arrays.

1.8.1 Unitary groups and their representations. Signatures

Let $U(N)$ denote the Nth unitary group, a group of $N \times N$ unitary matrices (complex matrices with $UU^* = U^*U = 1$, where * means conjugation inverse). This is a compact group; and thus to understand representations of $U(N)$ one can argue similarly to finite groups (more will follow in subsequent lectures). Let us briefly present the necessary definitions and facts.

All irreducible representations of $U(N)$ are parametrized by *signatures* (or, in other terminology, *highest weights*) of length N which are nonincreasing N-tuples of integers:

$$\lambda = (\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N), \qquad \lambda_i \in \mathbb{Z}.$$

Let us denote by $\mathbb{GT}_N$ the set of all signatures of length N.[4] Let also denote by π_N^λ the corresponding representation of $U(N)$.

Remark 1.8.1. For $N = 1$, the unitary group is simply the unit circle in $\mathbb{C}$, and its representations are parametrized by integers $\mathbb{Z}$, which is the same as $\mathbb{GT}_1$.

1.8.2 Branching rule

We need not a single unitary group, but the chain of unitary groups of growing order:

$$U(1) \subset U(2) \subset U(3) \subset \ldots,$$

[4] The letters G and T stand for Gelfand and Tsetlin who investigated these problems in the '50s.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

where the inclusions are defined as

$$U(N-1) \ni U \mapsto \begin{bmatrix} U & 0 \\ 0 & 1 \end{bmatrix} \in U(N)$$

(that is, we add zero column and row and a singe 1 on the diagonal).

A natural question in connection with this chain of unitary groups is:

Question 1.8.2. What happens with the representation π_λ^N of the unitary group if we restrict it to $U(N-1) \subset U(N)$?[5]

This restricted representation in completely reducible (we discuss the necessary representation-theoretic notions in subsequent lectures), and the explicit decomposition looks as

$$\pi_N^\lambda|_{U(N-1)} = \bigoplus_{\mu \in \mathbb{GT}_{N-1}\colon\, \mu \prec \lambda} \pi_{N-1}^\mu.$$

A nice feature is that this decomposition is *simple*, i.e., every direct summand appears only once. The representations of $U(N-1)$ which enter the decomposition are indexed by signatures $\mu \in \mathbb{GT}_{N-1}$ which *interlace* with λ:

$\mu \prec \lambda$ by definition means that

$$\lambda_1 \ge \mu_1 \ge \lambda_2 \ge \mu_2 \ge \ldots \ge \lambda_{N-1} \ge \mu_{N-1} \ge \lambda_N.$$

This is the so-called *branching rule* for irreducible representations of unitary groups.

1.8.3 Branching rule and Gelfand–Tsetlin basis in π_N^λ

The branching rule can be continued. Fix $\lambda \in \mathbb{GT}_N$, and consider the decomposition of π_N^λ into one-dimensional representations; in this way we find a basis in π_N^λ every vector in which is parametrized by an interlacing sequence of signatures

$$\lambda^{(1)} \prec \lambda^{(2)} \prec \ldots \prec \lambda^{(N)} = \lambda, \qquad (1.8.1)$$

where $\lambda^{(i)} \in \mathbb{GT}_i$. Another name for such an interlacing sequence of signatures is *Gelfand–Tsetlin scheme*. Such basis in the representation is called a *Gelfand–Tsetlin basis*.

Remark 1.8.3. A consequence of this construction is the following combinatorial description of the dimension $\mathrm{Dim}_N \lambda := \dim(\pi_N^\lambda)$ of the irreducible representation π_N^λ. Namely, $\mathrm{Dim}_N \lambda$ is the number of all Gelfand–Tsetlin schemes with fixed top (Nth) row $\lambda \in \mathbb{GT}_N$.

Using the Weyl dimension formula, we may in fact conclude that

$$\mathrm{Dim}_N \lambda = \prod_{1 \le i < j \le N} \frac{\lambda_i - i - \lambda_j + j}{j - i}.$$

[5]Of course, a representation of a group is also a representation of its subgroup.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

1.8.4 Gelfand–Tsetlin schemes. Interlacing arrays

A Gelfand–Tsetlin scheme (1.8.1) may be also viewed as a triangular array of integers as on Fig. 1.2.

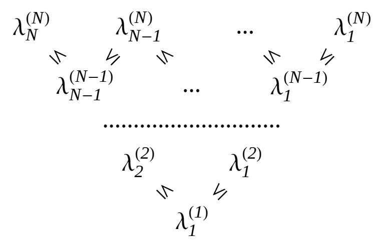


Figure 1.2: A Gelfand–Tsetlin scheme of depth N.

For technical convenience (i.e., to be able to interpret signatures as point configurations on $\mathbb{Z}$) sometimes shifted coordinates are used:

$$x_j^m := \lambda_j^{(m)} - j.$$

The advantage is that the λ-coordinates may have coinciding points, while the x-coordinates x_j^m for any chosen m are always distinct.

These shifted x-coordinates satisfy the following interlacing constraints:

$$x_{j+1}^m < x_j^{m-1} \le x_j^m$$

(for all j's and m's for which these inequalities can be written out).

The interlacing array $\{x_j^m\}$ can be then interpreted as a point configuration which is evolved under the push-block dynamics. On Fig. 1.3 one particle with coordinate x_1^1 is put on the first level, two particles $x_2^2 < x_1^2$ are put on the second (upper) level, and so on.

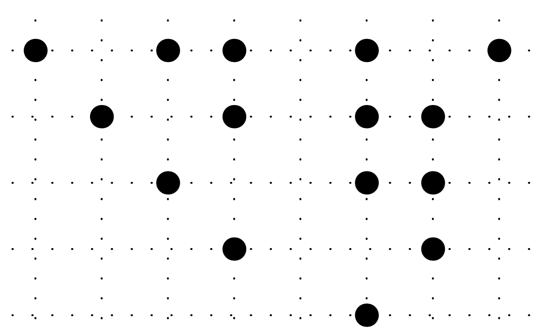


Figure 1.3: Interlacing particle array.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

In this way, we see at least some (may be seemingly remote) connection of probabilistic models we've seen on slides and representations of unitary groups. In fact, these connections allow to investigate fine properties of probabilistic models; and in many models these methods are so far the only ones available.

1.9 References

Details on the method of characteristic functions may be found in, e.g., [Shi95, Ch. III]. De Moivre-Laplace theorem is discussed in Ch. I of that book. The PNG model is investigated in [PS02], a slightly more general model is discussed also in [BO06]. About TASEP and push-block dynamics see [BF08], and references therein. Representation theory and branching of irreducible representations of unitary groups is discussed in, e.g., [Wey97].

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Chapter 2

De Finetti's Theorem

In this and the next lecture we will discuss a "toy model" of ideas which will come in subsequent lectures. This is de Finetti's theorem which may be viewed as one of the limit theorems in classical probability along with the LLN and CLT (Chapter 1). This model possesses some algebraic features which will be interesting for us.

2.1 Formulation

Definition 2.1.1. Let $X_1, X_2, \dots$ be a sequence of random variables. We do not assume that they are independent. This sequence is called *exchangeable* if the law of the sequence does not change under permutations of the random variables. In detail, for any N and $\sigma \in \mathfrak{S}(N)$ we have

$$\operatorname{Law}(X_1, \dots, X_N) = \operatorname{Law}(X_{\sigma(1)}, \dots, X_{\sigma(N)}).$$

Note that this implies that the random variables X_i are identically distributed. Of course, if the variables X_i are iid (independent identically distributed), then the sequence $X_1, X_2, \dots$ is exchangeable.

Definition 2.1.2. Let $\mathfrak{S}(\infty)$ denote the *infinite symmetric group* which is the group consisting of finitary permutations of the natural numbers $\mathbb{N}$. That is, every permutation $\sigma \in \mathfrak{S}(\infty)$ fixes almost all elements of $\mathbb{N}$.

One can say that the exchangeable property means that the law of the sequence $X_1, X_2, \dots$ is $\mathfrak{S}(\infty)$-*invariant.*

We will consider exchangeable *binary* sequences, that is, every X_i can take values 0 and 1.[1] They are naturally identified with Borel measures on the infinite

[1] There is a generalization which allows to consider random variables of any nature, but for binary sequences the discussion seems clearer.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

product space (with product topology)

$$\{0,1\}^\infty := \{(x_1, x_2, \ldots)\colon x_i = 0 \text{ or } 1\}. \tag{2.1.1}$$

A natural question is

Question 2.1.3. How do *all* the exchangeable binary sequences look like?

Before formulating the answer, let us note that for binary sequences the whole distribution of $X_1, X_2, \ldots$ is completely determined (EX) by the following probabilities:

$$\begin{aligned} P_{N,k} &:= \mathbb{P}(X_1 = 1, \ldots, X_k = 1, X_{k+1} = 0, \ldots, X_N = 0) \\ &= \frac{\mathbb{P}(X_1 + \ldots + X_N = k)}{\binom{N}{k}}. \end{aligned} \tag{2.1.2}$$

The rule of addition of probabilities yields the backward recursion

$$P_{N,k} = P_{N+1,k} + P_{N+1,k+1}, \quad 0 \le k \le N,\ N = 0, 1, \ldots, \tag{2.1.3}$$

for the numbers (2.1.2).

The classification theorem of exchangeable binary sequences is:

Theorem 2.1.4 (de Finetti)**.** *Exchangeable binary sequences $X_1, X_2, \ldots$ are in one-to-one correspondence with Borel probability measures μ on the segment $[0,1]$. The correspondence looks as*

1. *For a measure μ, the law of the sequence is determined via the numbers $P_{N,k}$ (2.1.2) as follows:*

$$P_{N,k} = \int_0^1 x^k (1-x)^{N-k} \mu(dx)$$

2. *If the sequence $X_1, X_2, \ldots$ is exchangeable, there is a limit in distribution*

$$\frac{X_1 + \ldots + X_N}{N} \to \mu.$$

2.2 Remarks on de Finetti's theorem

2.2.1 Bernoulli sequences

The iid binary sequences are often called the *Bernoulli sequences.* They are determined by a single number $p = \mathbb{P}(X_1 = 1)$, $0 \le p \le 1$ (we include constant sequences into Bernoulli sequences as well). Let us denote by ν_p the measure on $\{0,1\}^\infty$ (2.1.1) corresponding to the Bernoulli sequence with parameter p. Then de Finetti's theorem is equivalent to the following:

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Theorem 2.2.1. *For every exchangeable law ν on $\{0,1\}^\infty$ there exists a unique Borel probability measure μ on $[0,1]$ such that*

$$\nu = \int_0^1 \nu_p \cdot \mu(dp).$$

To *sample* an exchangeable sequence corresponding to a measure μ there is thus a two-step procedure: first choose at random the value of parameter p from the probability distribution μ on $[0,1]$, then sample an infinite Bernoulli sequence with probability p for 1's.

2.2.2 Ergodicity

Measures ν_p that appear in Theorem 2.2.1 can be characterized as follows. Consider a general setting: $\mathfrak{X}$ is a measurable space (that is, a set with a distinguished sigma–algebra of subsets) and G is a group acting on $\mathfrak{X}$ by measurable transformations (in our context, $\mathfrak{X} = \{0,1\}^\infty$ and $G = \mathfrak{S}(\infty)$).

Consider the *convex set* of all G–invariant probability measures on $\mathfrak{X}$. For a measure ν from this set the following conditions are equivalent:

(1) ν is extreme.

(2) Any invariant modulo 0 measurable set in $\mathfrak{X}$ has ν–measure 0 or 1. [2]

(3) The subspace of G–invariant vectors in $L^2(\mathfrak{X}, \nu)$ is one–dimensional, that is, it consists of the constant functions.

Condition (2) is usually taken as the definition of *ergodic* measures. If G is discrete then the words "modulo 0" in condition (2) can be omitted.

Theorem 2.1.4 (or 2.2.1) asserts that extreme (= ergodic, in this context) exchangeable binary sequences are precisely the Bernoulli sequences.

2.2.3 Hausdorff moment problem

One approach to this classical result, as presented in Feller [Fel71, Ch. VII, §4], is based on the following exciting connection with the Hausdorff moment problem.[3]

Recursion (2.1.3) readily implies that the array can be derived by iterated differencing of the sequence $(P_{n,0})_{n=0,1,\dots}$. Specifically, setting

$$u_l^{(k)} = P_{l+k,k}, \qquad l = 0,1,\dots, \quad k = 0,1,\dots, \tag{2.2.1}$$

and denoting by δ the difference operator acting on sequences $u = (u_l)_{l=0,1,\dots}$ as

$$(\delta u)_l = u_l - u_{l+1},$$

the recursion (2.1.3) can be written as

$$u^{(k)} = \delta u^{(k-1)}, \qquad k = 1,2,\dots. \tag{2.2.2}$$

[2] A set $A \subset \mathfrak{X}$ is called *invariant modulo 0* if for any $g \in G$, the symmetric difference between A and $g(A)$ has ν–measure 0.

[3] However, we will use another way to establish this fact.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Since $P_{n,k} \geq 0$, the sequence $u^{(0)}$ must be completely monotone, that is, componentwise

$$\underbrace{\delta \circ \cdots \circ \delta}_{k} u^{(0)} \geq 0, \qquad k = 0, 1, \ldots,$$

but then Hausdorff's theorem implies that there exists a representation

$$P_{n,k} = u^{(k)}_{n-k} = \int_{[0,1]} p^k (1-p)^{n-k} \mu(\mathrm{d}p) \tag{2.2.3}$$

with uniquely determined probability measure μ.

De Finetti's theorem follows since $P_{n,k} = p^k(1-p)^{n-k}$ for the Bernoulli process with parameter p.

2.3 Pascal triangle

In this section we describe the Pascal triangle (or Pascal graph) $\mathbb{PT}$ which provides a convenient framework for our "toy example" of de Finetti's theorem.

2.3.1 Definition

Definition 2.3.1. The Pascal triangle $\mathbb{PT}$ is the lattice $\mathbb{Z}^2_{\geq 0}$ equipped with a structure of a branching graph as follows. Set

$$\mathbb{PT}_N := \{(a, b) \in \mathbb{Z}^2_{\geq 0} \colon a + b = N\},$$

and say that $\mathbb{PT}_N$ is the Nth floor of the Pascal triangle. Let the edges in this branching graph connect (a, b) with $(a + 1, b)$ and $(a, b + 1)$ for all (a, b). Thus, $\mathbb{PT}_N$ is precisely the vertices that lie at distance N from $(0, 0)$, the initial vertex. See Fig. 2.1.

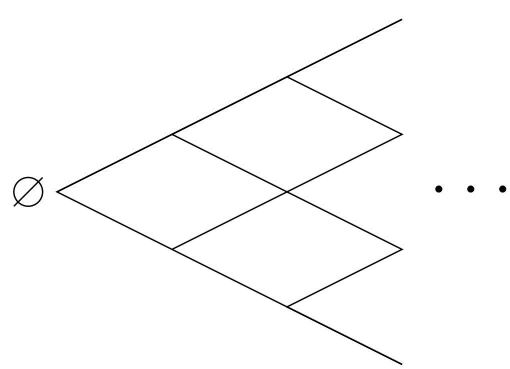


Figure 2.1: Pascal triangle (first four levels)

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Clearly, the total number of oriented paths from $(0,0)$ to (a,b) is equal to

$$\dim(a,b) = \binom{a+b}{a}.$$

2.3.2 Harmonic functions

The definition of $\mathbb{PT}$ is related to the backward recursion (2.1.3) in the following way. Place the number $\varphi(k, N-k) := P_{N,k}$ at the vertex $(k, N-k)$; then (2.1.3) means exactly that the number $\varphi(a,b)$ at each vertex is the sum of all of the "above" numbers:

$$\varphi(a,b) = \varphi(a+1,b) + \varphi(a,b+1). \tag{2.3.1}$$

De Finetti's theorem now becomes equivalent to the following question:

Question 2.3.2. Classify all functions φ on $\mathbb{Z}^2_{\geq 0}$ satisfying (2.3.1) which are nonnegative and normalized at 0: $\varphi(0,0) = 1$. Such functions are called *harmonic* (for some historical reasons; but in fact this name is not very accurate).

Extreme harmonic functions are precisely $\varphi_p(a,b) = p^a(1-p)^b$.

Remark 2.3.3. In fact, to show the equivalence of Questions 2.1.3 and 2.3.2, one needs to consider cylindrical sets of infinite paths in the Pascal triangle: the whole set of infinite paths is clearly $\{0,1\}^\infty$, and to define measure on it one must define it on cylindrical sets. We consider sets of paths which have a fixed beginning up to level, say, N, and then continue arbitrarily.

Exchangeable property tells that the measure of such a set depends only on the vertex on level N (say, it is $(k, N-k)$). This measure must be set equal to $\varphi(k, N-k)$ to obtain the desired equivalence. Then the convex sets of exchangeable laws $\{\nu\}$ and normalized nonnegative harmonic functions $\{\varphi\}$ are in a natural bijection.

2.3.3 Coherent systems of measures. Boundary

We will need one more definition. Consider an exchangeable measure ν on the set of paths $\{0,1\}^\infty$ in the Pascal triangle. Denote by $M_{a+b}(a,b)$ the measure of all paths which pass through the vertex (a,b). Clearly, M_N is a probability measure on $\mathbb{PT}_N$ for all N. Moreover,

$$M_N(a,b) = \binom{N}{a} \cdot \varphi(a,b), \qquad a+b = N.$$

These measures satisfy the following recurrence:

$$M_N(a,b) = \frac{a+1}{N+1} M_{N+1}(a+1,b) + \frac{b+1}{N+1} M_{N+1}(a,b+1). \tag{2.3.2}$$

Definition 2.3.4. The system of probability measures $\{M_N\}$ (each M_N lives on $\mathbb{PT}_N$) satisfying the above recurrence is called a *coherent system on* $\mathbb{PT}$.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Remark 2.3.5. Coherent system of measures $\{M_N\}_{N=0}^{\infty}$ is completely determined by any its "tail" $\{M_L, M_{L+1}, M_{L+2}, \ldots\}$. Indeed, using (2.3.2), one can readily project the coherent system down onto lower levels. This of course agrees with the backwards nature of the recurrence (2.3.2) (or (2.1.3)).

Coherent systems also form a convex set. We are interested in describing extreme coherent systems and all coherent systems.[4] This problem is again equivalent to the de Finetti's theorem.

Definition 2.3.6. The set of all extreme coherent systems on $\mathbb{PT}$ is called the *boundary* of the Pascal triangle.

The statement of the de Finetti's theorem is equivalent to saying that the boundary of the Pascal triangle is the unit interval $[0, 1]$. This can be seen geometrically: the floors of $\mathbb{PT}_N$ approximate $[0, 1]$ uniformly via the embeddings

$$\mathbb{PT}_N \ni (a, b) \mapsto \frac{a}{N} \in [0, 1].$$

Remark 2.3.7. The equation (2.3.2) for the coherent systems may be interpreted in the following form. Consider the original exchangeable random binary sequence $(X_1, X_2, X_3, \ldots)$, and consider the first N variables $(X_1, \ldots, X_N)$. Then the law of these completely determines the law of $(X_1, \ldots, X_{N-1})$. This is the way to understand (2.3.2).

One may also consider the (time-inhomogeneous; even the state spaces change with time) Markov process of cutting the last variable. Then the boundary of the Pascal triangle is the same as the "entrance boundary" for this Markov process, i.e., the set of all possible "starting points".

2.3.4 Approximation method

There is a powerful general approximation idea of Vershik–Kerov (employed in, e.g., [Ver74], [VK81a], [VK81b], [VK82], [VK87]) which explains how to approximate extreme coherent systems on the whole $\mathbb{PT}$ by "finite" analogues.

Fix N, and consider the part $\mathbb{PT}(N) := \mathbb{PT}_0 \cup \ldots \cup \mathbb{PT}_N$ of the Pascal triangle up to level N. This is again a branching graph, but its coherent systems are obviously described. Namely, all coherent systems are determined by the measures M_N on the level N (cf. Remark 2.3.5). Extremes of them are coming from the delta-measures on $\mathbb{PT}_N$.

The whole coherent system is given by:

$$M_K^{(a,b)}(c, d) = \frac{\#\text{ of paths from } 0 \text{ to } (a, b) \text{ passing through } (c, d)}{\text{total } \#\text{ of paths from } 0 \text{ to } (a, b)},$$

where $(a, b) \in \mathbb{PT}_N$ is the label of the coherent system, $K \le N$ and $(c, d) \in \mathbb{PT}_K$. Using the dim notation for the number of paths, we have

$$M_K^{(a,b)}(c, d) = \dim(c, d)\frac{\dim[(c, d); (a, b)]}{\dim(a, b)}. \tag{2.3.3}$$

[4] Any coherent system is a continuous convex combination of the extremes; this follows from the Chouquet's theorem which will be discussed in subsequent lectures.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

here $\dim(a,b) = \binom{a+b}{a}$ denotes the total number of paths from 0 to (a,b) (see above), and $\dim[(a,b);(c,d)]$ denotes the number of paths from (c,d) to (a,b) (if there are any; note that this number may be zero).

The essence of the approximation method is:

> *Any coherent system on the infinite graph $\mathbb{PT}$ is a limit of coherent systems $M^{(a(N),b(N))}$ on finite graphs $\mathbb{PT}_N$ as $N \to \infty$, where $(a(N),b(N)) \in \mathbb{PT}_N$ is a suitable sequence of vertices.*
>
> *In detail,[5] for any fixed K and $(c,d) \in \mathbb{PT}_K$, it must be*
>
> $$\lim_{N\to\infty} M_K^{(a(N),b(N))}(c,d) = M_K(c,d),$$
>
> *where M_K on the left belongs to the limiting coherent system.*

Thus, to describe the boundary of $\mathbb{PT}_N$ is the as to

> *Describe all sequences of numbers $(a(N),b(N)) \in \mathbb{PT}_N$ for which the quantities*
>
> $$\frac{\dim[(c,d);(a(N),b(N))]}{\dim(a(N),b(N))}$$
>
> *have a limit as $N \to \infty$ for any fixed $(c,d) \in \mathbb{PT}_K$ (for some fixed K).*

Here we simply removed the factor $\dim(c,d)$ from (2.3.3) because it does not depend on N. Note that $\frac{\dim[(c,d);(a(N),b(N))]}{\dim(a(N),b(N))}$ is bounded (in fact, it is ≤ 1, and it is of course nonnegative), so we need only to have a limit, and it will automatically be finite.

Remark 2.3.8. In other words (to further clarify the concept of the boundary), we have a way to project measures from $\mathbb{PT}_N$ to $\mathbb{PT}_K$ for $K < N$; and we would like to find a set "$\mathbb{PT}_\infty$" — the boundary of $\mathbb{PT}$ — from which we can project to every finite level $\mathbb{PT}_N$. Its points would correspond to extreme coherent systems on $\mathbb{PT}$.

2.3.5 Proof of de Finetti's theorem (= description of the boundary of the Pascal triangle)

The proof is based on explicit formulas for the dimension $\dim(a,b)$ and for the relative dimension $\dim[(c,d);(a,b)]$. Namely, it is clear that

$$\dim(a,b) = \binom{a+b}{a}, \qquad \dim[(c,d);(a,b)] = \binom{a+b-c-d}{a-c}$$

(with the understanding that the last binomial coefficient vanishes in appropriate cases).

[5]This statement explains what it means for one coherent system to converge to another.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

We have thus

$$\begin{aligned}\frac{\dim[(c,d);(a(N),b(N))]}{\dim(a(N),b(N))} &= \binom{a(N)+b(N)-c-d}{a-c}\Big/\binom{N}{a(N)}\\ &= \frac{(a(N))^{\downarrow c}(b(N))^{\downarrow d}}{N^{\downarrow(c+d)}}\\ &= \left(\frac{a(N)}{N}\right)^c\left(\frac{b(N)}{N}\right)^d + O(1/N),\end{aligned}$$

where we denote $x^{\downarrow k} := x(x-1)\dots(x-k+1)$.

We now need to find all sequences $a(N)$ (note that $b(N) = N - a(N)$) for which

$$\left(\frac{a(N)}{N}\right)^c\left(1-\frac{a(N)}{N}\right)^d + O(1/N)$$

has a finite limit for fixed (c,d). Clearly, $a(N)/N$ must converge to some number, say, p (take $(c,d) = (1,0)$). Moreover, $0 \le p \le 1$. We see that this convergence is sufficient.

We therefore have completely described the extreme coherent systems; they are given by

$$M_K^{ex;\,p}(c,d) = \binom{K}{c}p^c(1-p)^d.$$

The corresponding extreme harmonic function is

$$\varphi_p(c,d) = p^c(1-p)^d.$$

Any coherent system can be represented through extremes as

$$M_K(c,d) = \binom{K}{c}\int_0^1 p^c(1-p)^d\mu(dp), \tag{2.3.4}$$

where μ is the corresponding measure on $[0,1]$ (the so-called *boundary measure* of the coherent system $\{M_K\}$).

Remark 2.3.9. One can see this property (2.3.4) independently of the Choquet's theorem (which asserts that every coherent system can be represented as a convex combination of the extremes).

The harmonicity condition and the above computation imply

$$\begin{aligned}\psi(k_0,l_0) &= \sum_{k+l=n}\dim((k_0,l_0),(k,l))\cdot\psi(k,l)\\ &= \sum_{k+l=n}\frac{\dim((k_0,l_0),(k,l))}{\dim(k,l)}\,M_n(k,l)\end{aligned}$$

$$= \sum_{k+l=n} \left(\frac{k}{k+l}\right)^{k_0} \left(\frac{l}{k+l}\right)^{l_0} M_n(k,l) + O(n^{-1}).$$

For an arbitrary point $p \in [0,1]$, let $\langle p \rangle$ denote the Dirac delta-measure concentrated at p. Consider the sequence of probability measures

$$M^{(n)} = \sum_{k+l=n} M(k,l) \left\langle \frac{k}{k+l} \right\rangle$$

on [0,1]. It has a convergent subsequence because $[0,1]$ is compact and so the space of Borel probability measures on $[0,1]$ is also compact. Denoting the limit measure by μ, we obtain

$$\psi(k_0, l_0) = \int_0^1 p^{k_0}(1-p)^{l_0} \mu(dp)$$

for any $(k_0, l_0) \in \mathbb{Z}^2_{\geq 0}$. Uniqueness of μ follows from the density of the span of $\{p^k(1-p)^l \mid k, l \geq 0\}$ in $C[0,1]$.

2.4 Algebraic structure under de Finetti's theorem

Let us briefly discuss the underlying algebraic structure behind the Pascal graph and its boundary. The backward recurrence (2.1.3) for the harmonic functions

$$\varphi(a,b) = \varphi(a+1,b) + \varphi(a,b+1)$$

is mimicked by the property

$$(x+y)x^a y^b = x^{a+1}y^b + x^a y^{b+1} \tag{2.4.1}$$

(where x and y are formal variables).

Consider the polynomial algebra $\mathbb{C}[x,y]$. The basis $\{x^a y^b\}$ is indexed by vertices of the graph $\mathbb{PT}$. The identity (2.4.1) is said to *encode* the edges in the graph $\mathbb{PT}$.

The next important proposition is due to Vershik and Kerov (80's):

Proposition 2.4.1. *Normalized nonnegative extreme harmonic functions φ on $\mathbb{PT}$ are in a natural bijection with algebra homomorphisms $F\colon \mathbb{C}[x,y] \to \mathbb{R}$ such that*

- $F(x+y) = 1$
- $F(x^a y^b) \geq 0$ *for all vertices of* $\mathbb{PT}$

The bijection is given by $F(x^a y^b) = \varphi(a,b)$.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Proof. Step 0. It is clear that all harmonic functions are in bijection with linear functionals $F\colon \mathbb{C}[x,y] \to \mathbb{R}$ such that

- $F(1) = 1$
- F vanishes on the ideal $(x+y-1)\mathbb{C}[x,y]$
- $F(x^a y^b) \geq 0$ for all vertices of $\mathbb{PT}$

We will show that extreme harmonic functions correspond to multiplicative functionals. Denote $A = \mathbb{C}[x,y]$ for shorter notation, and let X be this set of linear functionals.

Step 1. Let A_+ be the cone in A formed by linear combinations of the basis elements $x^a y^b$ with nonnegative coefficients. Observe that this cone is closed under multiplication.

Step 2. For any $(a,b) \in \mathbb{PT}$ one has (by the binomial theorem)

$$(x+y)^n - \dim(a,b) \cdot x^a y^b \in A_+, \qquad n := a+b.$$

Step 3. Let $F \in X$ and $f \in A_+$ be such that $F(f) > 0$. Assign to f another functional defined by

$$F_f(g) = \frac{F(fg)}{F(f)}, \qquad g \in A$$

(the definition makes sense because $F(f) \neq 0$). It can be checked that F_f also belongs to X.

Step 4. Assume $F \in X$ is extreme and show that it is multiplicative. It suffices to prove that for any $(a,b) \neq (0,0)$,

$$F(x^a y^b g) = F(x^a y^b) F(g), \qquad g \in A. \tag{2.4.2}$$

There are two possible cases: either $F(x^a y^b) = 0$ or $F(x^a y^b) > 0$.

If the first case (2.4.2) is equivalent to saying that $F(x^a y^b \cdot x^c y^d) = 0$ for every $(c,d) \in \mathbb{PT}$. But we have, setting $n = c+d$,

$$0 \leq F(x^a y^b x^c y^d) \leq F(x^a y^b (x+y)^n) = F(x^a y^b) = 0,$$

where the first inequality follows from step 1, the second one follows from step 2, and the equality holds by virtue of properties of F. Therefore, $F(x^a y^b x^c y^d) = 0$, as desired.

In the second case we may form the functional $F_{x^a y^b}$, and then (2.4.2) is equivalent to saying that it coincides with F. Let $m = a+b$ and set

$$f_1 = \frac{1}{2} \dim(a,b)\, x^a y^b, \qquad f_2 = (x+y)^m - f_1.$$

Observe that F is strictly positive both on f_1 and on f_2, so that both F_{f_1} and F_{f_2} exist. On the other hand, for any $g \in A$ one has

$$F(g) = F((x+y)^m g) = F(f_1 g) + F(f_2 g), \qquad F(f_1) + F(f_2) = 1,$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

which entails that F is a convex combination of F_{f_1} and F_{f_2} with strictly positive coefficients:

$$F = F(f_1)F_{f_1} + F(f_2)F_{f_2}.$$

Since F is extreme, we conclude that $F_{f_1} = F$, as desired.

Step 5. Conversely, assume that $F \in X$ is multiplicative and let us show that it is extreme.

Observe that X satisfies the assumptions of Choquet's theorem. Indeed, as we may regard X as a subset of the vector space dual to A and equipped with the topology of simple convergence. This simply means that in X every element can be represented as a convex combination of extremes.

Let P be the probability measure on $E(X)$ (the set of extremes in X) representing F in accordance with Choquet's theorem. This implies that

$$F(f) = \int_{G \in E(X)} G(f)P(dG), \qquad f \in A. \tag{2.4.3}$$

We are going to prove that P is actually the delta-measure at a point of $E(X)$, which just means that F is extreme.

Write ξ_f for the function $G \to G(f)$ viewed as a random variable defined on the probability space $(E(X), P)$. Equality (2.4.3) says ξ_f has mean $F(f)$.

By virtue of step 4, every $G \in E(X)$ is multiplicative. On the other hand, F is multiplicative by the assumption. Using this we get from (2.4.3)

$$\left(\int_{G \in E(X)} G(f)P(dG)\right)^2 = (F(f))^2 = F(f^2) = \int_{G \in E(X)} G(f^2)P(dG)$$
$$= \int_{G \in E(X)} (G(f))^2 P(dG).$$

Comparing the leftmost and the rightmost expressions we conclude that ξ_f has zero variance. Hence $G(f) = F(f)$ for all $G \in E(X)$ outside a P-null subset depending on f.

It follows that $G(x^a y^b) = F(x^a y^b)$ for all $(a, b) \in \mathbb{PT}$ and all $G \in E(X)$ outside a P-null subset, which is only possible when P is a delta-measure.

This concludes the proof. □

Remark 2.4.2. This statement and the overall formalism which associates a branching graph with an algebra helps in other, more general cases. For instance, it is a tool to classify characters of the infinite symmetric group. We will discuss the corresponding graph later.

Remark 2.4.3. Note that the "kernels" $p^k(1-p)^l$ entering the integral representation for the harmonic functions form a basis in the factoralgebra $\mathbb{C}[x, y]/(x + y - 1)$ (if we understand that $x = p$).

Remark 2.4.4 (Choquet's theorem)**.** We have used the following Choquet's theorem

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Theorem 2.4.5 (Choquet's Theorem, see Phelps [Phe66])**.** *Assume that X is a metrizable compact convex set in a locally convex topological space, and let x_0 be an element of X. Then there exists a (Borel) probability measure P on X supported by the set of extreme points $E(X)$ of X, which represents x_0. In other words, for any continuous linear functional f*

$$f(x_0) = \int_{E(X)} f(x) P(dx).$$

($E(X)$ is a G_δ-set, hence it inherits the Borel structure from X.)

It assures the existence of the measure on extremes so that any element of the set is represented as the convex combination of the extremes.

2.5 References

De Finetti's theorem and generalizations, as well as related results are discussed in [Ald85]. Other examples of branching graphs of Pascal-type (when edges are equipped with some multiplicities) are considered in [Ker03, Ch. I, Section 2–4], and also in [GP05], [GO06], [GO09]. On general formalism regarding algebraic structures behind graphs considered in this lecture see [BO00].

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Chapter 3

Irreducible characters of finite and compact groups

In this chapter we discuss more traditional material: basics of representation theory, and especially theory of characters.

3.1 Basic definitions

Definition 3.1.1. Let G be a group which is assumed to be finite or compact finite-dimensional Lie group (such as the unitary group $U(N)$).

A representation of G is a homomorphism $G \to GL(V)$, where V is a vector space. We will work with vector spaces over $\mathbb{C}$.

Definition 3.1.2. If $G \to GL(V_1)$, $G \to GL(V_2)$ are two representations, one can define the direct sum $V_1 \oplus V_2$ in an obvious way.

A representation is called *irreducible* if it has no nontrivial invariant subspaces.

A representation is *indecomposable* if it cannot be written as $V_1 \oplus V_2$, where V_1, V_2 are nontrivial subspaces.

Every irreducible representation is indecomposable.

Every complex finite-dimensional representation of a group G as above (we call them "finite or compact groups") can be expressed as a direct sum of irreducibles. This can be seen from the following property:

Proposition 3.1.3. *Let $T\colon G \to GL(V)$ be a finite-dimensional representation of a finite or compact group G. Then in V there exists an inner product (a unitary form $(\cdot,\cdot)$) which is invariant under the representation T.*

Proof. This is a standard fact from representation theory. On the group G there exists a Haar measure dg, i.e., an invariant finite measure (assume its total mass

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

is 1). Let $(\cdot,\cdot)_0$ be any unitary form, define

$$(u,v) = \int_G (u, T(g)v)_0 dg.$$

Then it is easy to check that $(\cdot,\cdot)$ is invariant. □

3.2 Characters

Among all representations the irreducible ones are distinguished; they are building blocks for all other finite-dimensional representations. This is best seen using the concept of characters.

Definition 3.2.1. Let T be a finite-dimensional complex representation of a finite or compact group G; its *character* χ_T is a function on G defined as

$$\chi_T(g) := \operatorname{Tr} T(g).$$

Some properties of characters are:

1. $\chi(ab) = \chi(ba)$, i.e., character is constant on conjugacy classes in the group G

2. $\chi(e) = \dim V$, the dimension of the representation.

3. χ is positive-definite, that is,

$$\chi(g) = \overline{\chi(g^{-1})},$$

   and for any k, any complex $z_1, \dots, z_k$ and any $g_1, \dots, g_k \in G$ one has

$$\sum z_i \overline{z_j} \chi(g_i g_j^{-1}) \geq 0.$$

   The second property can be seen if we consider the operator

$$A := \sum z_i g_i.$$

   Then the above double sum over i and j is equal to $\operatorname{Tr}(AA^*)$, where A^* is the conjugate operator with respect to that unitary invariant form of Proposition 3.1.3.

4. Character defines a representation uniquely.

5. Consider the space of functions which are constant on conjugacy classes; define an inner product in it:

$$(\varphi,\psi)_G := \frac{1}{|G|} \sum_{g \in G} \varphi(g)\overline{\psi(g)} = \int_G \varphi(g)\overline{\psi(g)} dg$$

   (first formula for finite, and second formula for compact groups; here dg is the Haar measure).

   Then irreducible characters form an *orthonormal* basis in this space.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

6. There is one more structure related to the characters. Let us consider the case of finite groups. Let $\mathbb{C}[G]$ be the group algebra, it is a $*$-algebra with convolution $\varphi * \psi$ (which gives a function on G as a result) and $\varphi^*(g) := \overline{\varphi(g^{-1})}$. Irreducible characters satisfy $(\chi^\pi)^* = \chi^\pi$ (because the representations are unitarizable), and

$$\chi^{\pi_1} * \chi^{\pi_2} = \begin{cases} \dfrac{\chi^\pi}{\dim \pi}, & \pi_1 \text{ is equivalent to } \pi_2 \\ 0, & \text{else.} \end{cases}$$

Consider the normalized irreducible characters

$$\widehat{\chi^\pi}(\cdot) := \frac{\chi^\pi(\cdot)}{\dim \pi}.$$

Such characters have the properties:

- $\widehat{\chi^\pi}(e) = 1$
- $\widehat{\chi^\pi}$ is constant on conjugacy classes
- $\widehat{\chi^\pi}$ is positive definite.
- $\widehat{\chi^\pi}$ is continuous (if G is compact).

Let us consider the space of all functions φ on G which satisfy the above four properties.

Proposition 3.2.2. *A function φ on G is normalized, constant on conjugacy classes, positive definite, and continuous if and only if it is a convex combination of normalized irreducible characters:*

$$\varphi = \sum_\pi c_\pi \frac{\chi^\pi}{\dim \pi} \tag{3.2.1}$$

Proof. That a function (3.2.1) satisfies the four properties is obvious.

Let us take some function φ which is constant on conjugacy classes. Since the characters form a basis in this space, there exists a decomposition of the form (3.2.1). Moreover, $\sum c_\pi = 1$ because φ is normalized. It remains to show that $c_\pi \geq 0$.

This is done by using convolution operation in the group algebra $\mathbb{C}[G]$.[1] Let $\psi := \chi^{\pi_0}$ for some fixed irreducible representation π_0. The function

$$(\psi^*) * \varphi * \psi$$

is positive-definite (this is readily checked), and also

$$(\psi^*) * \varphi * \psi = c_{\pi_0} \frac{\chi^{\pi_0}}{(\dim \pi_0)^3}.$$

This is positive definite because χ^{π_0} is; so $c_{\pi_0} \geq 0$. □

[1]For compact groups we use $L^2(G)$ with the Haar measure.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

So we see that the functions on the (finite or compact) group which are normalized, positive definite, constant on conjugacy classes and continuous, form a convex set whose extreme points are precisely the normalized irreducible characters. This allows to give the following definition.

Definition 3.2.3 ((Normalized) characters of a topological group)**.** Let G be a topological group. A function φ on G which is

1. normalized, $\varphi(e) = 1$
2. constant on conjugacy classes
3. positive definite
4. continuous in the topology of G

is called a (formal) *character* of G.

Such characters form a convex set; extreme points of this convex set we will call extreme characters; they serve as a natural replacement for irreducible characters of finite or compact groups.

There is a theory (von Neumann factors; operator algebras) which associates true representations with these characters; basically, there the notion of a trace is also "axiomatized". We do not discuss those topics here.

Remark 3.2.4. For a finite group G, normalized traces $\chi^\pi / \dim \pi$ of irreducible representations can be characterized as unique functions on G taking value 1 at $e \in G$ such that

$$\frac{1}{|G|} \sum_{h \in G} \chi(g_1 h g_2 h^{-1}) = \chi(g_1)\chi(g_2)$$

(this is the so-called functional equality for characters).

If "$|G| \to \infty$", this equation may simplify. This comment is related to a certain property of "big" groups we will discuss in the course.

3.3 Examples

Our representation-theoretic side in the course is not to consider representations of just one group, but organize groups into a sequence, and study characters of all the groups simultaneously. The group sequences we will consider are symmetric and unitary groups.

3.3.1 Characters of finite symmetric and finite-dimensional unitary groups

Let us give some examples of characters of groups that we will be interested in — (finite) symmetric and (finite-dimensional) unitary groups.

Some normalized characters of $\mathfrak{S}(n)$:

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

1. $\chi(g) = 1$ — trivial representation
2. $\chi(g) = \mathrm{sgn}(g)$ — sign representation
3. $\chi(g) = \begin{cases} 1, & g = e; \\ 0, & g \neq e \end{cases}$

The first two representations are irreducible, and the third one is not; the third is regular representation: action of $\mathfrak{S}(n)$ in $\mathbb{C}[\mathfrak{S}(n)]$. It is decomposed as

$$\bigoplus_{\pi \in \widehat{\mathfrak{S}(n)}} \dim \pi \cdot \pi,$$

which leads to the identity:

$$\sum_{\pi \in \widehat{\mathfrak{S}(n)}} \frac{\dim^2 \pi}{|G|} \frac{\chi^\pi}{\dim} \pi = \begin{cases} 1, & g = e; \\ 0, & g \neq e \end{cases},$$

so

$$\sum_{\pi \in \widehat{\mathfrak{S}(n)}} \frac{\dim^2 \pi}{|G|} = 1$$

(this is also sometimes called Burnside's theorem).

Some normalized characters of $U(N)$:

1. $\chi(g) = 1$ — trivial representation
2. $\chi(g) = \det(g)^k$, $k \in \mathbb{Z}$.

All these representations are irreducible.

An example of a reducible representation:

$$\chi(g) = \det(1 + g).$$

Decomposition of this representation into irreducibles is rather straightforward once we know all the irreducibles. We will discuss this shortly.

3.3.2 Big groups

We need Definition 3.2.3 to discuss characters of big groups such as the infinite symmetric group and the infinite-dimensional unitary group.

Definition 3.3.1. Consider the increasing chain of symmetric groups,

$$\mathfrak{S}(1) \subset \mathfrak{S}(2) \subset \mathfrak{S}(3) \subset \ldots$$

The subgroup $\mathfrak{S}(n) \subset \mathfrak{S}(n+1)$ fixes the $(n+1)$st index.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

The union of these groups is called the infinite symmetric group

$$\mathfrak{S}(\infty) := \bigcup_{n=1}^{\infty} \mathfrak{S}(n).$$

This group may be viewed as a group of bijections of a countable set which fix almost all elements.

Definition 3.3.2. Consider the increasing chain of finite-dimensional unitary groups

$$U(1) \subset U(2) \subset U(3) \subset \dots, \tag{3.3.1}$$

where the inclusions are defined as

$$U(N-1) \ni U \mapsto \begin{bmatrix} U & 0 \\ 0 & 1 \end{bmatrix} \in U(N). \tag{3.3.2}$$

The union of these groups gives the infinite-dimensional unitary group

$$U(\infty) := \bigcup_{N=1}^{\infty} U(N).$$

Elements of $U(\infty)$ are infinite ($\mathbb{N} \times \mathbb{N}$) unitary matrices which differ from the identity matrix only in a fixed number of positions.

There are two separate (but related) theories of characters of the two big groups $\mathfrak{S}(\infty)$ and $U(\infty)$. In the course we will deal with $U(\infty)$ in detail, but also will give definitions and formulate results for the infinite symmetric group.

These two problems sound as "describe all extreme characters of the group $\mathfrak{S}(\infty)$ or $U(\infty)$". In both cases the set of parameters of extreme characters would be infinite-dimensional. Any character of the group can be uniquely decomposed into a (continuous) convex combination of extremes. This decomposition leads to a probability measure on this infinite-dimensional set of parameters.

3.4 References

The material on characters and representations of finite/compact groups can be found in any textbook on representation theory, e.g., [FH91].

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Chapter 4

Symmetric polynomials and symmetric functions. Extreme characters of $U(\infty)$ and $\mathfrak{S}(\infty)$

The theory of symmetric functions is a framework which we will use in describing characters of both the (finite-dimensional) unitary groups and the (finite) symmetric groups. In the case of unitary groups this can be done in a more straightforward way, and we start discussing symmetric functions from this point. Then we make appropriate remarks on characters of symmetric groups.

The theory of symmetric function is also very useful in many other areas.

4.1 Symmetric polynomials, Laurent-Schur polynomials and characters of unitary groups

4.1.1 Laurent symmetric polynomials

Consider $\mathbb{C}[u_1^{\pm 1}, \dots, u_N^{\pm 1}]$ — the algebra of Laurent polynomials in N variables.

1. The basis in $\mathbb{C}[u_1^{\pm 1}, \dots, u_N^{\pm 1}]$ is formed by monomials of the form $u_1^{m_1} \dots u_n^{m_N}$, $m_i \in \mathbb{Z}$.

2. This is a graded algebra with the usual grading.

3. The symmetric group $\mathfrak{S}(N)$ acts on $\mathbb{C}[u_1^{\pm 1}, \dots, u_N^{\pm 1}]$ by permutations of variables.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

The symmetric polynomials — i.e., elements of $\mathbb{C}[u_1^{\pm 1}, \dots, u_N^{\pm 1}]$ invariant under $\mathfrak{S}(N)$ — form a subalgebra in $\mathbb{C}[u_1^{\pm 1}, \dots, u_N^{\pm 1}]$:

$$LSym(N) := \mathbb{C}[u_1^{\pm 1}, \dots, u_N^{\pm 1}]^{\mathfrak{S}(N)}.$$

We call $LSym(N)$ the algebra of symmetric Laurent polynomials in N variables. This algebra is also graded.

A (simplest possible) basis in this algebra is formed by the homogeneous symmetrized monomials. We will index these monomials by the so-called *signatures* of length N:

$$\mu = (\mu_1 \ge \mu_2 \ge \ldots \ge \mu_N), \qquad \mu_i \in \mathbb{Z}$$

The symmetrized monomial looks as

$$m_\mu(u_1, \dots, u_N) = \sum_{\nu \in \mathfrak{S}(N)\mu} u_1^{\nu_1} \dots u_N^{\nu_N}$$

That is, the sum is taken over all *distinct* permutations of μ. That is, all individual monomials in m_μ have coefficients 1.

4.1.2 Anti-symmetric polynomials

Definition 4.1.1. A Laurent polynomial $f \in \mathbb{C}[u_1^{\pm 1}, \dots, u_N^{\pm 1}]$ is called anti-symmetric if

$$f(gu) = \operatorname{sgn}(g) \cdot f(u), \qquad g \in \mathfrak{S}(N).$$

The anti-symmetric polynomials form a subspace (also a module over *LSym*):

$$Alt(N) \subset \mathbb{C}[u_1^{\pm 1}, \dots, u_N^{\pm 1}].$$

A basis in $Alt(N)$ is formed by anti-symmetrized monomials (alternants)

$$a_\mu(u_1, \dots, u_N) := \sum_{g \in \mathfrak{S}(N)} \operatorname{sgn}(g) u^{g \cdot \mu} = \det \begin{bmatrix} u_1^{\mu_1} & \dots & u_N^{\mu_1} \\ \dots & \dots & \dots \\ u_1^{\mu_N} & \dots & u_N^{\mu_N} \end{bmatrix}.$$

If μ has equal parts, then $a_\mu = 0$, so the alternants are parametrized by signatures with distinct parts.

Denote the special "staircase" signature

$$\delta := (N-1, N-2, \dots, 1, 0).$$

It corresponds to the Vandermonde determinant

$$V_N(u_1, \dots, u_N) := a_\delta(u_1, \dots, u_N) = \prod_{1 \le i < j \le N} (u_i - u_j)$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Proposition 4.1.2. *Symmetric and anti-symmetric Laurent polynomials in N variables are in a bijective correspondence. Every anti-symmetric Laurent polynomial is divisible by the Vandermonde $V_N(u_1, \dots, u_N)$, and the result is a symmetric polynomial:*

$$Alt(N) = V_N \cdot LSym(N).$$

Proof. 1) First, we reduce this statement to ordinary polynomials because for every Laurent polynomial $f(u)$ there exists a k such that $f(u) \cdot (u_1, \dots, u_N)^k$ is an ordinary polynomial.

2) Let $f \in Alt(N)$. Clearly, $f(u_1, \dots, u_N)$ vanishes if $u_i = u_j$, therefore $u_i - u_j$ divides f for all $i < j$. Thus, $f(u_1, \dots, u_N)$ is divisible by the Vandermonde determinant. The result is clearly a symmetric polynomial. □

4.1.3 Schur polynomials

Taking the most obvious basis in the space $Alt(N)$ of anti-symmetric Laurent polynomials, namely, the alternants a_μ (μ runs over signatures with distinct parts), and transferring these alternants to symmetric polynomials via Proposition 4.1.2, we get a distinguished basis in the algebra of symmetric Laurent polynomials $LSym(N)$. These polynomial are called the *Laurent-Schur polynomials*:

Definition 4.1.3. Let λ be a signature of length N, $\lambda_1 \geq \lambda_2 \geq \dots \lambda_N$, $\lambda_i \in \mathbb{Z}$. Define the Laurent-Schur polynomials as

$$s_\lambda = s_\lambda(u_1, \dots, u_N) := \frac{a_{\lambda+\delta}(u_1, \dots, u_N)}{a_\delta(u_1, \dots, u_N)} = \frac{\det[u_j^{\lambda_i+N-i}]_{i,j=1}^N}{\det[u_j^{N-i}]_{i,j=1}^N} \in LSym(N).$$

Examples:

- A concrete example:

$$\begin{aligned} & s_{(3,2,0)}(u_1, u_2, u_3) \\ &= \frac{1}{(u_1-u_2)(u_1-u_3)(u_2-u_3)} \det \begin{bmatrix} u_1^5 & u_2^3 & 1 \\ u_2^5 & u_2^3 & 1 \\ u_3^5 & u_3^3 & 1 \end{bmatrix} \\ &= u_1^3u_2^2 + u_1^3u_2u_3 + u_1^3u_3^2 + u_1^2u_2^3 + 2u_1^2u_2^2u_3 + 2u_1^2u_2u_3^2 \\ &\quad + u_1^2u_3^3 + u_1u_2^3u_3 + 2u_1u_2^2u_3^2 + u_1u_2u_3^3 + u_2^3u_3^2 + u_2^2u_3^3 \\ &= m_{(3,2,0)} + m_{(3,1,1)} + 2m_{(2,2,1)} \end{aligned}$$

- We have for $\lambda = 0$, $s_\lambda = 1$, the character of the trivial representation.

- If $\lambda = (k, k, \dots, k)$, then (EX) $s_\lambda(u_1, \dots, u_N) = (u_1 \cdot \dots \cdot u_N)^k$, $k \in \mathbb{Z}$. The character has the form $(\det(U))^k$, this is one-dimensional representation.

Properties:

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

- These are homogeneous polynomials, $\deg s_\lambda = |\lambda| := \lambda_1 + \ldots + \lambda_N$ (this number is not necessary nonnegative)
- The polynomials s_λ form a basis in $\mathbb{C}[u_1^{\pm 1}, \ldots, u_N^{\pm 1}]$
- (Kostka numbers) The polynomials s_λ are expanded in the basis of symmetrized monomials as

  $$s_\lambda = \sum_\mu K_{\lambda\mu} m_\mu,$$

  where the numbers $K_{\lambda\mu}$ turn out to be nonnegative integers. They are called Kostka numbers. We also have $K_{\lambda\lambda} = 1$.

4.1.4 Orthogonality of Schur polynomials

Let now u_i — be not formal variables but belong to the unit circle:

$$u_i \in \mathbb{T} := \{z \in \mathbb{C} \colon |z| = 1\}, \qquad i = 1, \ldots, N. \tag{4.1.1}$$

On $\mathbb{T}$ consider the uniform measure: $l(du) = d\varphi/2\pi$, where $u = e^{i\varphi}$.

Definition 4.1.4. Define the inner product for $f, g \in LSym(N)$:

$$(f, g) := \frac{1}{N!} \int_{\mathbb{T}} \cdots \int_{\mathbb{T}} f(u)\overline{g(u)} |V_N(u)|^2 l(du_1) \ldots l(du_N). \tag{4.1.2}$$

Here $V_N(u)$ — is the Vandermonde determinant.

It can be readily seen that $(\cdot, \cdot)$ is nondegenerate sesquilinear form.

Proposition 4.1.5. *The Schur polynomials s_λ form an orthonormal basis with respect to the inner product $(\cdot, \cdot)$.*

Proof. By definition of s_λ, $s_\lambda = \frac{a_{\lambda+\delta}}{a_\delta}$, $s_\mu = \frac{a_{\mu+\delta}}{a_\delta}$. Set $\widetilde{\lambda} = \lambda + \delta$, $\widetilde{\mu} = \mu + \delta$. Note that $|V_N(u)|^2 = a_\delta(u)\overline{a_\delta(u)}$, so

$$(s_\lambda, s_\mu) = \frac{1}{N!} \int_{\mathbb{T}} \cdots \int_{\mathbb{T}} a_{\widetilde{\lambda}}(u)\overline{a_{\widetilde{\mu}}(u)} l(du_1) \ldots l(du_N). \tag{4.1.3}$$

By definition of alternants[1]

$$a_{\widetilde{\lambda}}(u)\overline{a_{\widetilde{\mu}}(u)} = \sum_{\sigma, \tau \in Sym_N} \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau) \cdot u^\alpha \overline{u}^\beta = \sum_{\sigma, \tau \in Sym_N} \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau) \cdot u^{\alpha - \beta}, \tag{4.1.4}$$

where the summation is over all $\sigma, \tau \in Sym_N$, and $\alpha = \sigma\widetilde{\lambda}$, $\beta = \tau\widetilde{\mu}$.

[1] Clearly, $\overline{u} = u^{-1}$.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Integral $\int_{\mathbb{T}^N} u^{\alpha-\beta} l(du_1)\dots l(du_N)$ is nonzero only if $\alpha = \beta$.[2] Since α and β are permutations of $\widetilde{\lambda}$ and $\widetilde{\mu}$, respectively, then for $\lambda \neq \mu$ we have $(s_\lambda, s_\mu) = 0$.

Let now $\lambda = \mu$. Since $\widetilde{\lambda}$ and $\widetilde{\mu}$ are signatures in which all parts are distinct, then Sym_N acts transitively on them and the size of the orbit is $N!$. This $N!$ cancels with $\frac{1}{N!}$ from (4.1.2). This concludes the proof. □

Remark 4.1.6 (How to decompose with respect to Schur polynomials)**.** Let $f \in LSym(N)$. Since the polynomials s_λ form a basis in $LSym(N)$, there exists a unique decomposition of the form

$$f = \sum_\lambda f_\lambda s_\lambda. \tag{4.1.5}$$

There exists a way to effectively obtain coefficients f_λ, namely,

$$f_\lambda = \left[u^{\lambda+\delta}\right](f \cdot V_N), \tag{4.1.6}$$

where $\left[u^{\lambda+\delta}\right](\dots)$ means the coefficient by $u^{\lambda+\delta}$ in $(\dots)$.

Indeed, from the definition of the Schur polynomials,

$$f = \sum_\lambda f_\lambda s_\lambda \qquad \Leftrightarrow \qquad f \cdot V_N = \sum_\lambda f_\lambda a_{\lambda+\delta}. \tag{4.1.7}$$

4.1.5 Characters of the unitary groups. Radial part of the Haar measure

We reproduce the following fact without proof:

Proposition 4.1.7. *The irreducible representations of the unitary group $U(N)$ are parametrized by signatures $\lambda \in \mathbb{GT}_N$. The corresponding irreducible characters look like*

$$\chi^\lambda(U) = s_\lambda(u_1, \dots, u_N),$$

where the u_i's are the eigenvalues of the matrix $U \in U(N)$.

An indication of this[3] is the fact that the Schur polynomials are orthonormal with respect to the inner product (4.1.2). Namely, the "radial part" of the Haar measure on $U(N)$ generates the same inner product as (4.1.2). Let us give more details.

First, as an example of what radial part means, note the polar change of variables:

$$\int_{\mathbb{R}^2} f(x,y)dxdy = \int_{\mathbb{R}_{>0}} \int_0^{2\pi} f(r,\varphi) \cdot rdr \cdot d\varphi.$$

[2]Note that this is the Lebesgue integral and not the complex integral: $\int_0^{2\pi} e^{ic\varphi} d\varphi = 0$ for all $c \in \mathbb{R}$, $c \neq 0$.

[3]But of course not the whole proof.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

If f depends only on r, then the radial part would be the distribution rdr on $\mathbb{R}_{>0}$.

Now let $f(U)$ be a function on the unitary group $U(N)$ which depends only on eigenvalues $(u_1, \dots, u_N)$. That is, f is constant on conjugacy classes (= a central function). We would like to write an integral of the form:

$$\int_{U \in U(N)} f(U) dU = \int_{\mathbb{T}^N} (?) l(du_1) \dots l(du_N)$$

where dU is the Haar measure on $U(N)$, and $l(du_1) \dots l(du_N)$ is the normalized Lebesgue measure on $\mathbb{T}^N$. We would like to find the Jacobian (?) of the change of variables.

Proposition 4.1.8. *The Jacobian is equal to*

$$|V(u_1, \dots, u_N)|^2 = \text{const} \prod_{1 \le i < j \le N} |u_i - u_j|^2$$

Proof. We present an argument which can be turned into proof. It belongs to Weyl [Wey97].

Let $U \in U(N)$ be the matrix we are integrating over; it can be diagonalized as

$$U = V \operatorname{diag}(u_1, \dots, u_N) V^{-1},$$

where $V \in U(N)$. That is, we introduce coordinates on $U(N)$,

$$U \leftrightarrow (V, (u_1, \dots, u_N)).$$

In fact, this is not bijection, and we must consider V modulo stabilizer of $u = (u_1, \dots, u_N)$.

We want to write

$$dU = (?) l(du_1) \dots l(du_N) d(V/stab)$$

Since the Haar measure is invariant, the Jacobian cannot depend on V.

Let $u_i = e^{2\pi \mathrm{i} \varphi_i}$, so $du_i = 2\pi \mathrm{i} u_i d\varphi_i$. Since

$$UV = Vu,$$

we have

$$dU \cdot V + U \cdot dV = dV \cdot u + V \cdot 2\pi \mathrm{i} u d\varphi.$$

dV is anti-Hermitian (EX), and we may assume that it has zeroes on the diagonal.

Multiply the left-hand side by $V^{-1}U^{-1}$, and the right-hand side by $u^{-1}V^{-1}$ (which is equal to that). We get

$$V^{-1}U^{-1} dU \cdot V + V^{-1} \cdot dV = u^{-1} V^{-1} dV \cdot u + 2\pi \mathrm{i} d\varphi;$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

We must consider $U^{-1}dU$, an infinitesimal element, which is the same (in terms of the volume element) as $V^{-1}U^{-1}dU \cdot V$. We have

$$(V^{-1}U^{-1}dU \cdot V)_{ij} = \begin{cases} 2\pi \mathrm{i} d\varphi_i, & i = j; \\ (u_j/u_i - 1)(V^{-1}dV)_{ij}. \end{cases}$$

Observe that the factors $(u_j/u_i - 1)$ come in pairs,

$$(u_i/u_j - 1) = \overline{(u_j/u_i - 1)}.$$

The infinitesimal volumes $U^{-1}dU$ and $V^{-1}dV$ just correspond to taking both d's around unit elements of the group. We conclude that the desired Jacobian must be equal to

$$\mathrm{const} \prod_{k<j} |u_j - u_k|^2 = \mathrm{const} |V(u_1, \ldots, u_N)|^2.$$

□

Remark 4.1.9. This statement is one of the foundations of the theory of random matrices: one starts with a matrix, and asks how the distribution of eigenvalues looks like.

In this way one gets the circular unitary ensemble; the distribution of N particles on the circle with Vandermonde-square density (repelling potential).

Remark 4.1.10. To complete the proof that the Schur polynomials are irreducible characters, one needs the following property from the theory of representations of Lie groups.

We say that signature $\lambda \in \mathbb{GT}_N$ *dominates* the signature $\mu \in \mathbb{GT}_N$ iff μ can be obtained from λ by a sequence of operations of the form

$$\nu \mapsto \nu - \varepsilon_i + \varepsilon_j, \qquad i < j,$$

where $\varepsilon_k = (0, \ldots, 0, 1, 0, \ldots, 0)$ (1 is a the kth position), and all intermediate $\nu - \varepsilon_i + \varepsilon_j$ must be signaturesas well. Equivalently:

$$\lambda \text{ dominates } \mu \text{ iff } \quad \sum_{i=1}^{j} \mu_i \le \sum_{i=1}^{j} \lambda_i \quad \text{for all } j = 1, \ldots, N.$$

In this way we have a partial order on the set of signatures $\lambda \in \mathbb{GT}_N$ with fixed $|\lambda|$.

Then the irreducible characters of $U(N)$ are determined by the two conditions:

1. They are parametrized by signatures $\lambda \in \mathbb{GT}_N$, and we have

$$\chi^\lambda = m_\lambda + \sum_{\mu\colon |\mu|=|\lambda|,\ \mu \ne \lambda,\ \lambda \text{ dominates } \mu} c_{\lambda\mu} m_\mu$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

for some coefficients $c_{\lambda\mu}$. This equality should be viewed as an identity of functions on the N-dimensional torus

$$\mathbb{T}^N = \{(u_1, \dots, u_N) \in \mathbb{C}^N \colon |u_i| = 1\}.$$

The summand m_λ is the highest weight of the representation, and all other summands are referred to as lower terms.

2. χ^λ's form an orthonormal basis in the (L^2) space of central functions on $U(N)$.

We've discussed the second property (at least for finite groups; but for compact groups this can be done in a similar way). The first property is coming from the theory of representations of Lie groups.

4.1.6 Branching of Schur polynomials and restriction of characters

Definition 4.1.11. If μ is a signature of length $N-1$, λ is a signature of length N (we write this as $\mu \in \mathbb{GT}_{N-1}$, $\lambda \in \mathbb{GT}_N$), we say that μ and λ interlace (notation $\mu \prec \lambda$) iff

$$\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \ldots \geq \mu_{N-1} \geq \lambda_N.$$

We will discuss the following problem now: assume that we take a representation of $U(N)$ corresponding to a signature $\lambda \in \mathbb{GT}_N$, and restrict it to the subgroup $U(N-1) \subset U(N)$ which is embedded as in §3.3.1. Then this same representation will become a representation of $U(N-1)$. In general, this new representation needs not to be irreducible. How this representation of $U(N-1)$ is decomposed into irreducibles?

The answer is given by the following branching rule for Schur polynomials:

Proposition 4.1.12. *We have*

$$s_\lambda(u_1, \dots, u_{N-1}; u_N = 1) = \sum_{\mu\colon \mu \prec \lambda} s_\mu(u_1, \dots, u_{N-1}).$$

Proof. Let $N = 3$ for simplicity (for general N all is done in the same way).

Let $\lambda = (\lambda_1 \geq \lambda_2 \geq \lambda_3)$, $\lambda_i \in \mathbb{Z}$. We have

$$s_\lambda(u_1, u_2, u_3) = \frac{1}{V_3(u_1, u_2, u_3)} \begin{vmatrix} u_1^{\lambda_1+2} & u_2^{\lambda_1+2} & u_3^{\lambda_1+2} \\ u_1^{\lambda_2+1} & u_2^{\lambda_2+1} & u_3^{\lambda_2+1} \\ u_1^{\lambda_3} & u_2^{\lambda_3} & u_3^{\lambda_3} \end{vmatrix}. \tag{4.1.8}$$

Setting $u_3 = 1$, we write:

$$s_\lambda(u_1, u_2, 1) = \frac{1}{V_2(u_1, u_2) \cdot (u_1 - 1)(u_2 - 1)} \begin{vmatrix} u_1^{\lambda_1+2} & u_2^{\lambda_1+2} & 1 \\ u_1^{\lambda_2+1} & u_2^{\lambda_2+1} & 1 \\ u_1^{\lambda_3} & u_2^{\lambda_3} & 1 \end{vmatrix}. \tag{4.1.9}$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Subtracting the 2nd row from the 1st, and then the 3rd from the 2nd, we get

$$s_\lambda(u_1, u_2, 1) = \frac{1}{V_2(u_1, u_2) \cdot (u_1 - 1)(u_2 - 1)} \begin{vmatrix} u_1^{\lambda_1+2} - u_1^{\lambda_2+1} & u_2^{\lambda_1+2} - u_2^{\lambda_2+1} \\ u_1^{\lambda_2+1} - u_1^{\lambda_3} & u_2^{\lambda_2+1} - u_2^{\lambda_3} \end{vmatrix}. \tag{4.1.10}$$

We see that the 1st column can be divided by $u_1 - 1$, and the second one by $u_2 - 1$.

Note that for $a > b$:

$$\frac{u^a - u^b}{u - 1} = \sum_{c\colon\, b \le c < a} u^c. \tag{4.1.11}$$

We see that by dividing by $u_1 - 1$ in the first column in the first place we get a sum which can be written as

$$\sum_{\mu_1} u_1^{\mu_1+1}, \tag{4.1.12}$$

where $1 + \lambda_2 \le \mu_1 + 1 < \lambda_1 + 2$ (i.e., $\lambda_2 \le \mu_1 \le \lambda_1$). We choose to call the summation index $\mu_1 + 1$ so that we will have the desired result. Considering three other matrix elements in a similar way, we conclude the proof. ☐

Therefore, we may say that

$$\pi_N^\lambda = \bigoplus_{\mu\colon\, \mu \prec \lambda} \pi_{N-1}^\mu.$$

That is, the representation of $U(N)$ indexed by λ restricted to $U(N-1)$ is decomposed into a multiplicity-free direct sum of representations of $U(N-1)$ indexed by μ's such that $\mu \prec \lambda$.

In fact, trivial degree considerations lead to the following property which is more general than Proposition 4.1.12:

Proposition 4.1.13. *We have*

$$s_\lambda(u_1, \ldots, u_{N-1}, u_N) = \sum_{\mu\colon\, \mu \prec \lambda} s_\mu(u_1, \ldots, u_{N-1}) u_N^{|\lambda| - |\mu|}.$$

In this way one can start to analyze the representations of all unitary groups $U(1) \subset U(2) \subset U(3) \subset \ldots$ as a chain (i.e., consider representations simultaneously), and this will lead to the theory of characters of the infinite-dimensional unitary group. We will pause here for a while, and consider more properties and definitions regarding symmetric polynomials. In this way we will describe the characters of finite symmetric groups, and derive a branching rule for them. Then we will return to the questions of characters of $\mathfrak{S}(\infty)$ and $U(\infty)$ in parallel.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

4.1.7 Combinatorial formula for Schur polynomials

Continuing the expansion of Proposition 4.1.13, we get the following combinatorial formula for Schur polynomials:

$$s_\lambda(u_1,\dots,u_N) = \sum_{\lambda^{(1)}\prec\lambda^{(2)}\prec\dots\prec\lambda^{(N-1)}\prec\lambda} u_1^{|\lambda^{(1)}|}u_2^{|\lambda^{(2)}|-|\lambda^{(1)}|}u_3^{|\lambda^{(3)}|-|\lambda^{(2)}|}\dots u_N^{|\lambda|-|\lambda^{(N-1)}|} \tag{4.1.13}$$

Here the sum is taken over all sequences of signatures (they are sometimes called Gelfand–Tsetlin schemes).

A Gelfand–Tsetlin scheme may be also viewed as a triangular array of integers as on Fig. 4.1.

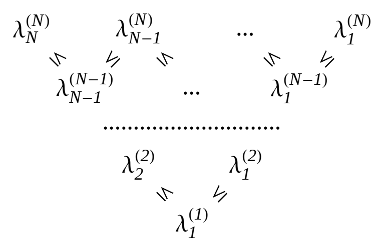
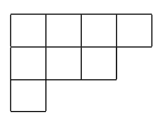


Figure 4.1: A Gelfand–Tsetlin scheme of depth N.

For nonnegative signatures, this has another combinatorial interpretation, see the next subsection:

4.1.8 Semistandard Young tableaux

Let us take a signature with nonnegative parts, we denote that by $\lambda \in \mathbb{GT}_N^+$. We associate with λ a so-called *Young diagram* — a set of boxes in the plane with λ_1 boxes in the first row, λ_2 boxes in the second row, etc. For example, if $\lambda = (4,3,1)$, then the Young diagram looks as:



A *Young tableaux* of shape $\lambda \in \mathbb{GT}_N^+$ with entries $\{1,2,\dots,N\}$ is a map $T\colon \square \to \{1,\dots,N\}$, where $\square \in \lambda$, i.e., a box of the Young diagram. In other words, a Young tableau is a filling of the boxes of the Young diagram with entries $\{1,2,\dots,N\}$.

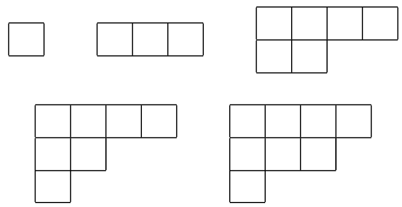
A Young tableau is called *semistandard* if the entries weakly increase along rows and strictly increase down columns. An example:

1	2	2	3
3	3	5	
4			

(4.1.14)

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Any semistandard Young tableau can be viewed as a sequence of Young diagrams which correspond to interlacing nonnegative signatures. In the above example:



One partition differs from another by adding a so-called horizontal strip — a set of boxes which contains no more than one box in each column. This condition comes from the interlacing constraint, and we readily see that it corresponds to semistandardness of the Young tableau.

Thus, we have a bijection between Gelfand–Tsetlin schemes with fixed top row $\lambda \in \mathbb{GT}_N^+$, and semistandard Young tableaux of shape λ filled with $\{1, \ldots, N\}$.

We may also restate the above combinatorial formula (4.1.13) for Schur polynomials indexed by nonnegative signatures (note that then these polynomials are honest, not Laurent polynomials) in terms of semistandard Young tableaux:

$$s_\lambda(u_1, \ldots, u_N) = \sum_T x^T, \tag{4.1.15}$$

where the sum is taken over all semistandard Young tableaux of shape λ, and x^T is the product of $x_{T(\square)}$ over all boxes $\square \in \lambda$. For example, to the semistandard Young tableau above corresponds $x^T = x_1 x_2^2 x_3^3 x_4 x_5$.

Remark 4.1.14. From (4.1.15) — as well as from (4.1.13) — it is not clear that $s_\lambda(u_1, \ldots, u_N)$ is symmetric. However, of course it is symmetric as the definition with determinants shows.

4.1.9 Elementary symmetric polynomials

One important example of symmetric polynomials are the elementary symmetric polynomials

$$e_k(u_1, \ldots, u_N) := \sum_{1 \le i_1 < \ldots < i_k \le N} u_{i_1} u_{i_2} \ldots u_{i_k}, \qquad k = 0, 1, 2, \ldots. \tag{4.1.16}$$

We have $e_0 = 1$ and $e_k = 0$ if $k > N$.

In fact, these polynomials are particular cases of Schur polynomials:

$$e_k(u_1, \ldots, u_N) = s_{(\underbrace{1,1,\ldots,1}_{k},0,\ldots,0)}(u_1, \ldots, u_N).$$

This is readily seen from the combinatorial formula for the Schur polynomials (4.1.15) (EX).

There are other distinguished symmetric polynomials which are indexed by integers — the so-called *complete homogeneous symmetric polynomials* which are defined for all $m = 0, 1, 2, \ldots$ as

$$h_m(u_1, \ldots, u_N) = s_{(m,0,0,\ldots,0)}(u_1, \ldots, u_N).$$

From the combinatorial interpretation of the Schur polynomials we may readily see that

$$h_m(u_1, \ldots, u_N) = \sum_{1 \le i_1 \le \ldots \le i_m \le N} u_{i_1} \ldots u_{i_N}. \tag{4.1.17}$$

Note that $h_0 = 1$, and h_m for $m \geq 0$ is the sum of all possible symmetrized monomials of degree m. Thus, the h-polynomials do not share the vanishing $e_{N+1} = e_{N+2} = \ldots = 0$ which holds for the e-polynomials.

Remark 4.1.15. We will later see that the h-functions and e-functions are algebraically independent if one considers them in infinitely many variables. However, in N variables the h-functions of higher order $h_{N+1}, h_{N+2}, \ldots$ can be expressed as polynomials in $h_1, \ldots, h_N$.

4.1.10 Two examples of extreme characters of $U(\infty)$

Let us consider particular families of extreme (=normalized irreducible) characters of the unitary groups $U(N)$, and let $N \to \infty$. We hope to get extreme characters of the infinite-dimensional unitary group $U(\infty)$. For the "baby" examples we consider, the limits can be obtained by direct computation.

Consider first the elementary symmetric polynomials $e_{L(N)}$ indexed by N, $0 \leq L(N) \leq N$. We are interested in limits of normalized characters

$$\frac{e_{L(N)}(u_1, \ldots, u_N)}{e_{L(N)}(1, \ldots, 1)}.$$

The convergence notion is as follows:

Definition 4.1.16. We say that a sequence of central functions f_N on $U(N)$ converge to a central function f on $U(\infty)$ if for every fixed K we have

$$f_N(u_1, \ldots, u_K, 1, 1, \ldots, 1) \to f(u_1, \ldots, u_K, 1, 1, \ldots, 1)$$

uniformly on the K-torus $\mathbb{T}^K$.

The function f on $U(\infty)$ may be viewed as defined on the infinite-dimensional *finitary torus* $\mathbb{T}^\infty_{fin}$ which is a set of sequences $(u_1, u_2, \ldots)$, $|u_i| = 1$, with only finitely many components different from 1.

Proposition 4.1.17. *Let* $L(N)/N \to \beta \in [0, 1]$. *Then*

$$\frac{e_{L(N)}(u_1, \ldots, u_N)}{e_{L(N)}(1, \ldots, 1)} \to \prod_{i=1}^{\infty} (1 + \beta(u_i - 1)), \qquad (u_1, u_2 \ldots) \in \mathbb{T}^\infty_{fin}.$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Proof. This statement can be proved literally by hand, but we will use a slightly different approach and refer to the de Finetti's theorem.

Consider the space of subsets of $\{1,\dots,N\}$ which have $L(N)$ elements. This is the same as binary sequences of length N with $L(N)$ ones and $N-L(N)$ zeroes. Thus,

$$\frac{e_{L(N)}(u_1,\dots,u_N)}{e_{L(N)}(1,\dots,1)} = \frac{\sum_I u^I}{\sum_I 1},$$

where the sum is taken over all $L(N)$-element subsets. For a fixed K we have

$$\frac{e_{L(N)}(u_1,\dots,u_K,1,1,\dots,1)}{e_{L(N)}(1,\dots,1)}$$
$$= \sum_{\text{binary } K\text{-sequences}} u_1^{\varepsilon_1}\dots u_K^{\varepsilon_K}\,\mathbb{P}_N\{\text{first } K \text{ coordinates of the binary sequence are } (\varepsilon_1,\dots\varepsilon_K)\}$$

From the de Finetti's theorem we know that the limiting distribution as $N\to\infty$ of the binary sequences converges to the Bernoulli measure with parameter β. For it we know that

$$\mathbb{P}_{N=\infty}\{\text{first } K \text{ coordinates of the binary sequence are } (\varepsilon_1,\dots\varepsilon_K)\}$$
$$= \beta^{\sum\varepsilon_i}(1-\beta)^{K-\sum\varepsilon_i}.$$

So,

$$\sum_{\text{binary } K\text{-sequences}} u_1^{\varepsilon_1}\dots u_K^{\varepsilon_K}\,\beta^{\sum\varepsilon_i}(1-\beta)^{K-\sum\varepsilon_i} = \prod_{i=1}^{K}(1-\beta+\beta u_i),$$

and this concludes the proof. □

Another example is with the h-polynomials

$$\frac{h_{L(N)}(u_1,\dots,u_N)}{h_{L(N)}(1,\dots,1)}, \qquad L(N)=0,1,2,\dots$$

Proposition 4.1.18. *Let $L(N)/N\to\alpha\ge 0$, $N\to\infty$. Then in the sense of the above definition we have*

$$\frac{h_{L(N)}(u_1,\dots,u_N)}{h_{L(N)}(1,\dots,1)} \to \prod_{i=1}^{\infty}\frac{1}{1-\alpha(u_i-1)}, \qquad N\to\infty, \qquad (u_1,u_2,\dots)\in\mathbb{T}^{\infty}_{fin}.$$

Proof. This is done similarly to the case of the e-polynomials, we just consider L-point multisets in $\{1,2,\dots,N\}$. Their limit is a product of (independent) geometric distributions on $\mathbb{Z}_{\ge 0}$ with weights

$$\big(1-\alpha/(\alpha+1)\big)\cdot\big(\alpha/(\alpha+1)\big)^m, \qquad m\in\mathbb{Z}_{\ge 0}.$$

Then this property implies that we have the convergence of the normalized h-polynomials to the desired limit. □

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Remark 4.1.19. The above two statements reflect the so-called boson-fermionic correspondence: in the first case we deal with repelling particles, and in the second case two (or more) particles can occupy the same position.

Remark 4.1.20. Using shifts, one can readily get also the following extreme characters of the infinite-dimensional unitary group $U(\infty)$:

$$\prod_{i=1}^{\infty}(1+\beta(u_i^{-1}-1)), \qquad \prod_{i=1}^{\infty}\frac{1}{1-\alpha(u_i^{-1}-1)}.$$

The first corresponds to signatures of the form $(0,0,\dots,0,-1,-1,\dots,-1)$, and the second to $(0,\dots,0,-L)$.

4.1.11 Extreme characters of $U(\infty)$

General theorem: Description of extreme characters

Theorem 4.1.21 (Edrei–Voiculescu)**.** *Extremal characters of $U(\infty)$ are functions $\mathbb{T}^{\infty}_{\text{fin}}$ depending on countably many parameters*

$$\begin{array}{rcl}\alpha^{\pm} & = & (\alpha_1^{\pm}\ge\alpha_2^{\pm}\ge\cdots\ge 0);\\ \beta^{\pm} & = & (\beta_1^{\pm}\ge\beta_2^{\pm}\ge\cdots\ge 0);\\ \gamma^{\pm} & \ge & 0,\end{array} \tag{4.1.18}$$

such that

$$\sum_i \alpha_i^+ + \sum_i \alpha_i^- + \sum_i \beta_i^+ + \sum_i \beta_i^- < \infty, \qquad \beta_1^+ + \beta_1^- \le 1. \tag{4.1.19}$$

These functions have the form

$$\chi_{\alpha^{\pm},\beta^{\pm},\gamma^{\pm}}(u_1,u_2,\dots)=\prod_{j=1}^{\infty}\Phi_{\alpha^{\pm},\beta^{\pm},\gamma^{\pm}}(u_j), \tag{4.1.20}$$

where $\Phi_{\alpha^{\pm},\beta^{\pm},\gamma^{\pm}}(\cdot)$ — is the following continuous function on the one-dimensional torus $\mathbb{T}^1$:

$$\Phi_{\alpha^{\pm},\beta^{\pm},\gamma^{\pm}}(u) := e^{\gamma^+(u-1)+\gamma^-(u^{-1}-1)}\prod_{i=1}^{\infty}\left(\frac{1+\beta_i^+(u-1)}{1-\alpha_i^+(u-1)}\cdot\frac{1+\beta_i^-(u^{-1}-1)}{1-\alpha_i^-(u^{-1}-1)}\right). \tag{4.1.21}$$

Remark 4.1.22. Note that the product of two extremal characters is again an extremal character (which is not the case for $U(N)$). So, the "infinite-dimensional" characters are more complicated (because they depend on infinitely many parameters) and at the same time simpler than the pre-limit ones.

Let us discuss some properties of this statement.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Limits of Schur polynomials as the number of variables goes to infinity. Connection with totally nonnegative sequences

First, let us prove the following:

Lemma 4.1.23. *Let a function $B(u)$, $u \in \mathbb{T}$, be expressed as a Laurent series $B(u) = \sum_{n\in\mathbb{Z}} b_n u^n$. Then for $u_1, \dots, u_N \in \mathbb{T}$, we have*

$$B(u_1) \cdot \ldots \cdot B(u_N) = \sum\nolimits_{\nu\in\mathbb{GT}_N} \det[b_{\nu_i - i + j}]_{i,j=1}^N \cdot s_\nu(u_1, \dots, u_N). \qquad (4.1.22)$$

Proof. Indeed, let us multiply by the Vandermonde, $V(u_1, \dots, u_N)$, and find the coefficient by $u^{\nu+\delta}$ in both parts (as suggested in Remark 4.1.6). Recall that by δ we denote $(N-1, N-2, \dots, 1, 0)$. We have

$$\begin{aligned} &V(u_1, \dots, u_N) B(u_1) \dots B(u_N) \\ &\quad = \sum_{(n_1,\dots,n_N)\in\mathbb{Z}^N} \sum_{\sigma\in\mathfrak{S}(N)} \operatorname{sgn}(\sigma)\, u^{\sigma\cdot\delta} b_{n_1} \dots b_{n_N} u_1^{n_1} \dots u_N^{n_N}. \end{aligned}$$

Here we have simply expanded the product of Laurent series. Now to obtain a coefficient by $u^{\nu\cdot\delta}$ we see that we need to take the corresponding determinant of the b's. This concludes the proof. ☐

The Edrei–Voiculescu theorem implies that every extreme character of $U(\infty)$ can be approximated using a suitable sequence of signatures $\lambda(N) \in \mathbb{GT}_N$, as $N \to \infty$. This approximation effectively looks as

$$\frac{s_{\lambda(N)}(u_1, \dots, u_N)}{s_{\lambda(N)}(1, \dots, 1)} \to \prod_{i=1}^{\infty} \Phi_{\alpha^\pm,\beta^\pm,\gamma^\pm}(u_i),$$

where $(u_1, u_2, \dots) \in \mathbb{T}^\infty_{fin}$; and the convergence is understood as uniform convergence of these functions on finite-dimensional sub-tori $\mathbb{T}^K$, where K is fixed.

We also have that the extreme character $\chi_{\alpha^\pm,\beta^\pm,\gamma^\pm}$ restricted to $U(K) \subset U(\infty)$ can be written as

$$\chi_{\alpha^\pm,\beta^\pm,\gamma^\pm}|_{U(K)}(u_1, \dots, u_K) = \sum_{\varkappa\in\mathbb{GT}_K} \det[\varphi_{\varkappa_i - i + j}]_{i,j=1}^K s_\varkappa(u_1, \dots, u_K).$$

Here φ_n's are the coefficients of the decomposition of $\Phi_{\alpha^\pm,\beta^\pm,\gamma^\pm}$:

$$\Phi_{\alpha^\pm,\beta^\pm,\gamma^\pm}(u) = \sum_{n\in\mathbb{Z}} \varphi_n(\alpha^\pm, \beta^\pm, \gamma^\pm) u^n.$$

Since $\chi_{\alpha^\pm,\beta^\pm,\gamma^\pm}|_{U(K)}$ is obviously a character of $U(K)$, we must have that the coefficients of its decomposition as a linear combination of characters of $U(K)$ are nonnegative:

$$\det[\varphi_{\varkappa_i - i + j}]_{i,j=1}^K \geq 0.$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

In fact, this implies that all minors of the infinite matrix

$$\begin{bmatrix} \multicolumn{7}{c}{\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots} \\ \dots & \varphi_{-1} & \varphi_0 & \varphi_1 & \varphi_2 & \dots & \dots \\ \dots & \dots & \varphi_{-1} & \varphi_0 & \varphi_1 & \dots & \dots \\ \dots & \dots & \dots & \varphi_{-1} & \varphi_0 & \varphi_1 & \dots \\ \multicolumn{7}{c}{\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots} \end{bmatrix}$$

are nonnegative. Here we see the connection with the infinite totally nonnegative Toeplitz matrices. (Toeplitz matrices are those whose elements depend only on the diagonal: $\varphi_{i,j} = \varphi_{j-i}$.)

Vershik–Kerov sequences of signatures approximating boundary points

We will also discuss how the sequence of signatures $\nu(N)$ approximating a particular boundary point (= label of the extreme character) $(\alpha^{\pm}, \beta^{\pm}, \gamma^{\pm})$ looks like.

Let us draw the signature $\nu = (\nu_1, \dots, \nu_N)$ as a union of two Young diagrams, $\nu_1^+ \geq \dots \geq \nu_{\ell^+}^+$ and $\nu_1^- \geq \dots \geq \nu_{\ell^-}^-$, where the original signature looks as

$$\nu = (\nu_1^+, \dots, \nu_{\ell^+}^+, 0, \dots, 0, -\nu_{\ell^-}^-, \dots, -\nu_1^-).$$

See Figure 4.2 to understand how the signatures are drawn.

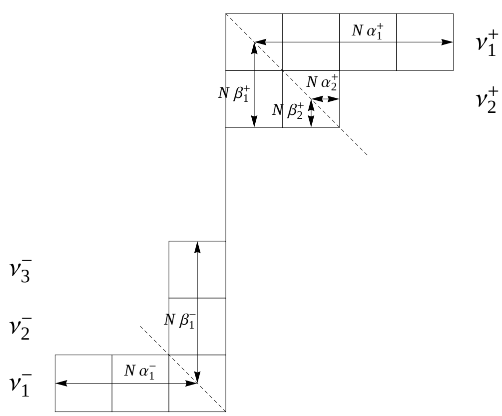


Figure 4.2: How Vershik–Kerov signatures grow.

As follows from the results of Vershik and Kerov, the sequences of signatures $\nu(N)$ for which there is convergence of the functions

$$\frac{s_{\nu(N)}(u_1, \dots, u_N)}{s_{\nu(N)}(1, \dots, 1)},$$

must behave in the following way (see Figure 4.2):

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

- The rows of $\nu^{\pm}$ must grow at most linearly in N, and

$$\frac{\nu_i^{\pm}(N)}{N} \to \alpha_i^{\pm}.$$

- The columns of $\nu^{\pm}$ must grow at most linearly in N, and

$$\frac{(\nu_i^{\pm}(N))'}{N} \to \beta_i^{\pm}.$$

  Here $(\nu_i^{\pm}(N))'$ denotes the ith row of a Young diagram.

- Moreover, the numbers of boxes in $\nu^{\pm}$ must also grow at most linearly:

$$\frac{|\nu^{\pm}(N)|}{N} \to \delta^{\pm}.$$

  We set $\gamma^{\pm} := \delta^{\pm} - \sum_i(\alpha_i^{\pm} + \beta_i^{\pm})$

We see why $\alpha_i^{\pm}$ and $\beta_i^{\pm}$ decrease, and also why $\beta_1^{+} + \beta_1^{-} \leq 1$ (because the number of parts in the signature must be N).

These sequences of signatures are called *Vershik–Kerov sequences*. For Vershik–Kerov sequences and only for them we have the convergence as in the Edrei–Voiculescu theorem:

$$\frac{s_{\nu(N)}(u_1, \dots, u_N)}{s_{\nu(N)}(1, \dots, 1)} \to \prod_{i=1}^{\infty} \Phi_{\alpha^{\pm},\beta^{\pm},\gamma^{\pm}}(u_i).$$

Gelfand–Tsetlin graph and coherent systems on it

Let us restrict the extreme character $\chi_{\alpha^{\pm},\beta^{\pm},\gamma^{\pm}}$ to $U(K)$ and expand it as a linear combination of extreme characters of $U(K)$:

$$\frac{s_{\varkappa}(u_1, \dots, u_K)}{s_{\varkappa}(1, \dots, 1)}.$$

We define the numbers $M_K(\varkappa)$ from the expansion:

$$\chi_{\alpha^{\pm},\beta^{\pm},\gamma^{\pm}} = \sum_{\varkappa \in \mathbb{GT}_K} M_K(\varkappa) \frac{s_{\varkappa}(u_1, \dots, u_K)}{s_{\varkappa}(1, \dots, 1)}.$$

Of course,

$$M_K(\varkappa) = s_{\varkappa}(1, \dots, 1) \det[\varphi_{\varkappa_i - i + j}]_{i,j=1}^{K}.$$

We easily see that the numbers M_K are nonnegative and sum, up to one, $\sum_{\varkappa \in \mathbb{GT}_K} M_K(\varkappa) = 1$. In fact, because these numbers for different K's come from one and the same character, they must be compatible in some way.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

They satisfy the following coherency condition (EX); use the branching rule for Schur polynomials which gives the interlacing condition:

$$M_K(\varkappa) = \sum_{\nu \in \mathbb{GT}_K \colon \varkappa \prec \nu} M_{K+1}(\nu) \frac{\mathrm{Dim}_K(\varkappa)}{\mathrm{Dim}_{K+1}(\nu)}.$$

This implies that the numbers $\{M_K\}_{K=1,2,3,\ldots}$ form a coherent system on a suitable branching graph which is called the *Gelfand–Tsetlin graph*. Its floors are

$$\varnothing,\ \mathbb{GT}_1,\ \mathbb{GT}_2,\ \mathbb{GT}_3,\ \ldots$$

(here $\varnothing$ is the empty signature on level zero for convenience). We connect two signatures $\mu \in \mathbb{GT}_{N-1}$ and $\lambda \in \mathbb{GT}_N$ iff they interlace, $\mu \prec \lambda$.

Here $\mathrm{Dim}_K \varkappa = s_\varkappa(1, \ldots, 1)$ is the number of paths in the Gelfand–Tsetlin graph from $\varnothing$ to $\varkappa \in \mathbb{GT}_K$.

In this sense the problem of describing all the extreme characters of the infinite-dimensional unitary group $U(\infty)$ is reduced to the problem of finding the boundary of the Gelfand–Tsetlin graph. This problem is the same as for the de Finetti's theorem, but of course the Gelfand–Tsetlin graph is more complicated.

4.1.12 Remark: Dimensions of irreducible representations of $U(N)$

For any finite-dimensional representation $G \to GL(V)$ of a finite or compact group G with character χ we have $\chi(e) = \dim V$. Therefore, if we evaluate the Schur polynomial at ones, we will have the dimension of the corresponding irreducible representation. So, let us denote

$$\mathrm{Dim}_N \lambda := s_\lambda(\underbrace{1, \ldots, 1}_{N}), \qquad \lambda \in \mathbb{GT}_N.$$

This also may be regarded as the number of Gelfand–Tsetlin schemes of depth N with fixed top row $\lambda \in \mathbb{GT}_N$, or, for $\lambda \in \mathbb{GT}_N^+$, as the number of semistandard Young tableaux of shape λ.

Here we aim to give two explicit formulas for $\mathrm{Dim}_N \lambda$.

Proposition 4.1.24. *We have for all $\lambda \in \mathbb{GT}_N$:*

$$\mathrm{Dim}_N \lambda = \prod_{1 \le i < j \le N} \frac{\lambda_i - \lambda_j + j - i}{j - i}. \tag{4.1.23}$$

Proof. Let us first compute explicitly the ratio of the two determinants in the definition of the Schur function for

$$s_\lambda(1, q, q^2, \ldots, q^{N-1}), \qquad 0 < q < 1.$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

We perform this because we cannot readily write both determinants for $1, \dots, 1$ (they both vanish). That is why we first deal with a geometric sequence, and then we will let $q \nearrow 1$.

We have

$$\begin{aligned} s_\lambda(1, q, q^2, \dots, q^{N-1}) &= \frac{\det[q^{(i-1)(\lambda_j+N-j)}]_{i,j=1}^N}{\det[q^{(i-1)(N-j)}]_{i,j=1}^N} \\ &= \frac{V(q^{\lambda_N - N}, \dots, q^{\lambda_1 - 1})}{V(q^{-1}, \dots, q^{-N})} \\ &= \prod_{1 \le i < j \le N} \frac{q^{\lambda_i - i} - q^{\lambda_j - j}}{q^{-i} - q^{-j}} \end{aligned}$$

where $V(\cdot)$ denotes the Vandermonde determinant.

Taking the $q \nearrow 1$ limit above, we arrive at the desired product formula. $\square$

Remark 4.1.25. There exists at least one more proof which is related to interpreting semistandard tableaux as families of nonintersecting paths. We will probably discuss that later (Gessel-Viennot determinants).

Let us state without proof one more formula for $\mathrm{Dim}_N \lambda$:

Proposition 4.1.26. *We have for all $\lambda \in \mathbb{GT}_N^+$:*

$$\mathrm{Dim}_N \lambda = \prod_{\square \in \lambda} \frac{N + c(\square)}{h(\square)}. \tag{4.1.24}$$

Here the product is taken over all boxes $\square$ of the diagram λ, $c(\square)$ — is the *content* of a box, and $h(\square)$ — is the *hook length* of that box. If $\square$ is in the j-th column and in the i-th row, than $c(\square)$ is, by definition, equal to $j - i$. The hook length $h(\square)$ is the number of boxes of the diagram λ which are strictly to the right or strictly below $\square$, plus one (i.e., the box itself belongs to the hook). An example of a hook is given on Figure 4.3.

Let us take our $\lambda = (4, 3, 1)$ and write the hook-length in its every box:

6	4	3	1
4	2	1	
1			

Proof of Proposition 4.1.26 See [Mac95, Ch. I] for a direct proof of equivalence of (4.1.23) and (4.1.24). $\square$

Let us make several remarks about the quantities $\mathrm{Dim}_N \lambda$.

Remark 4.1.27.

- From (4.1.24) we see that for a fixed $\lambda \in \mathbb{GT}_N^+$, the dimension $\mathrm{Dim}_N \lambda$ is a polynomial in N of degree $|\lambda|$ with leading term $\prod_{\square \in \lambda} \frac{1}{h(\square)}$.
- If $\ell(\lambda) > N$, then (4.1.24) implies that $\mathrm{Dim}_N \lambda = 0$. Indeed, for $\ell(\lambda) > N$ one can find a box $\square \in \lambda$ (in the first column) with content $(-N)$.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

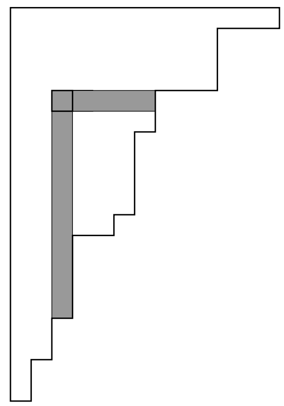


Figure 4.3: A hook in the Young diagram

Remark 4.1.28 (Recurrence for dimensions)**.** From the branching rule (4.1.12) we see that the quantities $\mathrm{Dim}_N \lambda$ satisfy the following recurrence relations:

$$\mathrm{Dim}_N \lambda = \sum_{\mu:\ \mu \prec \lambda} \mathrm{Dim}_{N-1}\ \mu \tag{4.1.25}$$

with initial conditions:

$$\mathrm{Dim}_1 \lambda = 1 \qquad \text{for all } \lambda \in \mathbb{GT}_1 = \mathbb{Z}_{\geq 0}. \tag{4.1.26}$$

The book [Mac95] contains an explicit proof of the fact that the right-hand side of (4.1.23) satisfies that recurrence.

Remark 4.1.29. It should be noted that

$$s_\lambda(u_1, \ldots, u_N) = (u_1 \cdot \ldots \cdot u_N)^{-k} s_{\lambda + \underbrace{(k,k,\ldots,k)}_{N}}(u_1, \ldots, u_N).$$

for any $k \in \mathbb{Z}$, where on the right we simply add k to all parts of the signature. This is transformed into property that $\mathrm{Dim}_N(\lambda + \underbrace{(k, k, \ldots, k)}_{N}) = \mathrm{Dim}_N \lambda$.

This implies that in many cases we may consider only nonnegative signatures. However, sometimes this is not so useful, and we must consider all signatures.

4.1.13 Proof of the Edrei-Voiculescu theorem

Outline of the proof: what we explicitly do

So we need to understand the behavior of the relative dimensions

$$\frac{\mathrm{Dim}_{K,N}(\varkappa, \nu)}{\mathrm{Dim}_N \nu}$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

for fixed $\varkappa \in \mathbb{GT}_K$, $\nu \in \mathbb{GT}_N$. We assume that N goes to infinity together with ν. If we understand the quantities $\frac{\mathrm{Dim}_{K,N}(\varkappa,\nu)}{\mathrm{Dim}_N \nu}$ explicitly, we will be able to establish the Edrei–Voiculescu theorem.

In the de Finetti's theorem, we showed that the relative dimensions behave as

$$\frac{\dim((c,d);(a,b))}{\dim(a,b)} = \left(\frac{a}{a+b}\right)^c \left(\frac{b}{a+b}\right)^d + o(1),$$

and thus we must have convergence $\left(\frac{a}{a+b}\right) \to p \in [0,1]$, and any coherent system must look like[4]

$$M_K((c,d)) = \dim(c,d) \int_0^1 p^c (1-p)^d \mu(dp),$$

where μ is the decomposing measure, and extreme coherent systems look like

$$\dim(c,d) p^c (1-p)^d$$

We expect a similar result in the context of Gelfand–Tsetlin graph. Namely, we will show that

$$\frac{\mathrm{Dim}_{K,N}(\varkappa,\nu)}{\mathrm{Dim}_N \nu} = \det[\varphi_{\varkappa_i - i + j}(\alpha^\pm, \beta^\pm, \gamma^\pm)]_{i,j=1}^K + o(1), \tag{4.1.27}$$

where $\nu = \nu(N)$ are signatures which converge to the triple of parameters $(\alpha^\pm, \beta^\pm, \gamma^\pm)$. This will imply that any coherent system has the form

$$M_K(\varkappa) = \mathrm{Dim}_K \varkappa \int\limits_{\{(\alpha^\pm,\beta^\pm,\gamma^\pm)\}} \det[\varphi_{\varkappa_i - i + j}(\alpha^\pm, \beta^\pm, \gamma^\pm)] \cdot \mu(d(\alpha^\pm, \beta^\pm, \gamma^\pm))$$

The integration is over the infinite-dimensional space of parameters, and μ is the decomposing measure on this infinite-dimensional space.

And this readily implies the statement of the Edrei–Voiculescu theorem above (Theorem 4.1.21). The extreme coherent systems have the form

$$M_K(\varkappa) = \mathrm{Dim}_K \varkappa \cdot \det[\varphi_{\varkappa_i - i + j}(\alpha^\pm, \beta^\pm, \gamma^\pm)]$$

See more:

About the other steps of the proof see the relevant research papers:

[Other material of Lectures 7 and also 8 may be found in the two research papers on the arXiv: `http://arxiv.org/abs/1109.1412`,

[4]Recall that in the Pascal triple case we simply represented every coherent system as a linear combination of extreme ones on finite level N, and then let $N \to \infty$.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

http://arxiv.org/abs/1208.3443. In the latter paper see section 2 for discussions about the boundary of the Gelfand–Tsetlin graph and its interpretations in terms of totally nonnegative sequences and in other aspects; the proof of the Edrei–Voiculescu theorem is outlined in sections 3–5 of the latter paper]

In these two references the asymptotics (4.1.27) is referred to as the **Uniform Approximation Theorem**.

4.2 Symmetric functions. Properties. Characters of symmetric groups

Here we will explain that for some symmetric polynomials (in particular, honest Schur polynomials — that is, those indexed by nonnegative signatures) it does not really matter at what number of variable we evaluate them. This leads to a notion of symmetric functions.

4.2.1 Symmetric polynomials. Stability. Algebra of symmetric functions

We've seen what happens with the Schur polynomials if one inserts 1 as one of the variables (e.g., see (4.1.12)). What happens if we insert 0 instead of 1?

Here we must restrict our attention to nonnegative signatures, so **in this section we will speak about honest symmetric polynomials, and not Laurent polynomials**. To indicate this agreement, we will use $x_1, x_2, \dots$ to denote the variables in symmetric honest polynomials (instead of u_i, $|u_i| = 1$, which were used as arguments in symmetric Laurent polynomials).

Proposition 4.2.1. *Let μ be a nonnegative signature, then*

$$s_\mu(x_1, \dots, x_N)|_{x_N=0} = \begin{cases} s_\mu(x_1, \dots, x_{N-1}), & \text{if } \ell(\mu) \le N-1; \\ 0, & \text{else.} \end{cases} \tag{4.2.1}$$

Here by $\ell(\mu)$ we denote the number of strictly positive parts in μ.

Proof. We have

$$s_\mu(x_1, \dots, x_N)|_{x_N=0} = \frac{\det[x_j^{\mu_i+N-i}]|_{x_N=0}}{\prod_{i<j}(x_i - x_j)|_{x_N=0}}. \tag{4.2.2}$$

The last column of the matrix in the denominator has the form

$$(x_N^{\mu_1+N-1}, \dots, x_N^{\mu_{N-1}+1}, x_N^{\mu_N})^T.$$

If $\mu_N > 0$ (i.e., $\ell(\mu) > N-1$), then this is a zero column, and $s_\mu(x_1, \dots, x_N)|_{x_N=0} = 0$. If $\mu_N = 0$, then this column has the form $(0, \dots, 0, 1)^T$. It remains to note that

$$\det\left[x_j^{\mu_i+N-i}\right]_{i,j=1,\dots,N-1} = x_1 \dots x_{N-1} \cdot \det\left[x_j^{\mu_i+(N-1)-i}\right]_{i,j=1,\dots,N-1} \tag{4.2.3}$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

and that $V_N(x_1,\dots,x_N)|_{x_N=0} = x_1 \dots x_{N-1} \cdot V_{N-1}(x_1,\dots,x_{N-1})$. This concludes the proof. $\square$

This property means that we can treat each Schur polynomial s_μ as a symmetric polynomial in any number of variables which is $\geq \ell(\mu)$. For example:

$$\begin{aligned}
s_{(2,1)}(x_1,x_2) &= x_2x_1^2 + x_2^2x_1 \\
s_{(2,1)}(x_1,x_2,x_3) &= x_2x_1^2 + x_3x_1^2 + x_2^2x_1 + x_3^2x_1 + 2x_2x_3x_1 + x_2x_3^2 + x_2^2x_3 \\
s_{(2,1)}(x_1,x_2,x_3,x_4) &= x_2x_1^2 + x_3x_1^2 + x_4x_1^2 + x_2^2x_1 + x_3^2x_1 + x_4^2x_1 + 2x_2x_3x_1 \\
&\quad + 2x_2x_4x_1 + 2x_3x_4x_1 + x_2x_3^2 + x_2x_4^2 + x_3x_4^2 + x_2^2x_3 \\
&\quad + x_2^2x_4 + x_3^2x_4 + 2x_2x_3x_4,
\end{aligned}$$

etc. Note a change happening when we pass from 2 variables to 3 variables.

In this way, we may regard every Schur polynomial as a "polynomial" in infinitely many variables $x_1, x_2, \dots$. One possible way is to express it as a linear combination of symmetrized monomials (for example, above $s_{(2,1)} = m_{(2,1)} + 2m_{(1,1,1)}$); and each symmetrized monomial can be readily defined for infinitely many variables:

$$m_\mu(x_1,x_2,x_3,\dots) = \sum_{\text{all distinct summands}} x_{i_1}^{\mu_1} x_{i_2}^{\mu_2} \dots x_{i_{\ell(\mu)}}^{\mu_{\ell(\mu)}}.$$

Here μ runs over all nonnegative signatures (= partitions = Young diagrams).

Let us give a more formal definition.

Definition 4.2.2 (Algebra of symmetric functions)**.** Let $\mathit{Sym}(N)$ be the algebra of (honest, not Laurent) polynomials in N variables, and $\mathit{Sym}_n(N)$ be its subalgebra consisting of homogeneous polynomials of degree exactly n:

$$\mathit{Sym}(N) = \bigoplus_{n=0}^{\infty} \mathit{Sym}_n(N).$$

We have projections

$$\mathit{Sym}(N) \leftarrow \mathit{Sym}(N+1)$$

which consist of setting the last variable x_{N+1} to 0. Also, these projections are compatible with the grading by $\mathit{Sym}_n(N)$.

Let Sym_n be the inverse (projective) limit of the chain of projections

$$\mathit{Sym}_n(1) \leftarrow \mathit{Sym}_n(2) \leftarrow \mathit{Sym}_n(3) \leftarrow \dots.$$

That is, elements of Sym_n are sequences (f_1, f_2, f_3) of degree n homogeneous symmetric polynomials with growing number of variables which are compatible with projections:

$$f_{N+1}(x_1,x_2,\dots,x_N;x_{N+1}=0) = f_N(x_1,x_2,\dots,x_N).$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Each Sym_n is the space of homogeneous symmetric functions of degree n.

Then we define the algebra of symmetric functions by

$$Sym := \bigoplus_{n=0}^{\infty} Sym_n(N).$$

This is indeed an algebra, and it is also graded. Elements of *Sym* — symmetric functions — are sequences of symmetric polynomials $(f_1, f_2, f_3, \ldots)$ which satisfy

(1) $f_N \in Sym(N)$

(2) $f_{N+1}(x_1, x_2, \ldots, x_N; x_{N+1} = 0) = f_N(x_1, x_2, \ldots, x_N)$

(3) $\sup_N \deg f_N < \infty$

In other words, *Sym* is the projective limit of *Sym*(N)'s in the category of graded algebras. The word "graded" is essential, because we want summands (monomials) in symmetric functions to have bounded degree. For example, we do not allow $\prod_{i=1}^{\infty}(1 + x_i)$ to be a symmetric function: it violates condition (3) above.

In this way we see that the Schur polynomials $s_\lambda(x_1, \ldots, x_N)$ indexed by nonnegative signatures (we use the notations $\mathbb{GT}^+ = \mathbb{Y}$ for nonnegative signatures; the letter $\mathbb{Y}$ just means the set of all Young diagrams) define symmetric functions "as $N \to \infty$". We call these symmetric functions the *Schur functions* $s_\lambda \in Sym$.

Both the symmetrized monomials $\{m_\mu\}_{\mu \in \mathbb{Y}}$ and the Schur functions $\{s_\lambda\}_{\lambda \in \mathbb{Y}}$ form linear bases of the algebra *Sym*.

4.2.2 Multiplicative bases in *Sym*: e- and h-functions

Here let us discuss some structure of the algebra *Sym*.

Proposition 4.2.3. *$Sym = \mathbb{R}[e_1, e_2, \ldots]$, where e_i's are the elementary symmetric functions (particular cases of the Schur functions). This proposition states that $\{e_i\}_{i \geq 1}$ is a multiplicative basis of Sym, i.e., every symmetric function can be written as a (usual, finite) polynomial in elementary symmetric functions. Equivalently, the symmetric functions*

$$e_\lambda := e_{\lambda_1} e_{\lambda_2} \ldots e_{\lambda_{\ell(\lambda)}},$$

where $\lambda \in \mathbb{Y}$ runs over all Young diagrams, form a linear basis of Sym.

Proof. We use the fact that the functions e_λ may be expresses through the basis of monomial symmetric functions (= symmetrized monomials) $\{m_\mu\}$ by means of a triangular transform. Let $<$ denote the usual lexicographic order on the set $\mathbb{Y}_n$ of all Young diagrams with n boxes. Then we will show that

$$e_{\lambda'} = m_\lambda + \sum_{\mu \in \mathbb{Y}_{|\lambda|},\, \mu < \lambda} a_{\lambda\mu} m_\mu, \tag{4.2.4}$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

where coefficients $a_{\lambda\mu} \in \mathbb{Z}_{\geq 0}$, and λ' denotes the transposed Young diagram (i.e., which is obtained by replacing rows by columns and vice versa; for example, $(4,3,1)' = (3,2,2,1)$). An example of the above identity would be

$$e_1 e_1 = m_{(2)} + 2m_{(1,1)}$$

(which can be readily checked).

To establish (4.2.4), note that the coefficients of the decomposition of a symmetric function in the basis $\{m_\mu\}$ are in fact coefficients of this function in front of suitable monomials.

Write

$$\begin{aligned} e_{\lambda'} &= e_1^{\lambda_1-\lambda_2} e_2^{\lambda_2-\lambda_3} e_3^{\lambda_3-\lambda_4} \dots \\ &= \left(\sum_i x_i\right)^{\lambda_1-\lambda_2} \left(\sum_{i<j} x_i x_j\right)^{\lambda_2-\lambda_3} \left(\sum_{i<j<k} x_i x_j x_k\right)^{\lambda_3-\lambda_4} \dots \\ &= x_1^{\lambda_1-\lambda_2} x_3 2^{\lambda_2-\lambda_3} x_3^{\lambda_3-\lambda_4} \dots + \text{lower terms in lexicographic order} \\ &= x^\lambda + \text{lower terms in lexicographic order} \end{aligned}$$

(the first equality is easy if we look at the transposed diagram, and in the third equality we just extract the highest lexicographic power of x: i.e., we choose the maximal possible power of x_1, then the maximal power of x_2 in what remains, etc.). This concludes the proof. □

Recall the complete homogeneous polynomials h_m; of course, they also define symmetric functions (called complete homogeneous symmetric functions). Now let us consider generating functions of the e- and h-functions.

Proposition 4.2.4. *We have*

$$\sum_{k=0}^{\infty} e_k(x_1, x_2, \dots) t^k = \prod_{i=1}^{\infty} (1 + x_i t),$$
$$\sum_{k=0}^{\infty} h_k(x_1, x_2, \dots) t^k = \prod_{i=1}^{\infty} \frac{1}{1 - x_i t}$$

These are identities of formal power series in t with coefficients in Sym.

Proof. This is an easy combinatorial exercise; for the e-functions it is just the opening of the parentheses plus (4.1.16). For the h-functions one should first write each factor $1/(1 - x_i t)$ as $1 + x_i t + (x_i t)^2 + \dots$, and then also open the parentheses using (4.1.17). □

We have

$$E(t)H(-t) = 1,$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

or, looking at corresponding powers of t to the left and to the right, we have

$$\sum_{r=0}^{n}(-1)^r e_r h_{n-r} = 0, \qquad n \geq 1. \tag{4.2.5}$$

Proposition 4.2.5. *The complete homogeneous symmetric functions also form a multiplicative basis in Sym:*

$$Sym = \mathbb{R}[h_1, h_2, \ldots].$$

Proof. Define a map $\omega\colon Sym \to Sym$ by setting $\omega(e_i) = h_i$, $i = 0, 1, \ldots$. Since the relations (4.2.5) are symmetric in e- and h-functions, we see that ω is an involution, $\omega^2 = id$. This means that the h-functions cannot be algebraically dependent, so they form a multiplicative basis. □

4.2.3 Power sums

Consider the special symmetrized monomials, namely, the power sums

$$p_k(x_1, x_2, \ldots) := m_{(k)}(x_1, x_2, \ldots) = \sum_{i=1}^{\infty} x_i^k, \qquad k = 1, 2, \ldots.$$

Proposition 4.2.6. *The power sums also form a multiplicative basis for the algebra of symmetric functions:*

$$Sym = \mathbb{R}[p_1, p_2, p_3, \ldots].$$

Proof. We have

$$\begin{aligned} P(t) :&= \sum_{r=1}^{\infty} p_r t^r / r \\ &= \sum_{r=1}^{\infty}\sum_{i=1}^{\infty} x_i^r t^r / r \\ &= \sum_{i=1}^{\infty}(-\log(1 - x_i t)) \\ &= -\log \prod_{i=1}^{\infty}(1 - x_i t) \\ &= \log H(t), \end{aligned}$$

where $H(t)$ is the generating function of the h-functions.

We also note that this implies the following:

$$P'(t) = H'(t)/H(t), \qquad P'(-t) = E'(t)/E(t).$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Comparing powers of t, we see that

$$nh_n = \sum_{r=1}^{n} p_r h_{n-r}, \qquad ne_n = \sum_{r=1}^{n} (-1)^r p_r e_{n-r}.$$

From these identities we see that e_i's and h_i's can be written as polynomials in p_i's, and vice versa. This concludes the proof. □

4.2.4 Cauchy identity

Let us prove an important identity in the theory of symmetric functions. It has many important consequences.

Proposition 4.2.7. *We have*

$$\sum_{\lambda\in\mathbb{Y}} s_\lambda(x_1, x_2, \ldots) s_\lambda(y_1, y_2, \ldots) = \prod_{i,j=1}^{\infty} \frac{1}{1 - x_i y_j}. \tag{4.2.6}$$

The sum is taken over all Young diagrams, and this is a formal identity in $Sym^x \otimes Sym_y$.

Proof. This is based on the Cauchy determinant (EX)[5]:

$$\det\left(\frac{1}{1 - x_i y_j}\right)_{i,j=1}^{N} = \frac{\prod_{i<j}(x_i - x_j)(y_i - y_j)}{\prod_{i,j}(1 - x_i y_j)}.$$

Using this determinant, we will establish (4.2.6) for a finite number of variables, this is enough (see the definition of a symmetric function). We will show that

$$\sum_{\lambda\colon\ \ell(\lambda)\le N} a_{\lambda+\delta}(x_1, \ldots, x_N) a_{\lambda+\delta}(y_1, \ldots, y_N) = \det\left(\frac{1}{1 - x_i y_j}\right)_{i,j=1}^{N}. \tag{4.2.7}$$

Here $a_{\lambda+\delta}$ is the alternant used to define Schur functions. After (4.2.7) is obtained, the rest is easy.

We have by the definition of the determinant (writing every $1/(1 - x_i y_j)$ as a power series):

$$\det\left(\frac{1}{1 - x_i y_j}\right)_{i,j=1}^{N} = \sum_{k_1>\ldots>k_N\ge 0} \sum_{\sigma\in\mathfrak{S}(N)} \sum_{\tau\in\mathfrak{S}(N)} \operatorname{sgn}(\sigma)(x_{\sigma(1)} y_{\tau(1)})^{k_1} \ldots (x_{\sigma(N)} y_{\tau(N)})^{k_N}.$$

This can be readily transformed to

$$\sum_{\lambda\colon\ \ell(\lambda)\le N} a_{\lambda+\delta}(x_1, \ldots, x_N) a_{\lambda+\delta}(y_1, \ldots, y_N).$$

This concludes the proof. □

[5] After multiplying by $\prod_{i,j}(1 - x_i y_j)$, we get anti-symmetric polynomials in both x's and y's in both sides. They are divisible by the product of Vandermondes. Then one needs to compare degrees of the polynomials, and show that the leading coefficient is 1.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

4.2.5 Jacobi-Trudi identity

We know that the h-functions form a multiplicative basis in the algebra of symmetric functions; that is, every symmetric function can be written as a polynomial in the h-functions. For the Schur symmetric functions this polynomial can be presented explicitly in a determinantal form:

Proposition 4.2.8 (Jacobi-Trudi identity)**.** *For every $\lambda \in \mathbb{Y}$ one has*

$$s_\lambda = \det[h_{\lambda_i - i + j}]_{i,j=1}^N,$$

where N is any number $\geq \ell(\lambda)$. By agreement, $h_{-1} = h_{-2} = \ldots = 0$.

Proof. This is based on the general fact that if

$$f(u) = \sum_{m \geq 0} f_m u^m,$$

then

$$f(x_1) \ldots f(x_N) = \sum_{\lambda:\, \ell(\lambda) \leq N} s_\lambda(x_1, \ldots, x_N) \det[f_{\lambda_i - i + j}]_{i,j=1}^N.$$

(the left-hand side is a symmetric power series, and on the right we see an expression of this power series as an infinite linear combination of the Schur polynomials). This identity may be obtained by multiplication by a_δ and matching the coefficients. Then one should apply this identity with $f(u) = H(u)$, the generating function for the h-functions. □

4.2.6 Pieri formula. Standard Young tableaux

Let us write the operator of multiplication by $p_1 = x_1 + x_2 + \ldots$ in the basis of the Schur functions.

Proposition 4.2.9. *For any $\mu \in \mathbb{Y}$, we have*

$$p_1 s_\mu = \sum_\lambda s_\lambda,$$

where the sum is over all Young diagrams λ such that λ is obtained from μ by adding one box.

Proof. We have

$$\begin{aligned} p_1(x_1, \ldots, x_N) s_\mu(x_1, \ldots, x_N) &= a_\delta(x)^{-1} \sum_j x_j \sum_{\sigma \in S_N} \operatorname{sgn}(\sigma) x_{\sigma(1)}^{\mu_1 + N - 1} \cdots x_{\sigma(N)}^{\mu_N} \\ &= a_\delta(x)^{-1} \sum_{k=1}^N a_{\mu + \delta + \bar{e}_k}, \end{aligned}$$

where $\bar{e}_k$ is the kth standard basis vector. A basic argument shows that each $a_{\mu+\delta+\bar{e}_k}$ is either 0 or can be rearranged to $a_{\lambda+\delta}$ where $\lambda = \mu + \square$ is obtained from μ by adding a box. □

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

In fact, there is a generalization of the Pieri rule which can be obtained in the same way. A **rim hook** is a connected set of boxes on the border of a Young diagram that does not contain a 2×2 square.

Proposition 4.2.10. *For any $\mu \in \mathbb{Y}$ and $r \geq 1$,*

$$p_r s_\mu = \sum_\lambda (-1)^{\mathrm{ht}(\lambda \setminus \mu)} s_\lambda$$

where the sum is over all partitions λ such that $\lambda \setminus \mu$ is a rim hook with r boxes, and ht *is the number of rows of the rim hook minus* 1.

Corollary 4.2.11. *We have for $n = 1, 2, \ldots$:*

$$p_1^n = \sum_{\lambda \in \mathbb{Y}\colon |\lambda| = n} \dim \lambda \cdot s_\lambda,$$

where $\dim \lambda$ *is the number of ways to build λ by adding one box at a time.*

The number $\dim \lambda$ also has a combinatorial interpretation, as the number of *standard Young tableaux* of shape λ. A standard Young tableau is a filling of boxed of λ, $|\lambda| = n$, with numbers $1, 2, \ldots, n$ (every number must occur only once), such that they increase along rows and down columns. Example of a standard Young tableau:

1	2	4	6
3	5	8	
7			

(4.2.8)

Remark 4.2.12. The dimensions $\dim \lambda$ and $\mathrm{Dim}_N \lambda$ are very different. There exist some connections, however (based on the Schur-Weyl duality). For $\dim \lambda$ there is a hook formula:

$$\dim \lambda = n! \prod_{\square \in \lambda} \frac{1}{h(\square)}.$$

One may compare it to formula (4.1.24) for $\mathrm{Dim}_N \lambda$ and see that these quantities are different.

Also,

$$\dim \lambda = \frac{n! \prod_{1 \leq i < j \leq N} (\lambda_i - \lambda_j + j - i)}{\prod_{i=1}^N (\lambda_i + N - i)!},$$

where N is any number $\geq \ell(\lambda)$. This is an analogue of (4.1.23). We see the quantity $\mathrm{Dim}_N \lambda$ is simpler.

4.2.7 Frobenius characteristic map. Characters of the symmetric grous $\mathfrak{S}(n)$

Consider a map $\psi\colon \mathfrak{S}(n) \to Sym_n$ (homogeneous symmetric functions of degree n) defined by $\psi(\sigma) = p_{\rho(\sigma)} := p_{\rho(\sigma)_1} \ldots p_{\rho(\sigma)_\ell}$ where $\rho(\sigma) = (\rho(\sigma)_1, \ldots, \rho(\sigma)_\ell)$ is

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

the partition encoding the cycle lengths of σ. For any $f \in \mathbb{C}[\mathfrak{S}(n)]$, define the **characteristic map**

$$\mathrm{ch}(f) = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}(n)} f(\sigma)\psi(\sigma).$$

Theorem 4.2.13. $\mathrm{ch}(\chi^\lambda) = s_\lambda$ *where* χ^λ *is the character of the corresponding irreducible representation of* $\mathfrak{S}(n)$*.*

Let c_ρ be the conjugacy class of type ρ. In particular, this says that

$$s_\lambda = \mathrm{ch}(\chi^\lambda) = \frac{1}{n!} \sum_{|\rho|=n} |c_\rho| \chi^\lambda_\rho p_\rho.$$

Here χ^λ_ρ is the value of the irreducible character indexed by λ on the conjugacy class indexed by ρ.

Using the orthogonality relations for irreducible characters of $\mathfrak{S}(n)$, we get the Frobenius formula

$$p_\rho = \sum_{|\lambda|=n} \chi^\lambda_\rho s_\lambda. \tag{4.2.9}$$

(Surprisingly enough, this is currently the most efficient way to calculate the character values for the symmetric group.)

We define a product on characters of all symmetric groups via induction: $\chi^{\pi_1} \circ \chi^{\pi_2} = \chi^\pi$ where $\pi = Ind_{S_{n_1} \times S_{n_2}}^{S_{n_1+n_2}} \pi_1 \boxtimes \pi_2$. Under the characteristic map, this product goes to usual product of symmetric functions.

4.2.8 Branching of irreducible characters of symmetric groups

Proposition 4.2.14. *Let* $\lambda \in \mathbb{Y}$*,* $|\lambda| = n$*. Consider the irreducible character* χ^λ *of* $\mathfrak{S}(n)$ *corresponding to the Young diagram* λ*. Restrict it to* $\mathfrak{S}(n-1) \subset \mathfrak{S}(n)$*, and write it as a linear combination of irreducible characters of* $\mathfrak{S}(n-1)$*. We have*

$$\chi^\lambda|_{\mathfrak{S}(n-1)} = \sum_{\mu:\ \mu = \lambda - \square} \chi^\mu.$$

That is, the irreducible representation of $\mathfrak{S}(n)$ *corresponding to* λ *restricted to* $\mathfrak{S}(n-1)$ *decomposes into a multiplicity-free direct sum over irreducible representations indexed by* $\mu \in \mathbb{Y}$*,* $|\mu| = n-1$*, which are obtained from* λ *by deleting one box.*

Proof. We use (4.2.9). Write for $\mu, \rho \in \mathbb{Y}$, $|\mu| = |\rho| = n-1$:

$$p_\rho = \sum_{|\mu|=n-1} s_\mu \chi^\mu_\rho,$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

and multiply this by p_1. Then

$$p_\rho p_1 = \sum_{|\mu|=n-1} \sum_{\lambda:\ \lambda=\mu+\square} s_\lambda \chi^\mu_\rho,$$

and on the other hand,

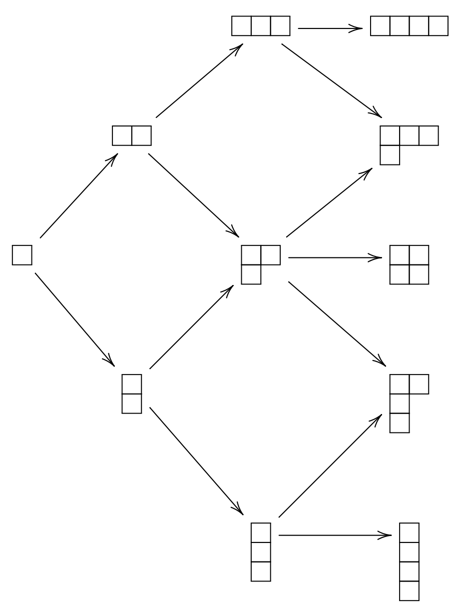
$$p_\rho p_1 = p_\sigma = \sum_{\lambda:\ |\lambda|=n} s_\lambda \chi^\lambda_\sigma$$

where $\sigma = \rho \cup 1$. This implies the desired branching rule if we compare coefficients of s_λ in two previous formulas. $\square$

In some sense, this branching of characters of symmetric groups is encoded by the operator of multiplication by p_1 in the algebra of symmetric functions *Sym*. We will present a branching-graph interpretation, and this will become clearer.

4.2.9 Young graph

Young graph's vertices are Young diagrams corresponding to partitions $\lambda = (\lambda_1 \geq \ldots \geq \lambda_\ell > 0)$, $\lambda_i \in \mathbb{Z}$. We connect two diagrams by an edge iff one of them can be obtained from the other by adding one box. This corresponds to the branching rule of characters of symmetric groups (Proposition 4.2.14). The first several floors of the Young graph look as follows:



Since we have a branching graph, we can readily define coherent systems, harmonic functions, etc. Note that $\dim \lambda$ defined above as the number of stan-

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

dard Young tableaux, is the same as the number of paths in the Young graph from the initial vertex to the diagram λ.

Proposition 4.2.15. *Extreme coherent systems on the Young graph are in bijection with extreme characters of the infinite symmetric group $\mathfrak{S}(\infty)$.*

Proof. Coherent systems correspond to characters χ of $\mathfrak{S}(\infty)$ as

$$\chi|_{\mathfrak{S}(n)} = \sum_{\lambda\colon |\lambda|=n} M_n(\lambda)\frac{\chi^\lambda(\cdot)}{\chi^\lambda(e)} = \sum_{\lambda\colon |\lambda|=n} M_n(\lambda)\frac{\chi^\lambda(\cdot)}{\dim\lambda}.$$

Here M_n is a (nonnegative) probability measure on $\mathbb{Y}_n := \{\lambda\colon |\lambda| = n\}$ for all n.

From the branching rule of characters of symmetric groups, restricting the above identity to $\mathfrak{S}(n-1)$, we see that the measures $\{M_n\}$ satisfy

$$M_{n-1}(\mu) = \sum_{\lambda\colon \lambda=\mu+\square} \frac{\dim\mu}{\dim\lambda} M_n(\lambda),$$

which is precisely the coherence condition for the measures $\{M_n\}$.

It is also possible to reconstruct the character back using the inductive structure of the infinite symmetric group. □

4.2.10 Classification theorem for extreme characters of $\mathfrak{S}(\infty)$

Let us formulate the answer — the Edrei-Thoma theorem classifying extreme characters of the infinite symmetric group $\mathfrak{S}(\infty)$.

Theorem 4.2.16. *Extreme characters of $\mathfrak{S}(\infty)$ depend on countably many continuous parameters α_i, β_i such that*

$$\alpha = (\alpha_1 \ge \alpha_2 \ge \cdots \ge 0);$$
$$\beta = (\beta_1 \ge \beta_2 \ge \cdots \ge 0);$$

and

$$\sum_i (\alpha_i + \beta_i) \le 1.$$

The character has the form

$$\chi_\rho^{\alpha,\beta} = \prod_{k=2}^{\infty} \left(\sum_{i=1}^{\infty} \alpha_i^k + (-1)^{k-1} \sum_{i=1}^{\infty} \beta_i^k \right)^{m_k},$$

where m_k is the number of parts of ρ which are equal to k (i.e., the number of k-cycles in ρ).

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

A useful generating function is

$$\exp\left\{\sum_{r\geq 1}\chi_{(r)}^{\alpha,\beta}\frac{t^r}{r}\right\} = e^{\gamma t}\prod_{i=1}^{\infty}\frac{1+\beta_i t}{1-\alpha_i t},$$

where $\gamma = 1-\sum(\alpha_i+\beta_i)\geq 0$.

4.2.11 Specizations of the algebra of symmetric functions. 'Ring theorem'

When we will discuss the proof of this Theorem, we will also mention the so-called "ring theorem" connecting characters of $\mathfrak{S}(\infty)$ and algebra homomorphisms $F\colon \mathit{Sym}\to\mathbb{R}$. This is related to the branching-graph interpretation of characters of $\mathfrak{S}(\infty)$.

In §2.4 we have proven a theorem that extreme coherent systems on the Pascal triangle are the same as algebra homomorphisms $\mathbb{C}[x,y]\to\mathbb{R}$ such that $F(x+y)=1$ and $F(x^a y^b)\geq 0$ for all vertices $(a,b)\in\mathbb{PT}$. In the same manner (using the same argument), one can prove the following:

Theorem 4.2.17 (Vershik–Kerov "ring theorem")**.** *If a nonnegative normalized harmonic function φ on $\mathbb{Y}$ is extreme, then the linear map $\hat\varphi\colon \mathit{Sym}\to\mathbb{R}$ defined by $\hat\varphi(s_\lambda)=\varphi(\lambda)$ is multiplicative, i.e., $\hat\varphi(fg)=\hat\varphi(f)\hat\varphi(g)$. Furthermore, this correspondence is a bijection between extreme nonnegative normalized harmonic functions and algebra homomorphisms $F\colon \mathit{Sym}\to\mathbb{R}$ such that $F(s_1)=1$ and $F(s_\lambda)\geq 0$ for all $\lambda\in\mathbb{Y}$.*

We know that nonnegative normalized extreme harmonic functions on $\mathbb{Y}$ bijectively correspond to extreme characters of $\mathfrak{S}(\infty)$.

Proof. First we show that φ extreme implies that $\hat\varphi$ is multiplicative. For any $\lambda\in\mathbb{Y}$ such that $\varphi(\lambda)\neq 0$, define $\varphi_\lambda\colon \mathbb{Y}\to\mathbb{R}$ by $\varphi_\lambda(\mu)=\hat\varphi(s_\lambda s_\mu)/\hat\varphi(s_\lambda)$. We claim that φ_λ is harmonic. We have

$$\sum_{\nu:\nu\searrow\mu}\varphi_\lambda(\nu)=\sum_{\nu\searrow\mu}\frac{\hat\varphi(s_\lambda s_\nu)}{\hat\varphi(s_\lambda)}=\frac{\hat\varphi(s_\lambda\sum_{\nu\searrow\mu}s_\nu)}{\hat\varphi(s_\lambda)}=\frac{\hat\varphi(s_\lambda p_1 s_\mu)}{\hat\varphi(s_\lambda)}. \tag{4.2.10}$$

Since φ is harmonic, we have

$$\hat\varphi(s_\kappa)=\varphi(\kappa)=\sum_{\rho\searrow\kappa}\varphi(\rho)=\hat\varphi(\sum_{\rho\searrow\kappa}s_\rho)=\hat\varphi(p_1 s_\kappa).$$

This implies that $\hat\varphi(p_1 f)=\hat\varphi(f)$ for any $f\in \mathit{Sym}$. Going back to (4.2.10), we get that the result is equal to $\hat\varphi(s_\mu s_\lambda)/\hat\varphi(s_\lambda)=\varphi_\lambda(\mu)$. Also, from Schur-positivity of products of Schur functions, φ_λ is nonnegative. This proves the claim.

Similarly, for any $f=\sum_\mu c_\mu s_\mu$ with $c_\mu\geq 0$ and $\hat\varphi(f)\neq 0$, we define $\varphi_f(\mu)=\hat\varphi(fs_\mu)/\hat\varphi(f)$. The argument above shows that φ_f is harmonic, normalized, and nonnegative.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Now we claim that that $\hat{\varphi}(s_\mu s_\lambda) = \hat{\varphi}(s_\mu)\hat{\varphi}(s_\lambda)$. If $\varphi(\lambda)\hat{\varphi}(s_\lambda) = 0$, then $\varphi(\rho) = 0$ for any ρ that contains λ from the harmonicity condition. Hence the claim holds in this case.

So assume that $\hat{\varphi}(s_\lambda) \neq 0$ and set $n = |\lambda|$. Take $f_1 = s_\lambda/2$, $f_2 = p_1^n - f_1 = \sum_{|\mu|=n}(\dim\mu)s_\mu - s_\mu/2$. So f_2 is Schur-positive. Then we have

$$\varphi = \hat{\varphi}(f_1)\varphi_{f_1} + \hat{\varphi}(f_2)\varphi_{f_2}.$$

To see this, pick μ. Then the left-hand side is

$$\varphi(s_\mu) = \hat{\varphi}(f_1)\frac{\hat{\varphi}(f_1 s_\mu)}{\hat{\varphi}(f_1)} + \hat{\varphi}(f_2)\frac{\hat{\varphi}(f_2 s_\mu)}{\hat{\varphi}(f_2)} = \hat{\varphi}((f_1+f_2)s_\mu) = \hat{\varphi}(p_1^n s_\mu) = \hat{\varphi}(s_\mu).$$

Since φ was assumed to be extreme, we must have either $\varphi = \varphi_{f_1}$ or $\varphi = \varphi_{f_2}$. In the first case, we get the identity $\hat{\varphi}(s_\lambda s_\nu/2)/\hat{\varphi}(s_\lambda/2) = \hat{\varphi}(s_\nu)$ which proves multiplicativity.

Now we will show the reverse direction. Let $H_{\mathbb{Y}}$ be the set of all nonnegative normalized harmonic functions on $\mathbb{Y}$, which lives in the space of all functions on $\mathbb{Y}$, with the topology of pointwise convergence. By Choquet's theorem (below), for any φ, we have $\varphi(y) = \int_{E(H_{\mathbb{Y}})}\psi(\lambda)P(\mathrm{d}\psi)$ for some probability measure P. This is equivalent to saying $\hat{\varphi}(s_\lambda) = \int_{E(H_{\mathbb{Y}})}\hat{\psi}(s_\lambda)P(\mathrm{d}\psi)$. By the first part of the proof, each $\hat{\psi}$ is multiplicative. Now assume that φ is multiplicative. Then we have

$$\left(\int_{E(H_{\mathbb{Y}})}\hat{\psi}(f)P(\mathrm{d}\psi)\right)^2 = \hat{\varphi}(f)^2 = \hat{\varphi}(f^2) = \int_{E(H_{\mathbb{Y}})}\hat{\psi}(f^2)P(\mathrm{d}\psi) = \int_{E(H_{\mathbb{Y}})}(\hat{\psi}(f))^2P(\mathrm{d}\psi).$$

If we think of $\hat{\psi}(f)$ as a random variable, this equality says $\mathrm{Var}(\hat{\psi}(f)) = 0$. This implies that $\hat{\psi}(f)$ is a constant with probability 1. Now take $f = s_\lambda$ and take a set $A_\lambda \subset E(H_{\mathbb{Y}})$ such that $\psi \in A_\lambda$ if $\psi(s_\lambda) = \varphi(s_\lambda)$. We know that $P(A_\lambda) = 1$. Then $\bigcap_\lambda A_\lambda \neq \emptyset$. So for any $\psi \in \bigcap_\lambda A_\lambda$, we will have $\psi = \varphi$ which implies that φ is extreme. □

Theorem 4.2.18 (Choquet [Phe66])**.** *Assume that X is a metrizable compact convex set in a locally convex topological linear space V and pick $x_0 \in X$. Then there exists a (Borel) probability measure P on X supported by the set $E(X)$ of extreme points of X which represents x_0, i.e., for any continuous linear functional $f \in V^*$, we have $f(x_0) = \int_{E(X)} f(x)P(\mathrm{d}x)$.*

Remark 4.2.19. Vague analogy: positive-definite normalized continuous functions on $\mathbb{R}$ are $\psi\colon \mathbb{R} \to \mathbb{C}$ such that $\psi(0) = 1$ and $[\psi(x_i - y_j)] \geq 0$ when $x_1 > x_2 > \cdots$ and $y_1 > y_2 > \cdots$.

Theorem 4.2.20 (Bochner)**.** *Such ψ are in bijection with probability measures μ on $\mathbb{R}$: $\psi(x) = \int_{\mathbb{R}}\exp(ixp)\mu(\mathrm{d}p)$.*

Note that $\exp(ixp)\colon \mathbb{R} \to \mathbb{C}$ are the characters of $\mathbb{R}$.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Edrei-Thoma's theorem (Theorem 4.2.16) asserts that all algebra homomorphisms (= specializations) $F\colon Sym \to \mathbb{R}$ such that $F(p_1) = 1$ and $F(s_\lambda) \geq 0$ for all λ (so-called Schur-positive specializations) have the following form on the linear basis $\{p_\rho\}$:

$$F(p_\rho) = \chi(\rho) = \prod_{k=2}^{\infty} \left(\sum_{i=1}^{\infty} \alpha_i^k + (-1)^{k-1} \sum_{i=1}^{\infty} \beta_i^k \right)^{m_k},$$

where m_k is the number of parts of ρ which are equal to k (i.e., the number of k-cycles in ρ). Here $\chi(\rho)$ is the value of the corresponding extreme character.

Equivalently, one can say that on the multiplicative basis $\{p_k\}$ one has

$$F(p_1) = 1, \qquad F(p_k) = \sum_{i=1}^{\infty} \alpha_i^k + (-1)^{k-1} \sum_{i=1}^{\infty} \beta_i^k, \qquad k \geq 2.$$

Another equivalent formulation which is important for our future developments is that the generating function for the complete homogeneous symmetric functions h_n is specialized as

$$H(t) := 1 + \sum_{n=1}^{\infty} h_n t^n \to e^{\gamma t} \prod_{i=1}^{\infty} \frac{1 + \beta_i t}{1 - \alpha_i t} \tag{4.2.11}$$

where $\gamma = 1 - \sum(\alpha_i + \beta_i) \geq 0$.

Remark 4.2.21. Note that in abstract setting, when we represent symmetric functions as formal power series in infinitely many variables $x_1, x_2, \ldots$, the generating series is equal to

$$H(t) = 1 + \sum_{n=1}^{\infty} h_n(x_1, x_2, x_3, \ldots) t^n = \prod_{i=1}^{\infty} \frac{1}{1 - x_i t}.$$

(See §4.2.2). Schur-positive specializations are *richer*[6] in the sense that we may include exponential term $e^{\gamma t}$ as well as terms of the form $1 + \beta_i t$.

4.3 References

The material on characters of finite-dimensional unitary groups can be located in the Weyl's book, [Wey97]. Other material on symmetric functions exists in the Macdonald's beautiful monograph [Mac95].

The Edrei–Voiculescu theorem is a result which has manifestations in various areas, and it has a long story. First a statement equivalent to that theorem was established in 1950's in context of classifying totally nonnegative Toeplitz matrices [Edr53] (see also [AESW51], [ASW52]). Extreme characters of $U(\infty)$ were

[6] If the variables x_i are nonnegative and sum to one, then p_1 specializes to 1, and we have a Schur-positive specialization (cf. §4.1.7). We may then treat x_i's as α_i's.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

classified by Voiculescu [Voi76]. A connection of Voiculescu's work with earlier results on totally nonnegative Toeplitz matrices was discovered in [Boy83]. The connection with Gelfand–Tsetlin graph and idea of approximation of characters of $U(\infty)$ by characters of finite unitary groups was suggested by Vershik and Kerov [VK81a], [VK82] and carried out by Okounkov and Olshanski [OO98]. The latter three papers considered asymptotics of Schur polynomials as they are generating series (in a certain sense) of the coherent systems on the Gelfand–Tsetlin graph. Recently [BO12], [Pet12] a more direct derivation using explicit formulas for relative dimensions was suggested. In these lectures we present the latter approach.

The Edrei-Thoma theorem is also a very important result, and its history goes along similar lines: the classification of totally nonnegative one-sided sequences (which is equivalent to the Edrei-Thoma theorem) is also due to Edrei in the 1950-s; then the result in terms of characters was established by E. Thoma in [Tho64]. Works of Vershik and Kerov in the 1980-s related the result to asymptotic behavior of Young diagrams. A generalization of the result may be found in [KOO98].

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Chapter 5

Totally nonnegative sequences, planar networks, Geseel-Viennot determinants

5.1 Unified picture of extreme characters of $U(\infty)$ and $\mathfrak{S}(\infty)$ and connection to totally nonnegative sequences

Now that we know classifications of extreme characters of $U(\infty)$ and $\mathfrak{S}(\infty)$, we are able to present a somewhat unified description of these results. These results (Edrei-Voiculescu and Edrei-Thoma theorems) are closely connected at the level of certain totally nonnegative sequences.

5.1.1 Totally nonnegative matrices: definition and examples

Definition 5.1.1. Let $X = [X_{ij}]$ be a matrix over $\mathbb{R}$ (finite or infinite). Let $I = \{i_1 < \cdots < i_n\}$ and $J = \{j_1 < \cdots < j_n\}$ be two finite sets of indices. A matrix $X_{IJ} := [X_{i_k, j_l}]_{k,l=1,\ldots,n}$ is called a *submatrix* of the matrix X. Its determinant $|X_{IJ}|$ is called a *minor* of X.

Definition 5.1.2. A matrix X is called *totally nonnegative* ($X \in \mathcal{TN}$) iff $|X_{IJ}| \geq 0$ for all finite sets of indices I and J (of same size), i.e., if all minors of X (of all orders) are nonnegative.

For example, for a 2×2 matrix $X = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, total nonnegativity means

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

precisely that $a, b, c, d \geq 0$ and $ad - bc \geq 0$.

Let us give more examples of totally nonnegative matrices:

1. Let $N \in \mathbb{Z}_{>0}$ and $1 > x_1 \geq \cdots \geq x_N \geq 0$ and $1 > y_1 \geq \cdots \geq y_N \geq 0$, then the $N \times N$ matrix
$$X_{ij} = \frac{1}{1 - x_i y_j} \tag{5.1.1}$$
is totally nonnegative (EX).

2. Let $N \in \mathbb{Z}_{>0}$ and $x_1 \geq \cdots \geq x_N > 0$ and $y_1 \geq \cdots \geq y_N > 0$, then the $N \times N$ matrix
$$X_{ij} = \frac{1}{x_i + y_j} \tag{5.1.2}$$
is totally nonnegative (EX).

3. Let $N \in \mathbb{Z}_{>0}$ and $x_1 \geq \cdots \geq x_N$ and $y_1 \geq \cdots \geq y_N$, then the $N \times N$ matrix
$$X_{ij} = e^{x_i y_j} \tag{5.1.3}$$
is totally nonnegative (EX).

Proposition 5.1.3 ((EX))**.** *IF X is a nondegenerate totally nonnegative finite matrix, then the matrix Y defined as $Y_{ij} = (-1)^{i+j} \left(X^{-1}\right)_{ij}$, is also totally nonnegative.*

Proposition 5.1.4. *If $X, Y \in \mathcal{TN}$ are matrices of finite size, then $XY \in \mathcal{TN}$.*

That is totally nonnegative matrices form a semigroup.

Proof. This follows from the Cauchy-Binet summation rule which we can write using the minors notation as
$$|(XY)_{IJ}| = \sum_{K\colon \#K = \#I} |X_{IK}| \cdot |Y_{KJ}|. \tag{5.1.4}$$
Here I and J are finite subsets of indices of the same size. □

5.1.2 Unitary group

Here we describe a correspondence between extreme characters of the infinite-dimensional unitary group $U(\infty)$ and totally nonnegative doubly infinite Toeplitz matrices with some normalization condition.

(from extreme characters to totally nonnegative matrices) Having an extreme character χ of $U(\infty)$, consider its values on the one-dimensional torus $\mathbb{T}^1 = U(1) \subset U(\infty)$, and consider Laurent decomposition of this function on the torus:
$$\chi|_{\mathbb{T}}(u) = \sum_{n \in \mathbb{Z}} \varphi_n u^n.$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Then, as we know (Lemma 4.1.23), the restrictions of χ to every finite-dimensional subgroup $U(N) \subset U(\infty)$ can be decomposed into a convex combination of extreme characters (= irreducible normalized characters) of $U(N)$ as follows:

$$\chi\big|_{U(N)}(u_1, \ldots, u_N) = \sum_{\nu \in \mathbb{GT}_N} \Big(\mathrm{Dim}_N\, \nu \cdot \det[\varphi_{\nu_i - i + j}]_{i,j=1}^N \Big) \frac{s_\nu(u_1, \ldots, u_N)}{s_\nu(1, \ldots, 1)}.$$

Since the coefficients of the convex combination $M_N(\nu) = \mathrm{Dim}_N\, \nu \cdot \det[\varphi_{\nu_i - i + j}]_{i,j=1}^N$ form a probability measure (i.e., they are nonnegative and sum to one), we conclude that the determinants $\det[\varphi_{\nu_i - i + j}]_{i,j=1}^N$ must be all nonnegative.

Observe that the above determinants $\det[\varphi_{\nu_i - i + j}]_{i,j=1}^N$ are minors of a doubly infinite Toeplitz[1] matrix $[\varphi_{j-i}]_{i,j \in \mathbb{Z}}$:

$$\begin{pmatrix} & \vdots & & & & \\ \cdots & \varphi_0 & \varphi_1 & \varphi_2 & \varphi_3 & \cdots \\ & \varphi_{-1} & \varphi_0 & \varphi_1 & \varphi_2 & \\ & \varphi_{-2} & \varphi_{-1} & \varphi_0 & \varphi_1 & \\ & \varphi_{-3} & \varphi_{-2} & \varphi_{-1} & \varphi_0 & \\ & \vdots & & & & \end{pmatrix} \tag{5.1.5}$$

Although these minors do not exhaust all possible minors of that matrix, their nonnegativity is sufficient to conclude that the matrix (5.1.5) is totally nonnegative. This is implied by the Fekete's Lemma, cf. [FZ00] and references therein.[2]

Remark 5.1.5. Instead of the matrix (5.1.5) one could also take a matrix which is infinite only in one direction:

$$\begin{pmatrix} \varphi_0 & \varphi_1 & \varphi_2 & \varphi_3 & \cdots \\ \varphi_{-1} & \varphi_0 & \varphi_1 & \varphi_2 & \\ \varphi_{-2} & \varphi_{-1} & \varphi_0 & \varphi_1 & \\ \varphi_{-3} & \varphi_{-2} & \varphi_{-1} & \varphi_0 & \\ \vdots & & & & \end{pmatrix}$$

Minors of these two matrices are the same.

The totally nonnegative matrices (5.1.5) have an additional restriction: $\sum_{n \in \mathbb{Z}} \varphi_n = 1$. That is, one more possible interpretation of the numbers $\{\varphi_n\}$ is that they define probability weights on $\mathbb{Z}$, or, in other words, an integer-valued random variable ξ with

$$\mathbb{P}(\xi = n) = \varphi_n.$$

[1]Recall that Toeplitz matrices are those whose elements are constant along diagonals.

[2]In fact, nonnegativity of all *solid* minors is enough to conclude that a matrix is totally nonnegative [FZ00]. A solid minor is a minor which occupies consecutive rows and columns.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Definition 5.1.6. A sequence $\{\varphi_n\}_{n\in\mathbb{Z}}$ is called *totally nonnegative* if the corresponding Toeplitz matrix (5.1.5) is totally nonnegative. Note that this implies that of course $\varphi_n \ge 0$ for all n.

Thus, we have the following correspondence:

> Extreme characters of $U(\infty)$ give rise to doubly infinite totally nonnegative sequences $\{\varphi_n\}_{n\in\mathbb{Z}}$ with additional property $\sum_{n\in\mathbb{Z}}\varphi_n = 1$.

(from totally nonnegative matrices to extreme characters) Passage from doubly infinite totally nonnegative sequences $\{\varphi_n\}_{n\in\mathbb{Z}}$ to extreme characters of $U(\infty)$ is less straightforward. First, let us understand how restrictive is the property $\sum_{n\in\mathbb{Z}}\varphi_n = 1$.

Log-concavity

Let φ_n be a totally nonnegative sequence, then $\varphi_n \ge 0$ for all n. The condition on 2×2 minors implies that

$$\det\begin{bmatrix}\varphi_n & \varphi_{n-1}\\ \varphi_{n+1} & \varphi_n\end{bmatrix} \ge 0, \qquad n\in\mathbb{Z}$$

This condition is equivalent to the *log-concavity* of the sequence $\{\varphi_n\}$:

$$\varphi_{n+1}\varphi_{n-1} \le (\varphi_n)^2.$$

From the log-concavity of a sequence one can already deduce some interesting properties.

Lemma 5.1.7 ((EX)). *1. The sequence $\varphi_n = cR^n$, where $c\ge 0$ and $R>0$, is totally nonnegative.*

2. If φ_n is a totally nonnegative sequence, then for all $c\ge 0$ and $R>0$ the sequence $M(n)\cdot cR^n$ is also totally nonnegative.

Lemma 5.1.8. *(EX) If φ_n is a log-concave sequence and $\varphi_n \ne cR^n$, then by replacing φ_n by $cR^n\varphi_n$ for suitable c and R, one can obtain the normalization condition $\sum_{n\in\mathbb{Z}}\varphi_n = 1$.*

Thus, the property $\sum_{n\in\mathbb{Z}}\varphi_n = 1$ of a sequence is not very restrictive as it rules out only the trivial case $\varphi_n = cR^n$. So from now on we can assume that $\{\varphi_n\}_{n\in\mathbb{Z}}$ are probability weights on $\mathbb{Z}$.

From the log-concavity condition it is also possible to get some bounds on tails of these probability measures:

Lemma 5.1.9. *If $\{\varphi_n\}$ is a log-concave sequence, then*

$$\varphi_n \le C\varepsilon^{|n|} \tag{5.1.6}$$

for some $0<\varepsilon<1$.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Proof. For simplicity, let all φ_n be nonzero (otherwise the same proof works with small modifications).

Consider $\frac{\varphi_{n+1}}{\varphi_n}$, $n \in \mathbb{Z}$. Since we assume that $\sum_{n\in\mathbb{Z}} \varphi_n = 1$, then all of $\frac{\varphi_{n+1}}{\varphi_n}$ cannot be ≥ 1. So, there exists n for which $\frac{\varphi_{n+1}}{\varphi_n} < 1$. Denote this quantity (for this particular n) by ε. Then by log-concavity, we have $\frac{\varphi_{n+2}}{\varphi_{n+1}} \leq \frac{\varphi_{n+1}}{\varphi_n} = \varepsilon$, and we get (5.1.6) for $n \to +\infty$. For $n \to -\infty$ the estimate is obtained in a similar way. □

Multiplicativity of extreme characters

Let us now take a probability measure φ_n on $\mathbb{Z}$ which is totally nonnegative in the sense of Definition 5.1.6. One can readily form a family of probability measures $M_N(\nu)$ on $\mathbb{GT}_N$ for all N which are coherent on the Gelfand–Tsetlin graph:

$$M_N(\nu) := \operatorname{Dim}_N \nu \cdot \det[\varphi_{\nu_i - i + j}]_{i,j=1}^{N}, \qquad \nu \in \mathbb{GT}_N. \tag{5.1.7}$$

(EX): these measures indeed satisfy coherency property.

However, it is not obvious that the measures M_N form an extreme coherent system. This fact may be established using an a priori *multiplicativity property* of extreme characters of $U(\infty)$. This multiplicativity property (discussed below in this subsection) means that the character χ is a product of functions of eigenvalues:

$$\chi(u_1, u_2, \ldots) = \prod_{i=1}^{\infty} f(u_i).$$

Starting from a totally nonnegative sequence and defining the measures M_N by (5.1.7), we (by Lemma 4.1.23) conclude that the character of $U(\infty$ corresponding to $\{M_N\}$ is extreme. Thus,

> Any doubly infinite totally nonnegative sequence $\{\varphi_n\}_{n\in\mathbb{Z}}$ with additional property $\sum_{n\in\mathbb{Z}} \varphi_n = 1$ gives rise to an extreme character of $U(\infty)$.

This fact is established modulo the multiplicativity property. We will not discuss this in every detail, but just give some definitions and explain briefly how the infinite-dimensional nature of the group $U(\infty)$ suggests this multiplicativity.

Let $g \in U(\infty)$. Note that g can be understood as a finite matrix because g differs from the unity of the group $U(\infty)$ only at finitely many positions. By $[g]$ denote the conjugacy class of $g \in U(\infty)$.

Definition 5.1.10. Let Γ be the set of all conjugacy classes in $U(\infty)$. This set Γ may be understood as a set of all finite unordered tuples of the form $\{u_1, \ldots, u_n\}$ (where $u_i \in \mathbb{T}^1$), under the equivalence relation

$$\{u_1, \ldots, u_n\} \sim \{u_1, \ldots, u_n, 1\}. \tag{5.1.8}$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

It is possible to equip Γ with a structure of a commutative semigroup. For $\gamma_1, \gamma_2 \in \Gamma$, set

$$\gamma_1 \circ \gamma_2 := \gamma_1 \sqcup \gamma_2 \tag{5.1.9}$$

(disjoint union of γ_1 and γ_2).

On the level of matrices this operation looks as follows:

$$[g_1] \circ [g_2] := \left[\begin{pmatrix} g_1 & 0 \\ 0 & g_2 \end{pmatrix} \right]. \tag{5.1.10}$$

Denote the matrix in the right-hand side by $g_1 \oplus g_2$.

Note that this operation and this semigroup structure are specific for the infinite-dimensional nature of $U(\infty)$; it is not possible to define such an operation for $U(N)$. This is a manifestation of the principle that the group $U(\infty)$ is at the same time simpler and more complicated than the finite unitary groups $U(N)$.

Theorem 5.1.11 (Multiplicativity)**.** *Any extreme character χ of $U(\infty)$ is multiplicative on conjugacy classes:*

$$\chi(\gamma_1 \circ \gamma_2) = \chi(\gamma_1)\chi(\gamma_2), \qquad \gamma_1, \gamma_2 \in \Gamma. \tag{5.1.11}$$

(A character of $U(\infty)$ is essentially a function on the semigroup Γ.)

Of course, the multiplicativity statement is straightforward from the Edrei-Voiculescu theorem, but there is an independent argument proving this theorem.

Similar multiplicativity theorems hold for the infinite symmetric group $\mathfrak{S}(\infty)$ (see below), and also for other infinite-dimensional classical groups $O(\infty)$ and $Sp(2\infty)$.

In this way, we see that classification of extreme characters of $U(\infty)$ is equivalent to classification of totally nonnegative sequences, and we have just explained how one of these statements follows from the other one.

5.1.3 Symmetric group

Extreme characters of the infinite symmetric group $\mathfrak{S}(\infty)$ are in one-to-one correspondence with specializations of the algebra of symmetric functions *Sym* (§4.2.11): algebra homomorphisms $F\colon \mathit{Sym} \to \mathbb{R}$ such that

- $F(p_1) = 1$
- $F(s_\lambda) \geq 0$ for all Young diagrams λ.

A specialization F is completely determined by the sequence $c_n = F(h_n)$, $n = 0, 1, 2, \ldots$ (of course, $c_0 = 1$). The sequence $\{c_n\}$ is totally nonnegative because (by Jacobi–Trudi formula)

$$F(s_\lambda) = \det[c_{\lambda_i - i + j}]_{i,j=1}^N$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

(where, by agreement, $c_{-1} = c_{-2} = \ldots = 0$, and N is any number $\geq \ell(\lambda)$). The nonnegativity of these minors is sufficient (see the previous subsection) for the whole sequence c_n to be totally nonnegative. This corresponds to total nonnegativity of the unitriangular Toeplitz matrix:

$$\begin{pmatrix} 1 & c_1 & c_2 & c_3 & \cdots \\ 0 & 1 & c_1 & c_2 & \\ 0 & 0 & 1 & c_1 & \\ 0 & 0 & 0 & 1 & \\ \vdots & & & & \end{pmatrix}$$

Thus,

> Extreme characters of $\mathfrak{S}(\infty)$ correspond to totally nonnegative sequences $\{c_n\}_{n\in\mathbb{Z}_{\geq 0}}$ with additional property $c_0 = 1$.

The passage from totally nonnegative sequences to extreme characters is more straightforward with the help of the ring theorem (§4.2.11). Namely, we know that extreme characters correspond to specializations (= *algebra* homomorphisms) of the algebra of symmetric functions, so from a sequence $\{c_n\}_{n\in\mathbb{Z}_{\geq 0}}$ one readily reconstructs a character by considering specialization defined as

$$F(h_n) = c_n, \qquad n = 0, 1, 2, \ldots.$$

The total nonnegativity of the sequence ensures that this specialization is nonnegative on Schur functions.

Thus,

> Any totally nonnegative sequence $\{c_n\}_{n\in\mathbb{Z}_{\geq 0}}$ with $c_0 = 1$ gives rise to an extreme character of $\mathfrak{S}(\infty)$.

Note that the condition $c_0 = 1$ is not very restrictive because of the normalization, see the discussion of log-concavity in the previous subsection.

Remark 5.1.12. There is an a priori argument showing that extreme characters of $\mathfrak{S}(\infty)$ must be multiplicative with respect to a semigroup structure on conjugacy classes. The semigroup operation is defined similarly to the unitary group $U(\infty)$ and is also specific for the infinite-dimensional situation. The simplest way to define this operation is to represent permutations as permutation matrices (so elements of $\mathfrak{S}(\infty)$ are infinite permutation matrices which differ from the identity matrix at finitely many positions), then the operation $\circ$ on conjugacy classes is defined as in Definition 5.1.10.

Extreme characters of $\mathfrak{S}(\infty)$ are multiplicative with respect to this operation:

$$\chi(g_1 \circ g_2) = \chi(\rho_1 \sqcup \rho_2) = \chi(\rho_1)\chi(\rho_2),$$

where $\rho_1 \sqcup \rho_2$ corresponds to taking union of the cycles in cycle representation of permutations g_1 and g_2. This fact of course follows from the Edrei-Thoma theorem, but there also exist straightforward independent proofs of this multiplicativity.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

5.1.4 Conclusion

We have seen that extreme characters of $U(\infty)$ and $\mathfrak{S}(\infty)$ correspond to doubly infinite and one-sided infinite totally nonnegative sequences, respectively. This (as well as relevant formulas (4.1.21) and (4.2.11)) reflects the similarly between these two problems of asymptotic representation theory. More connections at the level of branching graphs are discussed in [BO13].

In the rest of this chapter we will discuss total nonnegativity (and its connections with characters of the infinite-dimensional unitary group) in more detail.

5.2 Gessel-Viennot determinants

Here we discuss a celebrated combinatorial result about counting nonintersecting path ensembles. This result is first due to Karlin and McGregor [KM59], and then independently due to Lindstroem [Lin73] and Gessel and Viennot [GV85]. The theory of counting nonintersecting paths leads to explicit constructions of totally nonnegative matrices (not necessary Toeplitz).

We will first speak about finite matrices.

Let $\Gamma = (V, E, w)$ be a directed graph with vertices V, edges $E \subset V \times V$ and nonnegative weights on edges $w\colon E \to \mathbb{R}_{\geq 0}$. We assume that the graph does not have multiple edges.

By P we will denote oriented paths along edges of the graph, by $w(P) := \prod_{e\in P} w(e)$ we mean the weight of a path P. Denote by $\mathrm{Path}(v', v'')$ the set of all oriented paths from v' to v''. Assume that

(C1) For any two vertices v' and v'' the number of paths from v' to v'' is finite.

Note that this condition forbids cycles or loops in Γ.

Definition 5.2.1. Let there be two ordered sets of vertices of the same cardinality n:

$$V' = \{v'_1, \dots, v'_n\} \subset V, \qquad V'' = \{v''_1, \dots, v''_n\} \subset V. \tag{5.2.1}$$

The data (V, V', V'', E, w) defines a $n \times n$ matrix X as follows:

$$X_{ij} := \sum_{P \in \mathrm{Path}(v'_i, v''_j)} w(P). \tag{5.2.2}$$

By (C1), this matrix is well-defined.

By the definition of a determinant:

$$\det X = \sum_{s \in \mathfrak{S}(n)} \mathrm{sign}(s) X_{1,s(1)} \dots X_{n,s(n)} = \sum_{P_1,\dots,P_n} \mathrm{sign}(P_1, \dots, P_n) w(P_1) \dots w(P_n), \tag{5.2.3}$$

where the second sum is taken over all sets of paths $P_1, \dots, P_n$, such that the beginning of each path P_i is in V', and the end of each path is in V'', and the

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

paths $(P_1, \dots, P_n)$ define a bijection between V' and V''. The sign of a path ensemble $\operatorname{sign}(P_1, \dots, P_n)$ means the sign of this bijection between V' and V'' (recall that the sets V' and V'' are ordered). Denote the set of all such path ensembles by $\mathcal{P}$, and denote each ensemble $(P_1, \dots, P_n)$ by $\overrightarrow{P}$ (to shorten the notation). Also, by $\mathcal{P}^\circ \subset \mathcal{P}$ denote the subset consisting of all ensembles $\overrightarrow{P}$ such that any two paths P_i and P_j from $\overrightarrow{P}$ do not intersect (at a vertex).

Lemma 5.2.2 (Gessel-Viennot's lemma)**.** *In the above notation, we have*

$$\det X = \sum_{\overrightarrow{P} \in \mathcal{P}^\circ} \operatorname{sign}(\overrightarrow{P}) w(\overrightarrow{P}). \tag{5.2.4}$$

Proof. We will proof the lemma by the *involution method*. We know that

$$|X| = \sum_{\overrightarrow{P} \in \mathcal{P}} \operatorname{sign}(\overrightarrow{P}) w(\overrightarrow{P}) = \sum_{\overrightarrow{P} \in \mathcal{P}^\circ} \operatorname{sign}(\overrightarrow{P}) w(\overrightarrow{P}) + \sum_{\overrightarrow{P} \in \mathcal{P} \setminus \mathcal{P}^\circ} \operatorname{sign}(\overrightarrow{P}) w(\overrightarrow{P}). \tag{5.2.5}$$

To show that the last sum is zero, it suffices to construct an involution

$$*\colon \mathcal{P} \setminus \mathcal{P}^\circ \to \mathcal{P} \setminus \mathcal{P}^\circ, \qquad \overrightarrow{P} \mapsto \overrightarrow{P}^*, \tag{5.2.6}$$

such that

- $w(\overrightarrow{P}^*) = w(\overrightarrow{P})$;
- $\operatorname{sign}(\overrightarrow{P}^*) = -\operatorname{sign}(\overrightarrow{P})$.

We can construct such an involution as follows. Let $\overrightarrow{P} \in \mathcal{P} \setminus \mathcal{P}^\circ$. This means that some paths in the ensemble $\overrightarrow{P}$ intersect. Let i be the minimal number such that the path P_i intersects some other path from $\overrightarrow{P}$. Let v be the first vertex where P_i intersects some other path, and let j be the minimal number $(j > i)$ among the numbers of all paths $P_1, \dots, P_n$, which pass through the vertex v. These conditions uniquely define the numbers i and j, as well as the vertex v.

By definition, the involution $*$ consists in swapping the tails of the paths P_i and P_j after the vertex v, see Fig. 5.1. Clearly, $*$ satisfies our two conditions,

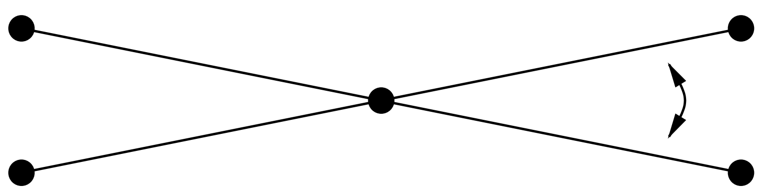


Figure 5.1: The involution $*$.

and this concludes the proof. $\square$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Remark 5.2.3. We note that Lemma 5.2.2 has a purely algebraic nature and the condition $w(\,\cdot\,) \in \mathbb{R}_{\geq 0}$ is not necessary for this lemma.

In what follows, however, we will need the condition $w(\,\cdot\,) \in \mathbb{R}_{\geq 0}$.

Let us impose one more condition on (V, V', V'', E, w):

(C2) If $i_1 < i_2$ and $j_1 < j_2$, $P \in \mathrm{Path}(v'_{i_1}, v''_{j_2})$ and $Q \in \mathrm{Path}(v'_{i_2}, v''_{j_1})$, then $P \cap Q \neq \varnothing$.

We say that the conditions (C1)–(C2) define a *planar network*.

Proposition 5.2.4. *Let $V', V'' \subset V$, $\#V' = \#V'' = N$, and the data (V, V', V'', E, w) defines a planar network (i.e., satisfies (C1)–(C2)). Let X be the matrix from Definition 5.2.1. Then the matrix X is totally nonnegative.*

Proof. This readily follows from Lemma 5.2.2. Indeed, consider the IJ-minor of the matrix X, where $\#I = \#J = n$, and $V'_I \subset V'$, $V''_J \subset V''$. Then by Lemma 5.2.2 we have

$$|X_{IJ}| = \sum_{\overrightarrow{P} \in \mathcal{P}^\circ} \mathrm{sign}(\overrightarrow{P}) w(\overrightarrow{P}) \tag{5.2.7}$$

(here $\mathcal{P}^\circ$ is defined in accordance with V'_I and V''_J). From (C2) follows that all ensembles of nonintersecting paths $\overrightarrow{P} \in \mathcal{P}^\circ$ preserve the ordering of V'_I and V''_J. So, we have $\mathrm{sign}(\overrightarrow{P}) = +1$ for all $\overrightarrow{P} \in \mathcal{P}^\circ$, and since $w \in \mathbb{R}_{\geq 0}$, then $|X_{IJ}| \geq 0$. This concludes the proof. □

This proposition gives a tool to prove total nonnegativity of matrices. Consider an example.

Example 5.2.5. Consider the graph V on Fig. 5.2 (here $n = 4$, but a similar graph can be constructed for any n). All edges are oriented to the right and have weight 1. The matrix X for it has the form

$$X = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 1 & 3 & 3 & 1 \end{bmatrix}. \tag{5.2.8}$$

We see that the entries of this matrix are binomial coefficients arranged like in the Pascal triangle.

From Proposition 5.2.4 it follows that X is a totally nonnegative matrix. The same holds for a matrix of any order n constructed from the Pascal triangle in a similar way.

Remark 5.2.6. The condition (C2) is satisfied if the graph V can be drawn on the plane such that all the edges are oriented to the right, the vertices from V' are located from top to bottom on the left, and the vertices from V'' — along a parallel vertical line on the right, and the order of vertices in V' and V'' is from top to bottom (or from bottom to top). See Fir. 5.2.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

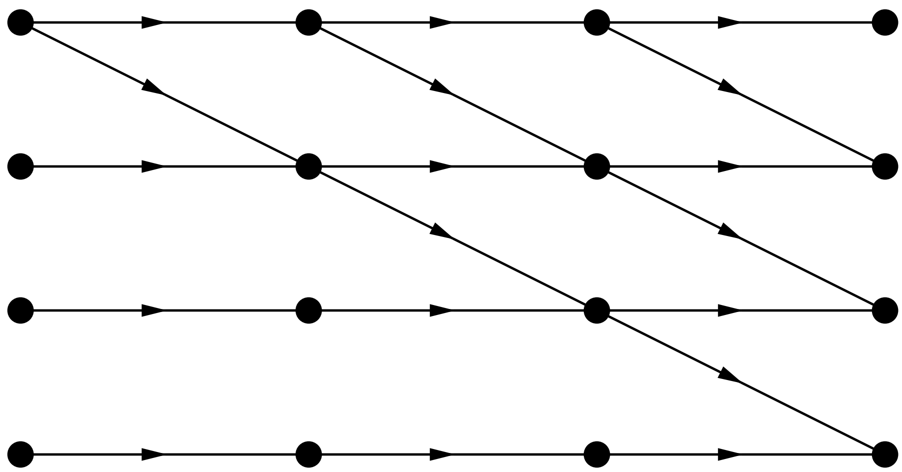


Figure 5.2: Graph V.

Remark 5.2.7. Let X and Y be totally nonnegative matrices of same size constructed from two graphs (X corresponds to starting vertices V' and ending vertices V'' in a graph Γ_X, and for Y the starting and the ending vertices are U' and U'', respectively, in a graph Γ_Y; assume that Γ_X and Γ_Y do not have common vertices).[3] The product XY is a totally nonnegative matrix which can be understood as a matrix constructed using the concatenation of the two graphs obtained by identifying the vertices V'' and U' (i.e., v_i'' with u_i', $i = 1, \ldots, \#V''$). The condition (C2) continues to hold.

5.3 Example with Catalan-Hankel determinants

Let us consider one more (not very related) example which illustrates the technique of Gessel–Viennot determinants. Let

$$C_n := \frac{1}{n+1}\binom{2n}{n}, \qquad n = 0, 1, 2, \ldots$$

be the Catalan numbers,

$$C_0 = 1, \qquad C_1 = 1, \qquad C_2 = 2, \qquad C_3 = 5, \qquad C_4 = 14, \qquad \ldots$$

Consider the Catalan-Hankel determinants:

$$H_k(n) := \det[C_{i+j+k-2}]_{i,j=1}^n, \qquad k = 0, 1, 2, \ldots$$

[3] We also assume that (C2) holds.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

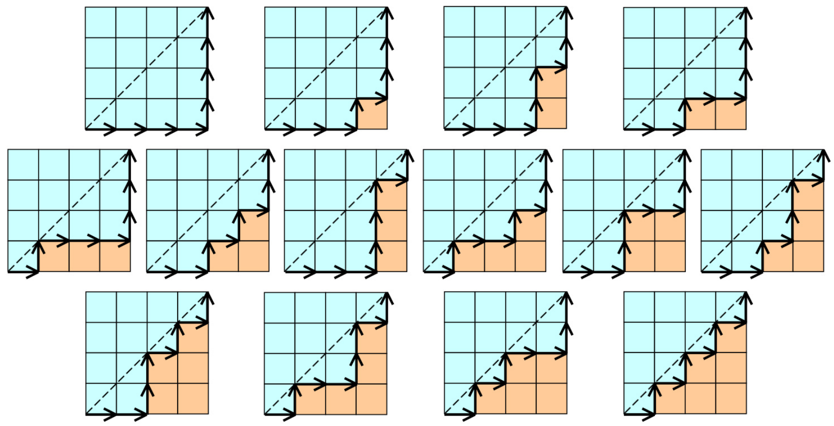


Figure 5.3: Staircase (Dyck) paths for $n = 4$.

We will prove (using Gessel-Viennot determinants) that

$$H_0(n) = H_1(n) = 1, \qquad H_2(n) = n + 1.$$

In fact, there is a general explicit formula for such determinants, e.g., see [ITZ10] and references therein.

As is known (e.g., see [Sta01][4]), Catalan numbers enumerate up-right lattice paths (see Fig. 5.3) from $(0, 0)$ to (n, n) which go (weakly) below the diagonal. They can be called Dyck paths (although Dyck paths are often from $(0, 0)$ to $(0, 2n)$ with steps $(1, 1)$ and $(1, -1)$ which do not go below the x-axis).

Let us represent

$$H_0(n) = \det \begin{bmatrix} C_0 & C_1 & C_2 & \cdots \\ C_1 & C_2 & C_3 & \\ C_2 & C_3 & C_4 & \\ \vdots & & & \end{bmatrix}$$

as a Gessel-Viennot determinant for some number of nonintersecting path ensembles, for which the total number of ensembles is obvious.

Consider the part of the up-right grid as on Fig. 5.3 below the diagonal, and let this be our graph V (with all edges having multiplicity 1). Let

$$V' := \{(0, 0), (-1, -1), (-2, -2), \ldots, (-(n - 1), -(n - 1))\},$$

and

$$V'' := \{(0, 0), (1, 1), (2, 2), \ldots, (n - 1, n - 1)\}$$

[4]Catalan numbers have numerous combinatorial interpretations, cf. `http://www-math.mit.edu/~rstan/ec/catadd.pdf`.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

(in this order). Consider the matrix X constructed from this planar network. Clearly, this matrix coincides with the matrix whose determinant is $H_0(n)$. But it is readily seen that the total number of nonintersecting path ensembles connecting V' and V'' is just 1. Thus, $H_0(n) = 1$.

(EX) The formulas for $H_1(n)$ and $H_2(n)$ are proven in a similar way.

Remark 5.3.1. The sequence of Catalan numbers is in fact the only integer sequence for which the Hankel determinants $H_0(n)$ and $H_1(n)$ are all equal to one.

5.4 Totally nonnegative Toeplitz matrices via Gessel-Viennot theory

The question of classification of totally nonnegative sequences was motivated (and raised by Schoenberg and co-authors in 1950-s) by data smoothing problems in numerical analysis. Let $x_i \in \mathbb{R}$, $i \in \mathbb{Z}$, be a data sequence, and $\{M(j)\}$, $j \in \mathbb{Z}$, be a 'template', or 'mask'. A 'smoothed' sequence is constructed as a convolution of x with M: $y_i = \sum_{j\in\mathbb{Z}} x_{i-j}M(j)$, denote the convolution by $x * M$. Schoenberg in 1950-s showed that the convolution $x \mapsto y$ with 'mask' M does not increase the number of oscillations (= sign changes) iff the Toeplitz matrix $X_{ij} = M(i-j)$ is totally nonnegative.

Convolution

There are four equivalent objects for which the property of total nonnegativity is defined:

1. Toeplitz matrices $[M(j-i)]_{i,j\in\mathbb{Z}}$
2. doubly infinite sequences $\{M(n)\}_{n\in\mathbb{Z}}$
3. probability measures on $\mathbb{Z}$, i.e., integer-valued random variables ξ with $\mathbb{P}(\xi = n) = M(n)$
4. generating functions $\psi(u) = \sum_{n\in\mathbb{Z}} M(n)u^n$, or 'characteristic functions' (= Polya frequency functions) of probability distributions.

In fact, every object can be used to reconstruct all other objects. A probability measure is uniquely determined by its characteristic functions which converges inside an annulus around the unit circle, and defines there a holomorphic function.

Also, Gessel-Viennot theory gives a graph construction which can lead to any of these objects. Here we will discuss how totally nonnegative matrices arise from graphs.

Totally nonnegative measures are log-concave If M' and M'' are two totally nonnegative measures, then by Lemma 5.1.9 the corresponding Toeplitz matrices

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

can be multiplied (despite their infinite size). This corresponds to convolution of probability measures

$$(M' * M'')(n) := \sum_{k\in\mathbb{Z}} M'(n-k)M''(k). \tag{5.4.1}$$

That is, $M' * M''$ is the distribution of the sum $\xi' + \xi''$ of two independent random variables corresponding to M' and M''.

The convolution of totally nonnegative distributions corresponds to the usual product of their characteristic functions.

Thus, totally nonnegative measures form a convolution semigroup.

Two examples

Consider first two examples of totally nonnegative matrices (cf. §4.1.10).

Example 5.4.1. Let M be the Bernoulli measure:

$$M(n) := \begin{cases} 1-\beta, & \text{if } n = 0; \\ \beta, & \text{if } n = 1; \\ 0, & \text{otherwise.} \end{cases} \tag{5.4.2}$$

Here $0 \le \beta \le 1$. This measure corresponds to the Toeplitz matrix of the form

$$X = \begin{bmatrix} \ddots & \ddots & \ddots & \ddots & \ddots \\ \ddots & 1-\beta & 0 & 0 & \ddots \\ \ddots & \beta & 1-\beta & 0 & \ddots \\ \ddots & 0 & \beta & 1-\beta & \ddots \\ \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}. \tag{5.4.3}$$

Clearly, this matrix arises from the graph on Fig. 5.4.[5] This graph satisfies (C2), that is, it is a planar network, so the Bernoulli measure is totally nonnegative. Its characteristic function is

$$\psi(u) = 1 + \beta(u-1). \tag{5.4.4}$$

Example 5.4.2. Let now M be a geometric distribution on nonnegative integers:

$$M(n) := \begin{cases} \dfrac{1}{1+\alpha}\left(\dfrac{\alpha}{1+\alpha}\right)^n, & \text{if } n \ge 0; \\ 0, & \text{otherwise.} \end{cases} \tag{5.4.5}$$

[5]This matrix as well as the graph are infinite, but this does no affect the considerations of the Gessel-Viennot theory.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

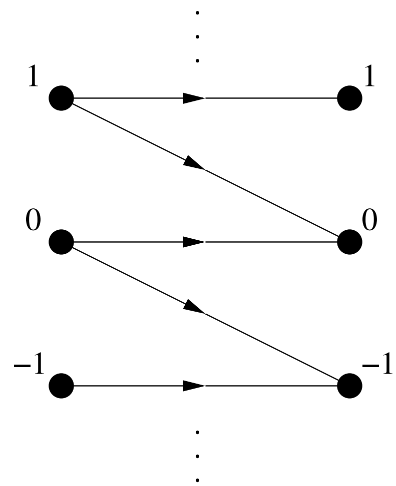


Figure 5.4: Graph for the Toeplitz matrix (5.4.3). The horizontal edges have weight $1-\beta$, and the down-right (= south-east) edges have weight β.

Here $\alpha \geq 0$. The corresponding Toeplitz matrix arises from the graph on Fig. 5.5. It also satisfies (C2), i.e., it is also a planar network. Thus, the measure M is totally nonnegative as well. Its characteristic function is

$$\psi(u) = \frac{1}{1-\alpha(u-1)}. \tag{5.4.6}$$

How about general totally nonnegative matrices?

Replacement $u \mapsto u^{-1}$ in a characteristic function $\psi(u)$ means reflection of the measure $M(n)$ with respect to zero. This also produces a totally nonnegative measure.

Using this observation and examples 5.4.1 and 5.4.2, we see that *finite* products of the form

$$\psi(u) = \frac{\prod(1+\beta_i^+(u-1))}{\prod(1-\alpha_i^+(u-1))} \cdot \frac{\prod(1+\beta_i^-(u^{-1}-1))}{\prod(1-\alpha_i^-(u^{-1}-1))} \tag{5.4.7}$$

also serve as characteristic functions of totally nonnegative sequences. (cf. the functions (4.1.21)).

It can be readily seen that weak limits of probability measures on $\mathbb{Z}$ (this corresponds to uniform convergence of characteristic functions on $\mathbb{T}^1$) preserve total nonnegativity.

Consider the following sequence of characteristic functions:

$$\psi_n(u) := \Big(1 + \frac{\gamma}{n}(u-1)\Big)^n, \qquad \gamma \geq 0. \tag{5.4.8}$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

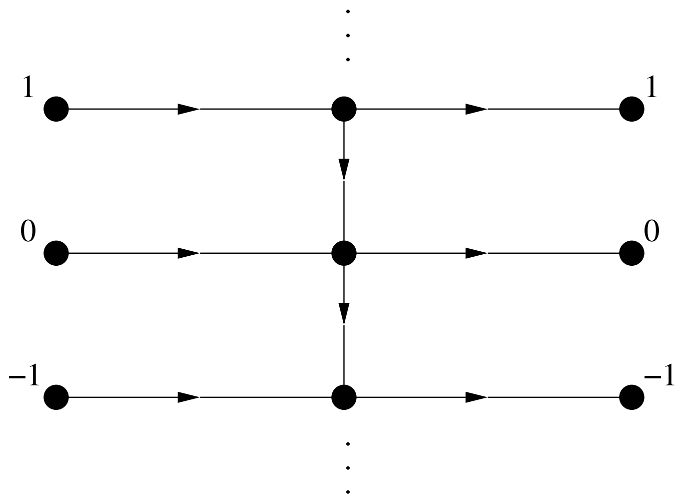


Figure 5.5: Planar network for the Toeplitz matrix corresponding to the geometric distribution. The weight of every right horizontal edge is $\frac{1}{1+\alpha}$, the weights of left horizontal edges are 1, and each vertical edge has weight $\frac{\alpha}{1+\alpha}$.

The characteristic function ψ_n is a convolution of n identical Bernoulli measure characteristic functions with parameters $\beta_n = \frac{\gamma}{n}$. As $n \to \infty$, the functions ψ_n converge (uniformly on $\mathbb{T}^1$) to a characteristic function of the Poisson distribution $e^{\gamma(u-1)}$.

Thus, limits of the characteristic functions (5.4.7) produce all the characteristic functions of the form (4.1.21):

$$\Psi_{\alpha^\pm,\beta^\pm,\gamma^\pm}(u) = e^{\gamma^+(u-1)+\gamma^-(u^{-1}-1)} \prod_{i=1}^{\infty} \left(\frac{1+\beta_i^+(u-1)}{1-\alpha_i^+(u-1)} \cdot \frac{1+\beta_i^-(u^{-1}-1)}{1-\alpha_i^-(u^{-1}-1)} \right). \tag{5.4.9}$$

Here $\sum \alpha_i^+ + \sum \alpha_i^- + \sum \beta_i^+ + \sum \beta_i^- < \infty$ and $\gamma^\pm \geq 0$.

In the 1950-s Schoenberg conjectured and Edrei proved that these characteristic functions exhaust characteristic functions of all totally nonnegative doubly infinite sequences.

Remark 5.4.3 (On the condition $\beta_1^+ + \beta_1^- \leq 1$)**.** In our asymptotic representation-theoretic constructions the condition $\beta_1^+ + \beta_1^- \leq 1$ arised as follows. We constructed the parameters $(\alpha^\pm, \beta^\pm, \gamma^\pm)$ from a regular sequence of signatures $\nu(N)$, so $\alpha^\pm$ and $\beta^\pm$ were limiting row/column lengths of the Young diagrams $\nu^\pm$, and $\gamma^\pm = \lim_{N\to\infty} \frac{|\nu(N)|}{N} - \sum(\alpha_i^\pm + \beta_i^\pm)$ are the limiting 'deficiencies' Indeed, $\beta_1^\pm = \ell(\nu^\pm) - \frac{1}{2}$, and $\ell(\nu^+) + \ell(\nu^-) \leq N$.

In Schoenberg's argument the condition $\beta_1^+ + \beta_1^- \leq 1$ does not arise naturally[6], but this condition, however, does not restrict the set of possible charac-

[6]Note that this condition is not formally necessary for the characteristic function to be

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

teristic functions:

> Take any two indices i and j and replace β_i^+ by $1-\beta_j^-$, and also replace β_j^- by $1-\beta_i^+$. The function (4.1.21) does not change under this transformation.
>
> If $\beta_1^+ + \beta_1^- \le 1$ does not hold, then by a finite number of these transformations one can always force this condition to hold.

Under the condition $\beta_1^+ + \beta_1^- \le 1$, the parameters of a totally nonnegative sequence are uniquely determined: if $(\alpha_1^\pm, \beta_1^\pm, \gamma_1^\pm) \ne (\alpha_2^\pm, \beta_2^\pm, \gamma_2^\pm)$, then the corresponding characteristic functions (5.4.9) are distinct.

The parameter set

$$\Omega = \Big\{(\alpha^\pm, \beta^\pm, \gamma^\pm)\colon \sum \alpha_i^+ + \sum \alpha_i^- + \sum \beta_i^+ + \sum \beta_i^- < \infty,\ \gamma^\pm \ge 0,\ \beta_1^+ + \beta_1^- \le 1\Big\} \tag{5.4.10}$$

can be embedded into $\mathbb{R}^{4\infty+2}$, and Ω will be a closed locally compact subset in $\mathbb{R}^{4\infty+2}$ (with the topology of coordinate-wise convergence).

Conclusion

Thus, we have explained how characteristic functions of *all possible* totally nonnegative sequences arise from certain (namely, Bernoulli and geometric) building blocks, and how these building blocks can be interpreted through graphs (more precisely, planar networks).

5.5 Schur polynomials via Gessel-Viennot theory

Let λ be a nonnegative signature of length N. Our aim here is to represent the Schur polynomial $s_\lambda(x_1, \dots, x_N)$ as a generating function (= partition function) of some ensembles of nonintersecting paths on a grid. This will reprove the Jacobi-Trudi identity. This is another use of the Gessel-Viennot determinants.

Using $\lambda = (\lambda_1 \ge \lambda_2 \ge \dots \ge \lambda_N \ge 0)$ construct a sequence of distinct nonnegative integers $l = (l_1 > l_2 > \dots > l_N \ge 0)$, where

$$l_i := \lambda_i + N - i, \qquad i = 1, \dots, N.$$

Interpret l as a point configuration in $\mathbb{Z}_{\ge 0}$. Also set

$$l^0 := (N-1, N-2, \dots, 1, 0)\,.$$

The configuration l^0 corresponds to an empty Young diagram.

Consider the following nonintersecting path model. Place points of the configurations l^0 and l on the lattice as on Fig. 5.6. The points of the configuration

defined. On the contrast, the condition $\sum \alpha_i^+ + \sum \alpha_i^- + \sum \beta_i^+ + \sum \beta_i^- < \infty$ ensures that the infinite product converges.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

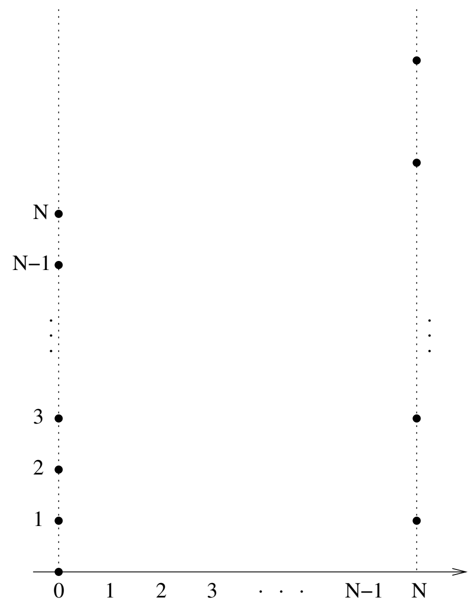


Figure 5.6

l^0 (i.e. $(N-1, N-2, \ldots, 1, 0)$) will be on the left as starting points, and the points of the configuration l — ending points — will be on the right.

Consider N nonintersecting paths connecting l_i^0 with l_i, $i = 1, 2, \ldots, N$. The paths will go along the edges of the oriented graph displayed on Fig. 5.7 (note that there are no edges going up on the left). An example of a nonintersecting path ensemble is given on Fig. 5.8.

There is a bijection between the set $SSYT(\lambda; N)$ of semistandard Young tableaux of shape λ and the set of nonintersecting path ensembles described above. Indeed, the semistandard tableau of shape λ is a sequence of Young diagrams

$$\varnothing = \lambda^{(0)} \subset \lambda^{(1)} \subset \cdots \subset \lambda^{(N)} = \lambda, \qquad \lambda^{(i)} \in \mathbb{Y}(i),$$

with condition that for all i the diagram $\lambda^{(i)}/\lambda^{(i-1)}$ is a horizontal strip. The configuration $l(i)$ (where i represents 'time' measured along the x-axis) corresponds to the diagram $\lambda^{(i)}$.

Let us first compute $s_\lambda(\underbrace{1, \ldots, 1}_{N}) \operatorname{Dim}_N \lambda = |SSYT(\lambda; N)|$, as a number of all ensembles of nonintersecting paths:

$$\operatorname{Dim}_N \lambda = \det \left[A(i,j)\right]_{i,j=1}^{N},$$

where $A(i,j)$ — is the number of all paths from $l_i^0 = N - i$ to $l_j = \lambda_j + N - j$ along the edges of the oriented graph on Fig. 5.7.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

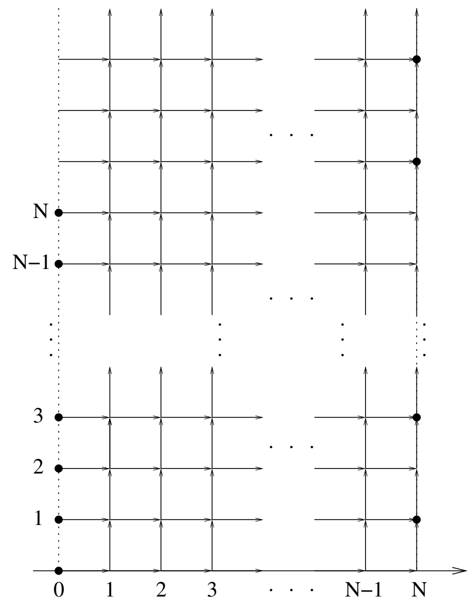


Figure 5.7

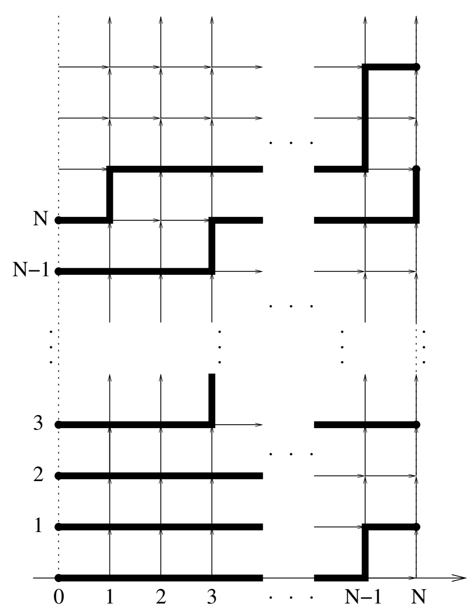


Figure 5.8

Clearly, $A(i,j)$ is the binomial coefficient

$$A(i,j) = \frac{(\lambda_j - j + i + N - 1)!}{(\lambda_j - j + i)!(N-1)!}. \tag{5.5.1}$$

The determinant

$$\mathrm{Dim}_N \lambda = \det[A(i,j)]_{i,j=1}^N = \det\left[\binom{l_j + i - 1}{N-1}\right]_{i,j=1}^N$$

may be computed by one of the techniques explained in the survey [Kra99]:

Proposition 5.5.1. *We have*

$$\det\left[\binom{l_j + i - 1}{N-1}\right]_{i,j=1}^N = \prod_{1 \le i < j \le N} \frac{l_i - l_j}{j - i}.$$

Now, put the weight x_i on every vertical edge in the graph (Fig. 5.7) which is on the line where the horizontal coordinate equals i.

Then one can readily check that $A(i,j) = h_{\lambda_j - j + i}(x_1, \ldots, x_N)$ (this is the weighed sum of paths from l_0^i to l_j), where h_k's are the complete homogeneous symmetric polynomials. Thus, we reproduce the Jacobi-Trudi identity from the Gessel-Viennot theory:

$$s_\lambda(x_1, \ldots, x_N) = \det[h_{\lambda_j - j + i}(x_1, \ldots, x_N)]_{i,j=1}^N.$$

Of course,

$$h_m(\underbrace{1, \ldots, 1}_{N}) = \frac{(m + N - 1)!}{m!(N-1)!}.$$

5.6 References

Totally nonnegative matrices are discussed in detail, e.g., in the books [GK02] and [Kar68], and in the paper [FZ00] (from various viewpoints).

Unified treatment of extreme characters of $U(\infty)$ and $\mathfrak{S}(\infty)$ — i.e., of boundaries of the Gelfand–Tsetlin and Young graphs — can be pushed even further, see [BO13].

Log-concave sequences are discussed in [Oko97].

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Chapter 6

Determinantal point processes

We begin to discuss *determinantal point processes* — a very useful formalism for studying probabilistic models arising from our representation-theoretic constructions.

6.1 Generalities [Bor11a]

6.1.1 Point processes

Let $\mathfrak{X}$ be a locally compact separable topological space. A *point configuration* X in $\mathfrak{X}$ is a locally finite collection of points of the space $\mathfrak{X}$. Any such point configuration is either finite or infinite. For our purposes it suffices to assume that the points of X are always pairwise distinct. The set of all point configurations in $\mathfrak{X}$ will be denoted as $\mathrm{Conf}(\mathfrak{X})$.

A relatively compact Borel subset $A \subset \mathfrak{X}$ is called *a window*. For a window A and $X \in \mathrm{Conf}(\mathfrak{X})$, set $N_A(X) = |A \cap X|$ (number of points of X in the window). Thus, N_A can be viewed as a function on $\mathrm{Conf}(\mathfrak{X})$. We equip $\mathrm{Conf}(\mathfrak{X})$ with the Borel structure generated by functions N_A for all windows A. That is, all these functions N_A are assumed measurable.

A *random point process* on $\mathfrak{X}$ is a probability measure on $\mathrm{Conf}(\mathfrak{X})$.

For algebraic purposes it is always enough to assume that the space $\mathfrak{X}$ is finite, and then $\mathrm{Conf}(\mathfrak{X}) = 2^{\mathfrak{X}}$ is simply the space of all subsets of $\mathfrak{X}$. The same space of configurations $\mathrm{Conf}(\mathfrak{X}) = 2^{\mathfrak{X}}$ may be taken for countable discrete $\mathfrak{X}$, then $\mathrm{Conf}(\mathfrak{X})$ is compact.

6.1.2 Correlation measures and correlation functions

Given a random point process, one can usually define a sequence $\{\rho_n\}_{n=1}^{\infty}$, where ρ_n is a symmetric measure on $\mathfrak{X}^n$ called the nth *correlation measure*. Under mild

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

conditions on the point process, the correlation measures exist and determine the process uniquely, cf. [Len73].

The correlation measures are characterized by the following property: For any $n \geq 1$ and a compactly supported bounded Borel function f on $\mathfrak{X}^n$ one has

$$\int_{\mathfrak{X}^n} f\rho_n = \left\langle \sum_{x_{i_1},\dots,x_{i_n}\in X} f(x_{i_1},\dots,x_{i_n}) \right\rangle_{X\in\mathrm{Conf}(\mathfrak{X})} \tag{6.1.1}$$

where the sum on the right is taken over all n-tuples of pairwise distinct points of the random point configuration X.

Often one has a natural measure μ on $\mathfrak{X}$ (called the *reference measure*) such that the correlation measures have densities with respect to $\mu^{\otimes n}$, $n = 1, 2, \dots$. Then the density of ρ_n is called the nth *correlation function* and it is usually denoted by the same symbol "ρ_n".

Often one has a natural measure μ on $\mathfrak{X}$ (called the *reference measure*) such that the correlation measures have densities with respect to $\mu^{\otimes n}$, $n = 1, 2, \dots$. Then the density of ρ_n is called the nth *correlation function* and it is usually denoted by the same symbol "ρ_n".

If $\mathfrak{X} \subset \mathbb{R}$ and μ is absolutely continuous with respect to the Lebesgue measure, then the probabilistic meaning of the nth correlation function is that of the density of probability to find an eigenvalue in each of the infinitesimal intervals around points $x_1, x_2, \dots x_n$:

$$\begin{aligned}\rho_n(x_1, x_2, \dots x_n)\mu(dx_1)\cdots\mu(dx_n)& \\ = \mathbb{P}\left\{\text{there is a particle in each interval } \ (x_i, x_i + dx_i)\right\}&.\end{aligned}$$

On the other hand, if μ is supported by a discrete set of points, then

$$\begin{aligned}\rho_n(x_1, x_2, \dots x_n)\mu(x_1)\cdots\mu(x_n)& \\ = \mathbb{P}\{\text{there is a particle at each of the points } \ x_i\}&.\end{aligned}$$

6.1.3 Determinantal processes

Assume that we are given a point process $\mathcal{P}$ and a reference measure such that all correlation functions exist. The process $\mathcal{P}$ is called *determinantal* if there exists a function $K : \mathfrak{X} \times \mathfrak{X} \to \mathbb{C}$ such that

$$\rho_n(x_1, \dots, x_n) = \det[K(x_i, x_j)]_{i,j=1}^n, \qquad n = 1, 2, \dots. \tag{6.1.2}$$

The function K is called a *correlation kernel* of $\mathcal{P}$.

The determinantal form of the correlation functions (6.1.2) implies that many natural observables for $\mathcal{P}$ can be expressed via the kernel K. The determinantal structure of a point process allows to study its asymptotics in various regimes, which lead to many interesting results.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

6.1.4 Remarks

Let us mention a few general remarks/facts about determinantal processes.

- **(kernel is not unique)**

  The correlation kernel $K(x,y)$ of a process is not defined uniquely. For example, transposition $K(x,y) \mapsto K(y,x)$, as well as 'gauge transformations' of the form

  $$K(x,y) \mapsto \frac{f(x)}{f(y)} K(x,y),$$

  with a nonvanishing function $f(x)$, do not change the correlation functions and hence the process. The change in the reference measure leads also to a gauge transformation of the above form.

- **(repelling property)**

  If the kernel $K(x,y)$ is Hermitean symmetric:

  $$K(x,y) = \overline{K(y,x)}, \qquad x, y \in \mathfrak{X}, \tag{6.1.3}$$

  then the point process possesses certain repelling property. Consider the first and the second correlation functions:

  $$\rho_1(x) = P(X \ni x) = K(x,x), \qquad \rho_2(x,y) = \rho_1(x)\rho_1(y) - |K(x,y)|^2 \le \rho_1(x)\rho_2(y). \tag{6.1.4}$$

  Points of a random configuration are often interpreted as random particles. We see that the probability to see particles at x and y simultaneously is not greater than the product of probabilities of the two separate events that (1) there is a particle at x and there is a particle at y. This can be understood as repelling of particles.

  This repelling property motivated an earlier name for the determinantal point processes — *fermionic processes.*

- **(processes with Hermitean symmetric kernel)**

  There exists a characterization of point processes with Hermitean symmetric kernels [Sos00]. Interpret K as an integral operator in $L^2(\mathfrak{X}, \mu)$ (where μ is the reference measure with respect to which the kernel is defined):

  $$(Kf)(x) := \int_{\mathfrak{X}} K(x,y) f(y) \mu(dy).$$

  The fact that the correlation functions are nonnegative means that $K \ge 0$, i.e., the operator K is nonnegative definite (all its diagonal minors which are the correlation functions are nonnegative). In fact, the operator $1 - K$ (here and below 1 is the identity operator) is also nonnegative definite.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Hermitian locally trace class[1] operator K in $L^2(\mathfrak{X}, \mu)$ defines a determinantal random point process if and only $0 \le K$ and $0 \le 1 - K$. If the corresponding random point process exists then it is unique.

A large subclass of processes with Hermitean symmetric kernel is formed by orthogonal polynomial ensembles whose kernels are finite-dimensional orthogonal projections. In fact, any Hermitean symmetric operator K with $0 \le K \le 1$ may be approximated by finite-dimensional projections.

- **(complementation principle)**

  If the space $\mathfrak{X}$ is discrete, there is an important operation on point processes which preserves the determinantal structure.

  For any subset $\mathfrak{Y} \subset \mathfrak{X}$ one can define an involution on point configurations $X \subset \mathfrak{X}$ by $X \mapsto X \triangle \mathfrak{Y}$ (here $\triangle$ is the symbol of symmetric difference). This map leaves intact the "particles" of X outside of $\mathfrak{Y}$, and inside $\mathfrak{Y}$ it picks the "holes" (points of $\mathfrak{Y}$ free of particles). This involution is called the *particle-hole involution* on $\mathfrak{Y}$.

  Given an arbitrary discrete state space $\mathfrak{X}$, a kernel $K(x, y)$ on $\mathfrak{X} \times \mathfrak{X}$, and a subset $\mathfrak{Y}$ of $\mathfrak{X}$, consider another kernel,

$$K^\circ(x, y) = \begin{cases} K(x, y), & x \notin \mathfrak{Y}, \\ \delta_{xy} - K(x, y), & x \in \mathfrak{Y}, \end{cases}$$

  where δ_{xy} is the Kronecker symbol.

  Proposition 6.1.1. *Let $\mathcal{P}$ be a determinantal point process with correlation kernel K on a discrete space $\mathfrak{X}$, and let $\mathcal{P}^\circ$ be the image of $\mathcal{P}$ under the particle–hole involution on $\mathfrak{Y} \subset \mathfrak{X}$. Then $\mathcal{P}^\circ$ is also a determinantal point process with correlation kernel $K^\circ(x, y)$ as defined above:*

$$\rho_m(x_1, \ldots, x_m \mid \mathcal{P}^\circ) = \det\left[K^\circ(x_i, x_j)\right]_{i,j=1}^m, \quad m = 1, 2, \ldots .$$

- **(killing)**

  Killing particles independently with some rate depending on location does not destroy the determinantal structure. This property is a relative of the above complementation principle.

6.1.5 Examples [Bor11b]

Poisson process.

Take $\mathfrak{X} = \mathbb{R}$. The (uniform) **Poisson process** is characterized by 2 properties:

[1] This means that the operator $K\chi_B$, where χ_B is an operator of restriction to a compact subset $B \subset \mathfrak{X}$, is trace class.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

- N_A is Poisson distributed for any window A as $P(N_A = k) = e^{-|A|}|A|^k/k!$.
- For any two disjoint windows A and B, N_A and N_B are independent (no interaction).

How to construct the Poisson process? Take $M \gg 0$, $M \in \mathbb{Z}$. Consider M i.i.d. uniform random variables on $[-M/2, M/2]$. Then there is approximately 1 point per unit length. For $M \to \infty$, the limit exists and it is the uniform Poisson process. The $M \to \infty$ picture is simpler than the case of M finite. For instance in the finite case, the probability for the number of points in a window is not as clean as an exponential function, and will be messy, but it limits to the exponential function. The standard definition is that the distances between neighboring particles are i.i.d. according to $e^{-x}\mathrm{d}x$.

The uniform Poisson process is determinantal with correlation kernel

$$K(x, y) = \delta_{xy}, \qquad x, y \in \mathbb{R}.$$

To define a non-uniform Poisson process with respect to a measure μ, we change the probabilities by $P(N_A = k) = e^{-\mu(A)}\mu(A)^k/k!$. This definition makes sense for any measure space.

Bernoulli process.

Take $\mathfrak{X} = \mathbb{Z}$ and $p \in [0, 1]$. The probability of containing a given particle is p, and the probability of not containing it is $1 - p$. Different locations are clearly independent.

The Bernoulli process is determinantal with correlation kernel

$$K(x, y) = p \cdot \delta_{xy}, \qquad x, y \in \mathbb{Z}.$$

Now embed $\mathbb{Z} \to \mathbb{R}$ via $m \mapsto pm$. With this embedding, as $p \to 0$, the Bernoulli process converges to uniform Poisson. Convergence means the following: for any set of windows $A_1, \ldots, A_m$, we have that $(N_{A_1}, \ldots, N_{A_m})$ converges in distribution.

6.2 Biorthogonal and orthogonal polynomial ensembles

Here we consider an important class of determinantal point processes which admit an explicit formula for the correlation kernel.

6.2.1 Biorthogonal ensembles [Bor11b]

Definition 6.2.1. Consider a state space $\mathfrak{X}$ with a reference measure μ. An N-point **biorthogonal ensemble** on $\mathfrak{X}$ is an N-point point process (= probability measure on $\mathfrak{X}^N$) of the form

$$P_N(\mathrm{d}x_1 \ldots \mathrm{d}x_n) = c_N \det(\varphi_i(x_j))_{i,j=1}^N \det(\psi_i(x_j))_{i,j=1}^N \mu(\mathrm{d}x_1) \cdots \mu(\mathrm{d}x_n)$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

where c_N is some constant and φ_i, ψ_i are arbitrary functions on $\mathfrak{X}$.

Proposition 6.2.2. *Any biorthogonal ensemble is a determinantal point process. Its correlation kernel has the form*

$$K(x,y) = \sum_{i,j=1}^{N} (G^{-T})_{i,j}\varphi_i(x)\psi_j(y)$$

where $G_{i,j} = \int_{\mathfrak{X}} \varphi_i(x)\psi_j(x)\mu(\mathrm{d}x)$ is the Gram matrix. The matrix G^{-T} is the inverse transposed matrix.

This was first considered by F. Dyson in 1962.

Proof. First let us obtain an expression for the normalizing constant.

$$\begin{aligned}\int_{\mathfrak{X}^N} \det(\varphi_i(x_j))_{i,j=1}^N \det(\psi_i(x_j))_{i,j=1}^N \mathrm{d}x &= \int_{\mathfrak{X}^N} \sum_{\sigma,\tau\in S_N} \mathrm{sign}(\sigma\tau) \prod_{i=1}^N \varphi_{\sigma(i)}(x_i)\psi_{\tau(i)}(x_i)\mathrm{d}x \\ &= \sum_{\sigma,\tau\in S_N} \prod_{i=1}^N G_{\sigma(i),\tau(i)} \\ &= N! \sum_{\rho\in S_N} \mathrm{sign}(\sigma) \prod_{i=1}^N G_{i,\rho(i)} = N! \det(G_{i,j})_{i,j=1}^N.\end{aligned}$$

This implies that G is invertible, and we will have $c_N = (N!\det(G))^{-1}$. Now we have[2]

$$\rho_n(x_1,\ldots,x_n) = \frac{N!}{(N-n)!}\int_{(x_{n+1},\ldots,x_N)\in\mathfrak{X}^{N-n}} \det(\varphi_i(x_j))_{i,j=1}^N \det(\psi_i(x_j))_{i,j=1}^N \mathrm{d}x \frac{1}{N!\det(G)}.$$

Take matrices A and B such that $AGB^T = 1$. Set

$$\Phi_k = \sum_{\ell=1}^N A_{k\ell}\varphi_\ell, \quad \Psi_k = \sum_{\ell=1}^N B_{k\ell}\psi_\ell.$$

Then we have

$$\langle \Phi_i, \Psi_j\rangle := \int_{\mathfrak{X}} \Phi_i(x)\Psi_j(x)\mathrm{d}x = [AGB^T]_{i,j} = 1_N.$$

Also we have

$$\begin{aligned}\det(\Phi_i(x_j))_{i,j=1}^N &= \det(A)\det(\varphi_i(x_j))_{i,j=1}^N, \\ \det(\Psi_i(x_j))_{i,j=1}^N &= \det(B)\det(\psi_i(x_j))_{i,j=1}^N.\end{aligned}$$

[2]This is a straightforward way to compute correlation functions of a N-point point process by integrating out several variables. The factor $N!/(N-n)!$ has a combinatorial nature.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

The result is that

$$P_N(\mathrm{d}x_1 \cdots \mathrm{d}x_N) = \frac{1}{N!} \det(\Phi_i(x_j))_{i,j=1}^N \det(\Psi_i(x_j))_{i,j=1}^N,$$

and

$$\begin{aligned}
\rho_n(x_1, \ldots, x_n) &= \frac{1}{(N-n)!} \int_{x_1,\ldots,x_n} \det(\varphi_i(x_j))_{i,j=1}^N \det(\psi_i(x_j))_{i,j=1}^N \mathrm{d}x \\
&= \frac{1}{(N-n)!} \int_{x_{n+1},\ldots,x_N} \sum_{\sigma,\tau \in S_N} \operatorname{sign}(\sigma\tau) \prod_{i=1}^N \Phi_{\sigma(i)}(x_i)\Psi_{\tau(i)}(x_i)\mathrm{d}x_i \\
&= \frac{1}{(N-n)!} \int_{x_{n+1},\ldots,x_N} \sum_{\substack{\sigma,\tau \in S_N \\ \sigma(k)=\tau(k) \\ \text{for } k=n+1,\ldots,N}} \operatorname{sign}(\sigma\tau) \prod_{i=1}^n \Phi_{\sigma(i)}(x_i)\Psi_{\tau(i)}(x_i)\mathrm{d}x_i \\
&= \sum_{1 \le j_1 < \cdots < j_n \le N} \det \Phi^{j_1,\ldots,j_n} \det \Psi^{j_1,\ldots,j_n},
\end{aligned}$$

where $\Phi^{j_1,\ldots,j_n}$ is the submatrix of $[\varphi_i(x_j)]$ with columns $j_1, \ldots, j_n$ and rows $i = 1, \ldots, n$. Now using the Cauchy–Binet theorem, this last expression becomes

$$\det(\Phi\Psi^T)_{i,j=1}^n = \det\left(\sum_{k=1}^N \Phi_k(x_i)\Psi_k(j)\right)_{i,j=1}^N.$$

Now define $K(x_i, x_j) = \sum_{k=1}^N \Phi_k(x_i)\Psi_k(x_j)$. We can write this as

$$\begin{aligned}
\sum_{k=1}^N \Phi_k(x)\Psi_k(y) &= \sum_{k,\ell,m} A_{k\ell}\varphi_\ell(x)B_{km}\psi_m(y) \\
&= \sum_{\ell,m} \varphi_\ell(x)\psi_m(y) \sum_k A_{k\ell}B_{km} = A^T B = G^{-T}. \qquad \square
\end{aligned}$$

Remark 6.2.3. In fact, there are more general point processes given by products of determinants whose correlation kernels admit explicit expressions. They are close relatives of the biorthogonal ensembles but have a certain "time" dependence. See [Bor11a, §4].

6.2.2 Orthogonal polynomial ensembles [Bor11b]

Take $\mathfrak{X} = \mathbb{R}$. Let $w(\mathrm{d}x)$ be a positive measure on $\mathbb{R}$ with finite moments, i.e., $\int_{\mathbb{R}} |x|^k w(\mathrm{d}x) < \infty$ for all $k \ge 0$.

Example 6.2.4. The natural map $\mathbb{C}[x]_{\le N} \to L^2(\mathbb{R}, w(\mathrm{d}x))$ is an embedding if and only if $\#\mathrm{supp}(w) > N + 1$.

We will assume that $\#\mathrm{supp}(w) = \infty$.

Notation: $V_N(x_1, \ldots, x_n) = \prod_{1 \le i < j \le N}(x_i - x_j) = \det(x_i^{N-j})_{i,j=1}^N$.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Definition 6.2.5. The N-**particle orthogonal polynomial ensemble** with weight w is the N-point random point process with joint probability density

$$P_N(\mathrm{d}x_1 \cdots \mathrm{d}x_N) = (V_N(x_1, \ldots, x_N))^2 \prod_{i=1}^{N} w(\mathrm{d}x_i).$$

Example 6.2.6. The most well-known example is the **Gaussian unitary ensemble** (GUE(N)). In this case the space is $\{H \in \mathrm{Mat}_N(\mathbb{C}) \mid H = H^*\}$ with measure $ce^{-\operatorname{Tr}(H^2)}$ for some constant c. The eigenvalues of H form an N-point orthogonal polynomial ensemble with $w(\mathrm{d}x) = e^{-x^2}\mathrm{d}x$.

An orthogonal polynomial ensemble is a biorthogonal ensemble with $\varphi_i(x) = \psi_i(x) = x^{i-1}\sqrt{w(x)}$ where w denotes the density function of the measure w. A kernel K defines a linear operator $K\colon L^2 \to L^2$ by $(Kf)(x) = \int K(x,y)f(y)\mathrm{d}y$.

Proposition 6.2.7. *The correlation kernel $K(x,y)$ is the kernel of the orthogonal projection operator onto* $\operatorname{span}(\sqrt{w(x)}, x\sqrt{w(x)}, \ldots, x^{N-1}\sqrt{w(x)})$ *in* $L^2(\mathbb{R}, w)$.

Proof. Let $U = \operatorname{span}(\varphi_j)_{j=1}^N$ and $V = \operatorname{span}(\psi_j)_{j=1}^N$ in $L^2(\mathbb{R}, w)$. We have

$$\int \sum_{i,j=1}^{N} (G^{-T})_{i,j}\varphi_i(x)\psi_j(y)\varphi_m(y)\mathrm{d}y = \sum_{i,j=1}^{N} (G^{-T})_{i,j} G^T_{mj}\varphi_i(x) = \sum_{i=1}^{N} 1_{i,m}\varphi_i(x) = \varphi_m(x).$$

Hence $K|_U = 1_U$. On the other hand $K|_{V^\perp} = 0$. Also, one can check that $K^2 = K$. So K is a projection onto U which is parallel to $V^\perp$. In the orthogonal polynomial ensemble, we have $U = V$. □

Definition 6.2.8. A **system of orthogonal polynomials** on $\mathbb{R}$ with weight w is a sequence $\{p_n(x)\}_{n\geq 0}$ with $p_n \in \mathbb{C}[x]$ and $\deg p_n = n$, such that $p_n \perp \mathbb{C}[x]_{\leq n-1}$ in $L^2(\mathbb{R}, w)$, i.e.,

$$\int_{\mathbb{R}} p_n(x)p_m(x)w(\mathrm{d}x) = \|p_n\|^2\delta_{m,n}.$$

Note that a system of orthogonal polynomials is an orthogonal basis in $\mathbb{C}[x]$ with inner product $\langle f, g\rangle = \int_{\mathbb{R}} f(x)g(x)w(\mathrm{d}x)$, so one can construct such systems using the Gram–Schmidt orthogonalization algorithm. Note that the degree constraint $\deg p_n = n$ uniquely determines the p_n up to a constant.

Standard notation: Let k_n be the leading coefficient of p_n and set $h_n = \|p_n\|^2_{L^2(\mathbb{R},w)}$.

Proposition 6.2.9. *Let $\{p_n\}$ be the sequence of monic orthogonal polynomials with weight w. Then the correlation kernel of the N-point orthogonal polynomial ensemble has the form*

$$K_N(x,y) = \sum_{j=0}^{N-1} \frac{p_j(x)p_j(y)}{h_j}$$

with respect to the reference measure $w(\mathrm{d}x)$ on $\mathbb{R}$.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Proof. Let π_{j-1} be monic polynomials of degree $j-1$. Then

$$\begin{aligned} P_N(\mathrm{d}x_1 \cdots \mathrm{d}x_N) &= c \det(x_i^{j-1}) \det(x_i^{j-1}) \prod w(\mathrm{d}x_i) \\ &= \det(\pi_{j-1}(x_i)) \det(\pi_{j-1}(x_i)) \prod w(\mathrm{d}x_i), \end{aligned}$$

where the last equality is via row operations. Then $G^{-T} = \operatorname{diag}(h_0^{-1}, \ldots, h_{N-1}^{-1})$. □

Proposition 6.2.10 (Christoffel–Darboux)**.**

$$\sum_{j=0}^{N-1} \frac{p_j(x)p_j(y)}{h_j} = \frac{p_N(x)p_{N-1}(y) - p_{N-1}(x)p_N(y)}{h_{N-1}(x-y)}.$$

Proof 1. Consider the operator given by multiplication by x in $\mathbb{R}[x]$. This operator is self-adjoint:

$$(x \cdot f, g)_{L^2(\mathbb{R},\omega)} = \int x f(x) g(x) w(\mathrm{d}x) = (f, x \cdot g)_{L^2(\mathbb{R},\omega)}.$$

The matrix of a self-adjoint operator in any orthonormal basis is symmetric. For our basis, we will use $\{p_n/\|p_n\|\}_{n\geq 0}$. By degree considerations, this matrix must be 0 below the subdiagonal. By symmetry, it must be 0 above the superdiagonal. Since the p_n are monic, we have

$$x p_n = A_{n,n+1} p_{n+1} + A_{n,n} p_n + A_{n,n-1} p_{n-1}$$

where $A_{n,n+1} = 1$. Now multiply the desired identity by $(x-y)$ and use this recurrence relation and symmetry of the matrix to finish (the left-hand side is a telescoping sum). □

Proof 2. Consider the average ($x_1, \ldots, x_N$ distributed as the orthogonal polynomial ensemble). Then

$$\mathbb{E}\Big(\prod_{i=1}^N (u-x_i)(v-x_i)\Big) = \text{constant} \cdot \int \prod_{i=1}^N (u-x_i)(v-x_i) \prod_{i<j} (x_i - x_j)^2 w(\mathrm{d}x_1) \cdots w(\mathrm{d}x_N).$$

(If the x_i are eigenvalues of a random matrix X, then $\prod_i (u-x_i) = \det(u-X)$.)

In the simpler case, we have

$$\begin{aligned} \mathbb{E}\Big(\prod_{i=1}^N (u - x_i)\Big) &= \text{constant} \cdot \int \prod_{i=1}^N (u-x_i) \prod_{i<j} (x_i - x_j)^2 w(\mathrm{d}x_1) \cdots w(\mathrm{d}x_N) \\ &= \text{constant} \cdot \int V_{N+1}(u, x_1, \ldots, x_N) V_N(x_1, \ldots, x_N) w(\mathrm{d}x_1) \cdots w(\mathrm{d}x_N) \\ &= \text{constant} \cdot \int \det \begin{bmatrix} p_N(u) & p_N(x_1) & \cdots & p_N(x_N) \\ p_{N-1}(u^{N-1}) & p_{N-1}(x_1^{N-1}) & \cdots & p_{N-1}(x_N^{N-1}) \\ \vdots & & & \\ 1 & 1 & \cdots & 1 \end{bmatrix}. \end{aligned}$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

$$\det \begin{bmatrix} p_{N-1}(x_1^{N-1}) & \cdots & p_{N-1}(x_N^{N-1}) \\ \vdots & & \\ 1 & \cdots & 1 \end{bmatrix} w(\mathrm{d}x_1)\cdots w(\mathrm{d}x_N)$$
$$= \text{constant} \cdot p_N(u) \int V_N(x_1,\ldots,x_N)^2 w(\mathrm{d}x_1)\cdots w(\mathrm{d}x_N) = p_N(u).$$

(In the third equality, we have done row operations.)

Using this, we get

$$\mathbb{E}(\prod_{i=1}^{N}(u-x_i)(v-x_i)) = \text{constant} \cdot \int V_{N+1}(u,x_1,\ldots,x_N)V_{N+1}(v,x_1,\ldots,x_N)w(\mathrm{d}x_1)\cdots w(\mathrm{d}x_N)$$
$$= \frac{\text{constant}}{u-v}\int V_{N+2}(u,v,x)V_N(x)w(\mathrm{d}x_1)\cdots w(\mathrm{d}x_N)$$

The first integral expression can be simplified as $\sum_{k=0}^{N} c_k p_k(u)p_k(v)$ for some coefficients c_k. The second integral expression can be simplified as $\frac{1}{u-v}\det\begin{bmatrix} p_{N+1}(u) & p_{N+1}(v) \\ p_N(u) & p_N(v)\end{bmatrix}$.

□

6.3 Examples: CUE, GUE, extreme characters of $U(\infty)$

We will consider three examples of determinantal point processes. The first two are of random-matrix nature, and the third ensemble is related to our asymptotic representation-theoretic constructions.

6.3.1 CUE

Dyson's circular unitary ensemble (CUE) was introduced by F. Dyson in 1962. This is a random matrix distribution, that is, a certain probability measure on some space of matrices.

The CUE is defined as the Haar measure on the set of unitary matrices $U(N)$ of size $N \times N$. We are interested in eigenvalues of these matrices. The random eigenvalues form a random N-point process on $\mathbb{T}^1$. As we observed in §4.1.5, the distribution of eigenvalues is

$$P_N(du_1,\ldots,du_N) = c_N \prod_{1\le i<j\le N} |u_i-u_j|^2 d\theta_1 \ldots d\theta_N,$$

where $d\theta_1$ denote the Lebesgue measure on the unit circle.

6.3.2 GUE

The *Gaussian Unitary Ensemble* (GUE) is a Gaussian measure on the space of $N \times N$ Hermitean matrices which is invariant under unitary rotations $H \mapsto$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

UHU^{-1}, $U \in U(N)$. The distribution is described often as

$$P(dH) = \frac{1}{Z} e^{-\frac{N}{2} \cdot \mathrm{Tr}(H^2)}.$$

The random eigenvalues of a GUE matrices are real and distinct with probability one, and they have density

$$P_N(d\lambda_1, \ldots, d\lambda_N) = c_N \prod_{1 \le i < j \le N} |\lambda_i - \lambda_j|^2 \prod_{i=1}^{N} e^{-N\lambda_i^2/2} d\lambda_i.$$

This random point process is an orthogonal polynomial ensemble, and its correlation kernel is expressed through the Hermite polynomials (orthogonal polynomials with respect to the weight $e^{-Nx^2/2}dx$ on $\mathbb{R}$).

6.3.3 Extreme coherent systems on the Gelfand-Tsetlin graph

Let M_N be a measure on $\mathbb{GT}_N$, the set of all signatures of length N, corresponding to a fixed extreme character of $U(\infty)$. That is, explicitly,

$$M_N(\lambda) = \mathrm{Dim}_N \lambda \cdot \det[\varphi_{\lambda_i - i + j}]_{i,j=1}^{N},$$

where φ_n are the Laurent coefficients of the function

$$\Psi_{\alpha^\pm, \beta^\pm, \gamma^\pm}(u) = e^{\gamma^+(u-1) + \gamma^-(u^{-1}-1)} \prod_{i=1}^{\infty} \left(\frac{1 + \beta_i^+(u-1)}{1 - \alpha_i^+(u-1)} \cdot \frac{1 + \beta_i^-(u^{-1}-1)}{1 - \alpha_i^-(u^{-1}-1)} \right)$$
$$= \sum_{n \in \mathbb{Z}} \varphi_n u^n.$$

(Where $\alpha^\pm, \beta^\pm, \gamma^\pm$ are fixed parameters of the extreme character.) Since

$$\mathrm{Dim}_N \lambda = c_N \det[(\lambda_i - i)^{N-j}]_{i,j=1}^{N},$$

we see that for every fixed N, the measure M_N viewed as a point process with configurations

$$\{\lambda_1 - 1 > \lambda_2 - 2 > \ldots > \lambda_N - N\} \subset \mathbb{Z}$$

(all points in these configurations are distinct), is a biorthogonal ensemble. In fact, one can compute the correlation kernel of such a measure explicitly, cf. [BK08]. In our lectures we consider a simpler case, when $\alpha^\pm = 0$, $\beta^\pm = 0$, $\gamma^- = 0$, and the only nonzero parameter is $\gamma^+ = \gamma > 0$.

Thus, we have

$$\Psi(u) = e^{\gamma(u-1)}. \tag{6.3.1}$$

Restricting to $U(N)$, we obtain

$$\chi(U) = \prod_{k=1}^{N} \Psi(u_k), \qquad U \in U(N), \tag{6.3.2}$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

where $\{u_1, \dots, u_k\}$ are eigenvalues of a matrix U.

Decomposing this function as a linear combination of Schur polynomials, we get:

Proposition 6.3.1. *We have*

$$\prod_{k=1}^{N} e^{\gamma(u_k-1)} = \sum_{\lambda\in\mathbb{GT}_N} \widetilde{c}_\lambda \frac{s_\lambda(u_1,\dots,u_N)}{\operatorname{Dim}_N \lambda}, \tag{6.3.3}$$

where the coefficients $\widetilde{c}_\lambda$ have the form (here and below we use the notation $l_i = \lambda_i + N - i$)

$$\widetilde{c}_\lambda = \frac{e^{-N\gamma}\gamma^{-\frac{N(N-1)}{2}}}{1!2!\dots(N-1)!} \cdot \prod_{i=1}^{N} \frac{\gamma^{l_i}}{l_i!} \cdot \prod_{1\le i<j\le N} (l_i - l_j)^2. \tag{6.3.4}$$

If any of l_i is negative, then $\widetilde{c}_\lambda = 0$.

Proof. We can write linear combination of usual Schur polynomials:

$$\prod_{k=1}^{N} e^{\gamma(u_k-1)} = \sum_{\lambda\in\mathbb{GT}_N} c_\lambda s_\lambda(u_1,\dots,u_N), \tag{6.3.5}$$

the coefficients $\widetilde{c}_\lambda$ and c_λ are related as

$$\widetilde{c}_\lambda = c_\lambda \cdot \frac{\prod_{1\le i<j\le N}(l_i - l_j)}{0!1!2!\dots(N-1)!}. \tag{6.3.6}$$

We have

$$c_\lambda = \det\left[\frac{\gamma^{\lambda_i - i + j}}{(\lambda_i - i + j)!} e^{-\gamma}\right]_{i,j=1}^{N} = e^{-\gamma N}\gamma^{\sum l_i}\gamma^{-\frac{N(N-1)}{2}} \cdot \det\left[\frac{1}{(\lambda_i - i + j)!}\right]_{i,j=1}^{N}. \tag{6.3.7}$$

It remains to compute the last determinant. Multiply each its ith row by $l_i! = (\lambda_i + N - i)!$. We get

$$\det\left[\frac{1}{(\lambda_i - i + j)!}\right]_{i,j=1}^{N} = \frac{1}{\prod l_i!} \cdot \det\left[l_i^{\downarrow(N-j)}\right]_{i,j=1}^{N}. \tag{6.3.8}$$

Here $l_i^{\downarrow(N-j)} = l_i(l_i-1)\dots(l_i-N+j+1)$ is the falling factorial power. Applying column transformations, we may replace the factorial powers in $\det[l_i^{\downarrow(N-j)}]$ by the usual powers. Thus, we arrive at a Vandermonde determinant $\prod_{1\le i<j\le N}(l_i - l_j)$. Putting all formulas together we conclude the proof, □

Thus, the measure M_N^γ with this gamma-parameter is an orthogonal polynomial ensemble. Its correlation kernel may be expressed through the classical Charlier orthogonal polynomials which are orthogonal on $\mathbb{Z}_{\ge 0}$ with the Poisson weight $e^{-\gamma}\frac{\gamma^x}{x!}$, $x = 0, 1, 2, \dots$.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Connection to Schur-Weyl duality

Let us state (without proof) a nice interpretation of the measure M_N^γ in terms of the Schur-Weyl duality between representations of the unitary $U(N)$ and the symmetric $\mathfrak{S}(n)$ groups.

Let $V = \mathbb{C}^N$, consider the nth tensor power $V^{\otimes n}$. The two groups $U(N)$ and $\mathfrak{S}(n)$ act in this space:

- The symmetric group acts by permutations of components. More precisely, a permutation $s \in \mathfrak{S}(n)$ acts on a decomposable tensor $v_1 \otimes \cdots \otimes v_n$ as

$$s(v_1 \otimes \cdots \otimes v_n) = v_{s^{-1}(1)} \otimes \cdots \otimes v_{s^{-1}(n)}.$$

- A matrix $X \in U(N)$ acts in $V^{\otimes n}$ as $X^{\otimes n}$ (component-wise).

These actions commute with each other.

The space $V^{\otimes n}$ can be decomposed into irreducible representations of $U(N)$ or $\mathfrak{S}(n)$, or, as a bimodule, as follows:

$$V^{\otimes n} = \bigoplus_\lambda V_\lambda^{\mathfrak{S}(n)} \otimes V_\lambda^{U(N)}. \tag{6.3.9}$$

The sum is taken over all Young diagrams λ with n boxes and $\le N$ rows. Here $V_\lambda^{\mathfrak{S}(n)}$ and $V_\lambda^{U(N)}$ denote the irreducible representations of the corresponding groups.

At the level of characters of $U(N)$, this decomposition gives

$$(u_1 + \ldots + u_N)^n = \sum_\lambda \dim\lambda \cdot s_\lambda(u_1, \ldots, u_N).$$

We already know this identity as a corollary of the Pieri rule (Corollary 4.2.11).

Further specializing to dimensions, we have

$$N^n = \sum_\lambda \dim\lambda \cdot \mathrm{Dim}_N\, \lambda.$$

This identity allows to define the probability measure on the set of all Young diagrams λ with n boxes and $\le N$ rows:

$$M_{N,n}(\lambda) := \frac{\dim\lambda \cdot \mathrm{Dim}_N\, \lambda}{N^n}.$$

The measure M_N^γ discussed in this subsection can be obtained as a *poissonization* of the measures $M_{N,n}$:

Proposition 6.3.2. *Let $\gamma > 0$. Then we have*

$$M_N^\gamma(\lambda) = e^{-\gamma N} \frac{(\gamma N)^{|\lambda|}}{|\lambda|!} M_{N,|\lambda|}(\lambda)$$

for all Young diagrams λ with $\le N$ rows.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Proof. Indeed, we have

$$\dim \lambda = |\lambda|! \prod_{1 \le i < j \le N} (\lambda_i - i - \lambda_j + j) \frac{1}{\prod_{i=1}^{N} (\lambda_i + N - i)!}.$$

This — together with the formula for $M_N^{(\gamma)}$ obtained above — implies the desired identity. ☐

In other words, to sample a Young diagram according to the distribution M_N^{γ}, one first samples a Poisson random variable n with parameter $N\gamma$, and then a diagram λ according to the measure $M_{N,n}$. (Equivalently: M_N^{γ} may be viewed as an ensemble $M_{N,n}$ with *random* n.)

Remark 6.3.3. The measure $M_N^{(\gamma)}$ coming from the asymptotic representation theory of the infinite-dimensional unitary group is itself a determinantal point process — Charlier orthogonal ensemble with parameter γ. This measure may be also viewed as a Poisson mixture of the Schur-Weyl measures $M_{N,n}$ with another parameter γN.

The measures $M_{N,n}$ in fact come from asymptotic representation theory of the infinite symmetric group, they form an extreme coherent system corresponding to parameters

$$\alpha = \Big(\underbrace{\frac{1}{N}, \dots, \frac{1}{N}}_{N}, 0, 0, \dots \Big), \qquad \beta = 0.$$

6.4 References

General definitions, facts and theorems (as well as examples) on determinantal point processes may be found in the surveys [Sos00], [HKPV06], [Bor11a].

Chapter 7

Asymptotics of determinantal point processes

The determinantal structure is especially useful when dealing with asymptotics of point processes. We will consider several aspects and methods, and apply them to simplest possible models. Our models have representation-theoretic origin: they arise from extreme characters of unitary and symmetric groups.

7.1 Difference/differential operators

7.1.1 Gamma-measures and formula for the kernel

To be concrete, we will first discuss the gamma-extreme coherent measures $M_N^{(\gamma)}$ arising from the asymptotic representation theory of unitary groups. Recall that these measures live on Young diagrams with $\leq N$ rows, and in shifted coordinates $l_i = \lambda_i + N - i$ the measure $M_N^{(\gamma)}$ is an orthogonal polynomial ensemble:

$$M_N^{(\gamma)}(l) = c_N \prod_{1 \leq i < j \leq N} |l_i - l_j|^2 \prod_{i=1}^{N} \pi_\gamma(l_i),$$

where π_γ is the Poisson distribution

$$\pi_\gamma(x) = e^{-\gamma} \frac{\gamma^x}{x!}, \qquad x = 0, 1, 2, \ldots.$$

This makes $M_N^{(\gamma)}$ a determinantal point process. Recall the formula for the correlation kernel of this process.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Consider the weighted space $\ell^2(\mathbb{Z}_{\ge 0}, \pi_\gamma)$ (corresponding to reference measure π_γ). By $\ell^2(\mathbb{Z}_{\ge 0})$ denote the un-weighted space ℓ^2 (corresponding to the counting reference measure). A obvious basis in this space is formed by the functions

$$\varepsilon_x(k) := \begin{cases} 1, & k = x; \\ 0, & k \ne x. \end{cases}$$

The map

$$\ell^2(\mathbb{Z}_{\ge 0}, \pi_\gamma) \to \ell^2(\mathbb{Z}_{\ge 0}), \qquad f \mapsto f\sqrt{\pi_\gamma}$$

is an isometry.

The algebra of polynomials $\mathbb{R}[x]$ belongs to $\ell^2(\mathbb{Z}_{\ge 0}, \pi_\gamma)$ because the measure π_γ has finite moments of all orders:

$$\sum_{x \in \mathbb{Z}_{\ge 0}} x^k \pi_\gamma(x) < \infty, \qquad k = 1, 2, \dots .$$

Moreover, one can show (EX) that this algebra is dense in $\ell^2(\mathbb{Z}_{\ge 0}, \pi_\gamma)$.

Thus, $\mathbb{R}[x]\sqrt{\pi_\gamma} \subset \ell^2(\mathbb{Z}_{\ge 0})$.

Remark 7.1.1. We use the isometry because we want a more invariant description of the kernel and other objects, i.e., not depending on the weight π_γ. We want this because this would allow to take limits of processes involving scaling of the parameter γ.

Let Q be the operator of orthogonal projection to the subspace

$$\operatorname{span}\left\{1, x, \dots, x^{N-1}\right\} \sqrt{\pi_\gamma} \subset \ell^2(\mathbb{Z}_{\ge 0}).$$

The correlation kernel is given by the formula

$$K(x, y) = (Q\varepsilon_x, \varepsilon_y)_{\ell^2(\mathbb{Z}_{\ge 0})}, \qquad x, y \in \mathbb{Z}_{\ge 0}$$

That is, $K(x, y)$ is the matrix of the projection operator Q in the standard basis $\{\varepsilon_x\}_{x \in \mathbb{Z}_{\ge 0}}$ of $\ell^2(\mathbb{Z}_{\ge 0})$.

If $p_n(x)$, $n = 0, 1, 2, \dots$ are the orthogonal polynomials with the weight function π_γ (they have the name *Charlier polynomials*), then the formula for the correlation kernel with the counting reference measure is

$$K(x, y) = \sum_{n=0}^{N-1} \frac{p_n(x)p_n(y)}{\|p_n\|^2_{\ell^2(\mathbb{Z}_{\ge 0}, \pi_\gamma)}} \sqrt{\pi_\gamma(x)\pi_\gamma(y)}. \tag{7.1.1}$$

7.1.2 Rodrigue's formula

In more detail on orthogonal polynomials (and especially discrete orthogonal polynomials) see the book [NSU91].

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Consider the elementary difference operators (forward and backward discrete derivatives) on functions on $\mathbb{Z}$

$$\Delta f(x) := f(x+1) - f(x), \qquad \nabla f(x) := f(x) - f(x-1).$$

Orthogonal polynomials $p_0(x) = 1, p_1(x), p_2(x), \dots$ with weight $\pi_\gamma(x)$ are uniquely characterized (up to a constant factor) by the conditions:

$$(p_n, f)_{\ell^2(\mathbb{Z}_{\geq 0}, \pi_\gamma)} = 0 \quad \text{for all polynomials } f \text{ with } \deg f \leq n-1, \quad n = 1, 2, \dots. \tag{7.1.2}$$

Proposition 7.1.2 (Rodrigue's formula)**.** *The orthogonal polynomials admit a representation*[1]

$$p_n = \frac{\nabla(\pi_\gamma p_{n-1})}{\pi_\gamma},$$

or

$$p_n = \left(\pi_\gamma^{-1} \circ \nabla \circ \pi_\gamma\right) p_{n-1},$$

where by π_γ we mean the operator of multiplication by the function $\pi_\gamma(x) = e^{-\gamma}\frac{\gamma^x}{x!}$.

Applying the above formulas several times we have

$$p_n = \frac{\nabla^n \pi_\gamma}{\pi_\gamma}, \qquad n = 1, 2, 3, \dots.$$

Proof. We show that the polynomials p_n defined as $\frac{\nabla^n \pi_\gamma}{\pi_\gamma}$, satisfy the orthogonality relations (7.1.2).

By induction we see that $\deg p_n = n$. Take polynomial g on $\mathbb{Z}_{\geq}$ with $\deg g \leq n-1$. Write

$$(p_n, g)_{\ell^2(\mathfrak{X}, W)} = \sum_{x \in \mathbb{Z}} p_n(x) g(x) W(x) = \sum_{x \in \mathbb{Z}} \nabla\left(W p_{n-1}\right)(x) g(x).$$

Because f decays fast on $+\infty$ and is zero on the left (i.e., at $x = -1, -2, \dots$) we may use the following summation by parts formula (EX):

$$\sum_{x \in \mathbb{Z}} (\nabla f)(x) g(x) = -\sum_{x \in \mathbb{Z}} f(x) \left(\Delta g\right)(x).$$

We have

$$\sum_{x \in \mathbb{Z}} \nabla\left(W p_{n-1}\right)(x) g(x) = \sum_{x \in \mathbb{Z}} W(x) p_{n-1}(x) (\Delta g)(x) = 0$$

by induction step because the operation Δ lowers the degree of the polynomial. □

[1] We take our concrete Poisson discrete weight, but the statements we make hold in more general situations as well.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

7.1.3 Charlier polynomials

From the Rodrigue's formula for the Poisson weight

$$p_n(x) = p_{n-1}(x) - \frac{x}{\gamma} p_{n-1}(x-1), \qquad p_0(x) = 1.$$

one can readily get an explicit formula for the Charlier polynomials p_n which are orthogonal polynomials on $\mathbb{Z}_{\geq 0}$ with Poisson weight π_γ. We have

$$p_n(x) = \sum_{m=0}^{n} \frac{n^{\downarrow m}}{(-\theta)^m m!} x^{\downarrow m}.$$

Here $a^{\downarrow m} := a(a-1)\dots(a-m+1)$ is the falling factorial power. Note that in fact the summation above is taken from 0 to the minimum of x and n because of these falling factorial powers, and so we conclude that

$$p_n(x) = p_x(n),$$

a nice self-duality property.

Proposition 7.1.3 (difference operator)**.** *We have*

$$\gamma p_n(x+1) + x p_n(x-1) - (x+\gamma) p_n(x) = -n p_n(x).$$

In other words, the Charlier polynomials are eigenfunctions of the second order difference operator

$$(Df)(x) = \gamma f(x+1) + x f(x-1) - (x+\gamma) f(x)$$

with eigenvalues $0, -1, -2, \dots$.

Proof. One can check the relation for the explicit formula for the Charlier polynomials in a straightforward way.

A more general argument shown the existence of such a difference operator. The Charlier polynomials[2] must satisfy three-term relations — linear relations on p_{n-1}, p_n, and p_{n+1}. This is shown as follows. Consider an operator of multiplication by x in the basis $\{p_n\}$:

$$x p_n(x) = p_{n+1}(x) + a_n p_n(x) + b_n p_{n-1}(x) + \text{terms with } p_{n-2}, p_{n-3}, \dots$$

(if the orthogonal polynomials are monic, then the coefficient by $p_{n+1}(x)$ is one). But the operator of multiplication by x is symmetric, so the terms with $p_{n-2}, p_{n-3}, \dots$ do not appear. Hence we have some three-term relations. Thus, the existence of such relations is proven. □

Under the isometry of $\ell^2(\mathbb{Z}_{\geq 0}, \pi_\gamma)$ with $\ell^2(\mathbb{Z}_{\geq 0})$, the above difference operator D maps to $\pi_\gamma^{\frac{1}{2}} \circ D \circ \pi_\gamma^{-\frac{1}{2}}$. Thus,

[2]Like a large class of orthogonal polynomials.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Proposition 7.1.4. *The functions $p_n\sqrt{\pi_\gamma} \in \ell^2(\mathbb{Z}_{\geq 0})$ are eigenfunctions with eigenvalues $0, -1, -2, \dots$ of the second order difference operator in $\ell^2(\mathbb{Z}_{\geq 0})$:*

$$(D_\gamma\varphi)(x) = \sqrt{\gamma(x+1)}\varphi(x+1) + \sqrt{\gamma x}\varphi(x-1) - (x+\gamma)\varphi(x),$$

where $\varphi \in \ell^2(\mathbb{Z}_{\geq 0})$.

Then operator D_γ may be viewed also as an operator in $\ell^2(\mathbb{Z})$ because of the factor $\sqrt{\gamma x}$. Thus, the operator does not map a function supported on $\mathbb{Z}_{\geq 0}$ into a function supported on a larger subset.

Thus, we have established our main interpretation of the correlation kernel:

Theorem 7.1.5. *The correlation kernel $K(x, y)$ (7.1.1) of the measure $M_N^{(\gamma)}$ is the matrix of the orthogonal projection operator onto the subspace $\{p_i\sqrt{\pi_\gamma}\}_{i=0}^{N-1} \subset \ell^2(\mathbb{Z}_{\geq 0})$.*

7.1.4 Plancherel measures

As a first application of the technique of difference/differential operators, let us consider the Plancherel measure on partitions.

The point process $M_N^{(\gamma)}$ is a Poisson mixture (with parameter γN) of the measures coming from the Schur-Weyl duality

$$M_{N,n}(\lambda) = \frac{\dim\lambda \cdot \mathrm{Dim}_N \lambda}{N^n},$$

which are supported on the set of all Young diagrams with $= n$ boxes and $\leq N$ rows.

That is,

$$M_N^{(\gamma)} = \sum_{n=0}^{\infty} \pi_{\gamma N}(n) M_{N,n}$$

(this is an informal identity).

From the hook-length formula

$$\mathrm{Dim}_N \lambda = \prod_{\square\in\lambda} \frac{N + c(\square)}{hook(\square)},$$

(where $c(\square)$ is the content), we see that as $N \to \infty$, we have

$$\frac{\mathrm{Dim}_N \lambda}{N^n} \to \prod_{\square\in\lambda} \frac{1}{hook(\square)} = \frac{\dim\lambda}{n!},$$

where we used the hook-length formula for $\dim\lambda$. Recall that $\dim\lambda$ is the number of standard, and $\mathrm{Dim}_N \lambda$ is the number of semistandard Young tableaux.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Thus, we see that the Schur-Weyl measures $M_{N,n}$ as $N \to \infty$ converge to the so-called Plancherel measures on Young diagrams with $= n$ boxes (of course, the restriction on the number of rows is lifted in the limit):

$$\mathsf{Pl}_n(\lambda) := \frac{(\dim \lambda)^2}{n!}.$$

Consider the Poisson mixture of these Plancherel measures

$$\mathsf{Pl}_\theta := \sum_{n=0}^{\infty} \pi_\theta(n) \mathsf{Pl}_n$$

This is a probability measure on the set of all Young diagrams.

It was shown in [Oko00], [Joh01], and [BOO00] (3 different proofs) that the poissonized Plancherel measure is a determinantal point process. To see the statement, we should interpret Pl_θ as a determinantal point process. One cannot use the correspondence between rows and particles

$$\lambda = (\lambda_1, \ldots, \lambda_N) \mapsto (l_1, \ldots, l_N), \qquad \text{where } l_i = \lambda_i + N - i$$

because N — the number of rows — varies.

Instead, assign to $\lambda \in \mathbb{Y}$ a semi-infinite point configuration as follows:

$$\mathbb{Y} \ni \lambda \mapsto \{y_1, y_2, \ldots\}, \qquad \text{where } y_i = \lambda_i - i + \tfrac{1}{2}, \qquad i = 1, 2, \ldots.$$

The configuration $\{y_i\}_{i=1}^{\infty}$ lives on the shifted lattice $\mathbb{Z}' = \mathbb{Z} + \frac{1}{2}$. E.g., to $\lambda = (3, 3, 1)$ corresponds a point configuration on Fig. 7.1. The number of

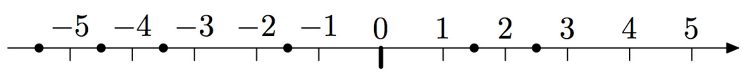


Figure 7.1

particles to the right of zero in this semi-infinite configuration equals the number of holes to the left of zero. The configuration is densely packed to the left of zero (starting from some $-K \ll 0$).

One can interpret the configuration $\{y_i\}_{i=1}^{\infty} = \left\{\lambda_i - i + \frac{1}{2}\right\}_{i=1}^{\infty}$ through the Young diagram, see Fig. 7.2.

7.1.5 Plancherel measures as limits of $M_N^{(\gamma)}$

As measures on Young diagrams, we clearly have

$$\mathsf{Pl}_\theta = \lim_{N \to \infty} M_N^{(\theta/N)}.$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

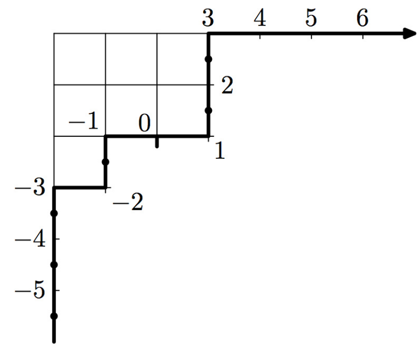


Figure 7.2

Theorem 7.1.6. *The determinantal correlation kernel of the poissonized Plancherel measure* Pl_θ *viewed as a point process on* $\ell^2(\mathbb{Z}')$ $(\mathbb{Z}' = \mathbb{Z} + \frac{1}{2})$ *coincides with the matrix of orthogonal projection onto the positive part of the spectrum of the following second order difference operator in* $\ell^2(\mathbb{Z}')$*:*

$$\left(D_\theta^{\mathsf{Pl}}\right)\psi(y) = \sqrt{\theta}\left(\psi(y+1) + \psi(y-1)\right) - y\psi(y).$$

Proof. We will look at the convergence of the operators $D_{\theta/N}$, where

$$(D_\gamma\varphi)(x) = \sqrt{\gamma(x+1)}\varphi(x+1) + \sqrt{\gamma x}\varphi(x-1) - (x+\gamma)\varphi(x),$$

to D_θ^{Pl}. This convergence is sufficiently nice (*strong resolvent convergence*), so that will imply the convergence of the corresponding projection operators. More technical details may be found in [Ols08].

The N-point configurations on $\mathbb{Z}_{\geq 0}$ are related to semi-infinite configurations on $\mathbb{Z}'$ as

$$x = y + N - \tfrac{1}{2}, \qquad x \in \mathbb{Z}_{\geq 0}, \qquad y \in \mathbb{Z}'.$$

Thus, we need to shift our operator D_γ (where $\gamma = \theta/N$). The shifted operator acts on the subspace $\ell^2(\mathbb{Z}')_{>-N} \subset \ell^2(\mathbb{Z}')$, of functions from $\ell^2(\mathbb{Z}')$ supported on $\left\{-N+\frac{1}{2}, -N+\frac{3}{2}, \dots\right\}$. The shifted operator $D_\gamma^{(N)}$ acts as

$$(D_\gamma^{(N)}\varphi)(y) = \sqrt{\gamma(y+N+\tfrac{1}{2})}\varphi(y+1) + \sqrt{\gamma(y+N-\tfrac{1}{2})}\varphi(y-1) - (y+\gamma+N-\tfrac{1}{2})\varphi(y).$$

Now, the kernel for $M_N^{(\gamma)}$ was the orthogonal spectral projection onto eigenvalues $\{-(N-1), \dots, -1, 0\}$ of D_γ, which is the same for the shifted $D_\gamma^{(N)}$. Equivalently, one can say that this kernel is the orthogonal spectral projection onto the positive part of the spectrum of the operator

$$((D_\gamma^{(N)} + (N-1/2)\mathbf{1})\varphi)(y)$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

$$= \sqrt{\gamma(y+N+\tfrac{1}{2})}\varphi(y+1) + \sqrt{\gamma(y+N-\tfrac{1}{2})}\varphi(y-1) - (y+\gamma)\varphi(y)$$

where $\mathbf{1}$ is the identity operator.

Recalling that $\gamma = \theta/N$, we may take the limit as $N \to \infty$ of the above operator in a straightforward way, and this concludes the proof. □

7.1.6 Discrete Bessel kernel

The correlation kernel of the measure Pl_θ can be expressed through the Bessel functions in an explicit way.

Definition 7.1.7. Let $m \in \mathbb{C}$ and $u > 0$. The *Bessel functions of the first kind* is defined as

$$J_m(u) = \left(\frac{u}{2}\right)^m \sum_{k=0}^{\infty} \frac{(-1)^k (u/2)^{2k}}{k!\Gamma m + k + 1}. \tag{7.1.3}$$

Set

$$J_m := J_m(2\sqrt{\theta}),$$

and view J_m as a function of m, not of the argument. Our m's will be integers.

For $m = -1, -2, \ldots$ the quantity J_m is defined as

$$J_m = (-1)^m J_{-m}.$$

These quantities satisfy

$$\sqrt{\theta}\,(J_{m+1} + J_{m-1}) = mJ_m, \qquad m \in \mathbb{Z}.$$

So, the functions J_m as functions on $\mathbb{Z}$ satisfy the difference equation $D_\theta^{\mathsf{Pl}} J_m = 0$. Define the functions

$$\psi_a(y) := J_{a+y}, \qquad a, y \in \mathbb{Z}'.$$

Note that $a+y$ — the indices — are integers and not half-integers. The functions $\varphi_a(y)$ are also self-dual, $\psi_a(y) = \psi_y(a)$.

These functions ψ are eigenfunctions for D_θ^{Pl}:

$$D_\theta^{\mathsf{Pl}} \psi_a = a\psi_a, \qquad a \in \mathbb{Z}'.$$

One can also show (e.g., see [Ols08]) that these functions form an orthonormal basis in the space $\ell^2(\mathbb{Z})$.

One can establish the following formula for the kernel of the Plancherel measure Pl_θ:

$$K^{\mathsf{Pl}_\theta}(y, y') = \sqrt{\theta} \frac{J_{y-\frac{1}{2}} J_{y'+\frac{1}{2}} - J_{y+\frac{1}{2}} J_{y'-\frac{1}{2}}}{y - y'}, \qquad y, y' \in \mathbb{Z}'. \tag{7.1.4}$$

For $y = y'$ one should forget that $y, y' \in \mathbb{Z}'$ and define the value $K^{\mathsf{Pl}_\theta}(y, y')$ by continuity.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

7.1.7 Limit behavior of the Plancherel measures

Remark 7.1.8. In fact, the $\theta \to \infty$ limit of the Plancherel measure Pl_θ is equivalent to the $n \to \infty$ limit of the non-poissonized measures Pl_n. This is because the Poisson distribution π_θ has mean θ and variance θ, so the standard deviation is $\sqrt{\theta}$, and thus the number of boxes in a Young diagram distributed according to Pl_θ is roughly in $(\theta - C\sqrt{\theta}, \theta + C\sqrt{\theta})$, i.e., concentrated. But we will consider $\theta \to \infty$ behavior of the determinantal point processes.

We will discuss two limit regimes and reconstruct the limit shape of Plancherel Young diagrams due to Vershik-Kerov-Logan-Shepp [VK77], [LS77]. The asymptotic results we will discuss may be found in [Oko00], [Joh01], and [BOO00].

The point process $\{\lambda_i - i + \frac{1}{2}\}$ can be understood geometrically, see Fig. 7.5 (this is the so-called "Russian picture" of Young diagrams). We place a particle

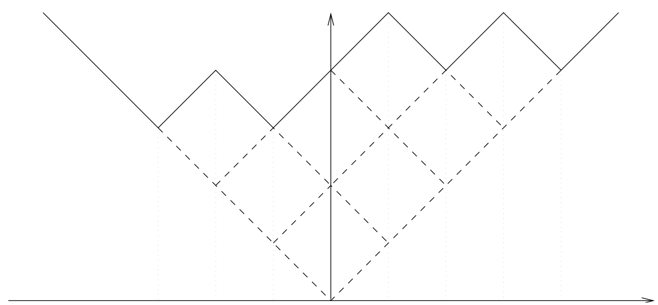


Figure 7.3: Picture taken from [Bia01].

onto each segment with slope -1.

The limit regime is $\theta \to \infty$, and we assume that we scale the Young diagrams so that the area of the diagram will be 1 (on average). This involves scaling $\mathbb{Z}'$ to $\sqrt{\theta}\mathbb{Z}'$. The limit shape of these Young diagrams has a curved part from -2 to 2, and flat parts outside $[-2, 2]$, see Fig. 7.4.

There are (at least) two limit regimes:

- (bulk) Zoom around a point "in the bulk", to see lattice picture; this leads to a nontrivial lattice limit.
- (edge) Around ± 2, the behavior is completely different: it is no longer lattice, and the scaling should be taken completely different.

Bulk limit

Consider our difference operator

$$\left(D_\theta^{\mathsf{Pl}}\right) \psi(y) = \sqrt{\theta} \left(\psi(y+1) + \psi(y-1)\right) - y\psi(y).$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

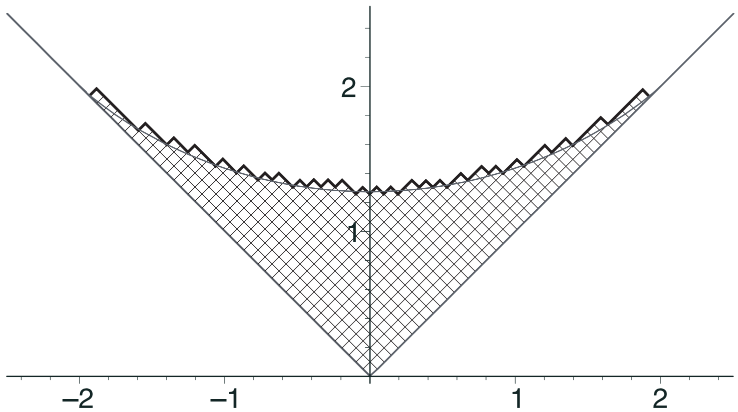


Figure 7.4: Picture taken from [BOO00].

The bulk limit means that we zoom around $2c$.[3] Thus, we need to plug in $y = 2c\sqrt{\theta} + \widetilde{y}$, $\widetilde{y} \in \mathbb{Z}$.

Then we will take limit as $\theta \to \infty$. Note that we may scale the difference operator by positive constants as this will not change the projection operator onto the positive part of its spectrum.

In the limit we get the following:

$$\widetilde{D}\varphi(\widetilde{y}) = \varphi(\widetilde{y}+1) + \varphi(\widetilde{y}-1) - 2c\varphi(\widetilde{y}). \tag{7.1.5}$$

This a difference operator with constant coefficients (and, moreover, translation invariant), we work with it using Fourier transforms (= Fourier series, as we are on the lattice).

Let $\mathbb{T}$ be the unit circle with complex coordinate ζ, i.e., $\mathbb{T} = \{\zeta : |\zeta| = 1\}$. From $\ell^2(\mathbb{Z})$ we pass to the dual space $L^2(\mathbb{T})$ with Lebesgue measure, it has orthonormal basis $\{\zeta^m\}_{m\in\mathbb{Z}}$.

The operator $\widetilde{D}$ in $\ell^2(\mathbb{Z})$ is translation invariant, and its Fourier transform is the operator of multiplication by a function. This function is

$$\psi(\zeta) = \zeta + \zeta^{-1} - 2c = 2\left(\operatorname{Re}\zeta - c\right).$$

This is a real function as it should be (because $\widetilde{D}$ is an Hermitean operator).

Spectrum of the operator of multiplication by $\psi(\zeta)$ is simply the range $\psi(\zeta)$, $\zeta \in \mathbb{T}$. The positive part is an arc $\{\operatorname{Re}\zeta \geq c\}$. So the projection operator is the inverse Fourier transform of the operator of multiplication by the indicator of this arc. Clearly, we see that meaningful values are $c \in [-1, 1]$.

The projection itself in the space $\ell^2(\mathbb{Z})$ is a convolution operator of conjugation. The kernel of this projection is translation invariant

$$S(\widetilde{y}, \widetilde{y}') = S(\widetilde{y} - \widetilde{y}'),$$

[3] It will be clear that $-1 \leq c \leq 1$.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

and has the form

$$S(k) = \begin{cases} \dfrac{\sin(Ak)}{\pi k}, & \text{if } k \neq 0, \\[2ex] \dfrac{A}{\pi}, & \text{if } k = 0. \end{cases} \tag{7.1.6}$$

Here $k \in \mathbb{Z}$ and $A = \arccos c$.

This correlation kernel is called *discrete sine kernel*, it describes the local limit behavior of the Young diagram. The limiting determinantal process is translation invariant and appears in many other models as a universal limit.

Limit shape

The proportion of particles under the sine kernel is just the density function, $\arccos c/\pi$. But this proportion can be written in terms of the slope of the limit shape curve. Represent the Young diagram as a graph of a broken line as on Fig. 7.5. Thus understood Young diagram is a function $\lambda(u)$, $-\infty < u < +\infty$. Scale the Young diagram so it has area 1 on average. Then the random Young diagrams concentrate as $\theta \to \infty$ around a deterministic curve $\Omega(u)$, which is curved above $[-2, 2]$, and flat elsewhere (Fig. 7.4). We have

$$\Omega'(u) = 1 - 2\rho(u) = 1 - \frac{2}{\pi} \arccos \frac{u}{2} = \frac{2}{\pi} \arcsin \frac{u}{2}.$$

Thus, by integration we reconstruct the limit shape

$$\Omega(u) = \begin{cases} \frac{2}{\pi}\left(u \arcsin \frac{u}{2} + \sqrt{4 - u^2}\right), & |u| \leq 2; \\ |u|, & |u| \geq 2. \end{cases}$$

Edge limit

Let $y = 2\sqrt{\theta} + v\theta^{\alpha}$. We will also extract the correct scaling exponent α from our formulas. Here $v \in \mathbb{R}$ is a new coordinate around the edge.

We have $\varphi(y) = F(v)$, let us write a differential operator on F. We have

$$\varphi(y \pm 1) = F(v \pm \theta^{-\alpha}) = F(v) \pm \theta^{-\alpha} F'(v) + \frac{1}{2}\theta^{-2\alpha} F''(v) + \dots.$$

The operator D_θ^{PI},

$$\left(D_\theta^{\mathsf{PI}}\right) \psi(y) = \sqrt{\theta}\,(\psi(y+1) + \psi(y-1)) - y\psi(y),$$

acts on F as (asymptotically)

$$\left(D_\theta^{\mathsf{PI}}\right) F(v) = \theta^{\frac{1}{2} - 2\alpha} F''(v) - \theta^{\alpha} F(v).$$

Multiply by $\theta^{-\alpha} > 0$, we have

$$\left(D_\theta^{\mathsf{PI}}\right) F(v) = \theta^{\frac{1}{2} - 3\alpha} F''(v) - F(v).$$

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Thus, to have nontrivial limit, we must have $\alpha = 1/6$, and so the limiting operator is

$$DF = F'' - uF.$$

This is the so-called Airy differential operator, and the Airy functions $Ai(x)$,

$$Ai(x) = \frac{1}{\pi}\int_0^\infty \cos(t^3/3 + tx)dt$$

are solutions of $D(Ai) = 0$, decaying exponentially fast at $x \to +\infty$.

The projection operator onto the positive part of the spectrum can be written as

$$\frac{Ai(v_1)Ai'(v_2) - Ai'(v_1)Ai(v_2)}{v_1 - v_2} = \int_0^\infty Ai(v_1 + y)Ai(v_2 + y)dy.$$

7.1.8 Ulam's problem

In early 1960-s S. Ulam formulated problem about maximal increasing subsequence. Take the symmetric group $\mathfrak{S}(n)$. Every permutation is a word, and with every such word take the length of the maximal increasing subword (such a subword may be not unique, but the length is well-defined). E.g., for $\sigma = (4213657)$ this length is 4. For the uniform probability measure on $\mathfrak{S}(n)$, the length of the maximal increasing subsequence becomes a random variable L_n. The asymptotics of this random variable was of interest for several decades.

Remark 7.1.9. The distribution of L_n coincides with the distribution of the first row of a Plancherel-distributed (Pl_n) random Young diagram. This may be seen from some combinatorial bijections (Robinson-Schensted-Knuth correspondence).

The answer is that
(LLN)

$$\frac{L_n}{\sqrt{n}} \to 2.$$

('CLT-like' result)

$$L_n \sim 2\sqrt{n} + n^{\frac{1}{6}} \cdot U,$$

where U is some random variable. This random variable is *not* Gaussian.

7.1.9 More general limit shapes: measures $M_N^{(\gamma)}$

Here we consider the measures $M_N^{(\gamma)}$ which are mixtures of the Schur-Weyl measures with Poisson parameter $\theta = \gamma N$, and at the same time they are orthogonal polynomial ensembles with Charlier (Poisson) weight with parameter γ.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

The measures $M_N^{(\gamma)}$ on the set of positive signatures $\mathbb{GT}_N^+$, or, which is the same, on Young diagrams with $\le N$ rows, depend on two parameters which may go to infinity simultaneously. This gives rise to a family of limit shapes which depend on the relative rate of this growth. The number of boxes in the Young diagram distributed as $M_N^{(\gamma)}$ is $\theta = \gamma N$ on average. The measures $M_N^{(\gamma)}$ are of *Plancherel-type* (an informal concept), and length of rows and columns of random Young diagrams is of order $\sqrt{\theta}$. So, if N grows faster than $\sqrt{\theta}$, then the condition that the number of rows is $\le N$ becomes trivial in the limit, and the limit shape is expected to be the Vershik-Kerov-Logan-Shepp one.

More generally, introduce an additional parameter $c \ge 0$:

$$\sqrt{\theta} \sim cN.$$

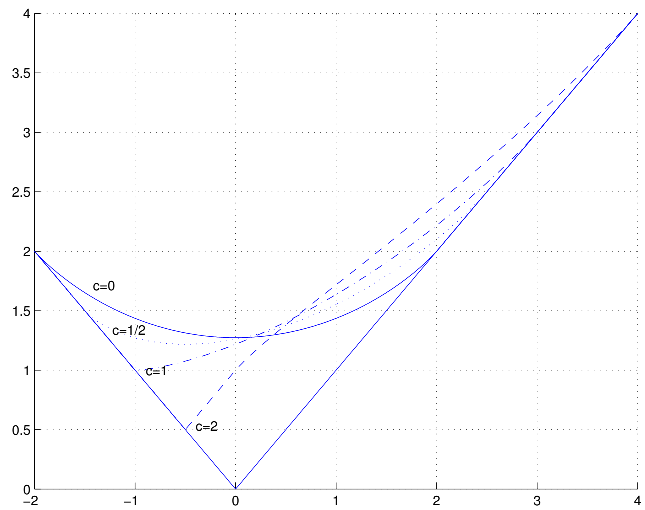


Figure 7.5: Limit shapes of $M_N^{(\gamma)}$, picture is from [Bia01].

The limit shapes $\Omega(c; u)$ have linear and curved parts. The curved part has the following derivative plus a boundary condition:

- $\dfrac{\partial \Omega(c; u)}{\partial u} = \dfrac{2}{\pi} \arcsin \dfrac{u + c}{2\sqrt{1 + uc}}$.
- $\Omega(c; u)\big|_{u=c+2} = c + 2$, i.e., the curve at $u = c + 2$ touches the right ray $u = v$.

Let us describe various particular cases of the limit shapes in more detail:

0. ($c = 0$) Vershik-Kerov-Logan-Shepp limit shape for the Plancherel measure.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

1. $(0 < c < 1)$ A "small" deformation. Curved part corresponds to $c - 2 \le u \le c + 2$, and
$$\frac{\partial}{\partial u}\Omega(c; u)|_{u=c\pm 2} = \pm 1,$$

2. $(c > 1)$ This has three linear parts: one more linear part from $u = -\frac{1}{c}$, $v = \frac{1}{c}$ to $u = c-2$, $v = c-2+\frac{2}{c}$, with slope $+1$. The curved part touches the new linear part: $\frac{\partial}{\partial u}\Omega(c; u)|_{u=c-2} = 1$. This is because the restriction on the number of rows in a diagram becomes crucial.

3. $(c = 1)$ The difference is that the curved part is not tangent to the left ray, the derivative there is 0.

The limit shape is obtained through difference operators in a straightforward way: we have the operator

$$((D + (N-1)\mathbf{1})\varphi)(x) = \sqrt{\gamma(x+1)}\varphi(x+1) + \sqrt{\gamma x}\varphi(x-1) - (x+\gamma-N+1)\varphi(x),$$

where $\gamma = \theta/N$. The kernel of the point process $x_i = \lambda_i + N - i$ is the projection onto positive part of the spectrum of this operator.

The bulk scaling is

$$x = [u\sqrt{\theta} + N + y], \qquad \sqrt{\theta} \sim cN.$$

The value u means limiting global position we are zooming in, and $y \in \mathbb{Z}$ is the new bulk lattice coordinate.

We have

$$\begin{aligned}
&x \sim u\sqrt{\theta} + N + y;\\
&N \sim c^{-1}\sqrt{\theta},\\
&\gamma = \frac{\theta}{N} = c\sqrt{\theta},\\
&\sqrt{\gamma x} \sim \sqrt{\gamma(x+1)} \sim \sqrt{\theta}\sqrt{1+uc};\\
&-(x+\gamma-N+1) \sim -\sqrt{\theta}(u+c).
\end{aligned}$$

In the limit as $N \to \infty$ we have:

$$((D + (N-1)\mathbf{1})\varphi)(x) = \sqrt{\theta}\sqrt{1+uc}(\varphi(x+1) + \varphi(x-1)) - \sqrt{\theta}(u+c)\varphi(x).$$

Divide the operator by $\sqrt{\theta}\sqrt{1+uc} > 0$. The limit operator acts on $F(y)$, $y \in \mathbb{Z}$ as

$$F(y) \mapsto F(y+1) + F(y-1) - 2\frac{u+c}{2\sqrt{1+uc}}F(y).$$

Here u means the global position, and c is the constant of the limit curve.

The constant $\frac{u+c}{2\sqrt{1+uc}}$ plays the same role as "$2c$" in the treatment of the Plancherel measures, it completely determines the limit shape curves:

$$\Omega'(u; c) = 1 - 2\rho(u) = 1 - \frac{2}{\pi}\arccos\frac{u+c}{2\sqrt{1+uc}} = \frac{2}{\pi}\arcsin\frac{u+c}{2\sqrt{1+uc}}.$$

Examining various cases, we reach the description of the limit shapes above.

New lattice process

We note that there is one more limit behavior in the case $c = 1$ at the edge; for $c \neq 1$ the tangency of the limit shape to the linear parts of the boundary leads to familiar Airy kernel.

For $c = 1$ and $u = -1$ we have

$$\frac{\partial \Omega(u;c)}{\partial u} = 1 - 2\rho(u) = 0,$$

so the density of particles is $\rho(u) = \frac{1}{2}$.

In fact, in the limit as $N \to \infty$, $\sqrt{n} \sim N$ (n is the same as θ under poissonization) *with no scaling*, looking at the smallest rows of the Young diagram, one sees a nontrivial local structure.

We look at the rows $0 \le \lambda_N \le \lambda_{N-1} \le \dots$, or

$$x_N = \lambda_N,\ x_{N-1} = \lambda_{N-1} + 1,\ \dots.$$

We need to consider finer asymptotics:

$$N = \theta^{\frac{1}{2}} - s\theta^{\frac{1}{4}} + o(\theta^{\frac{1}{4}}),$$

where s is one more parameter, i.e., $\theta \sim N^2 + sN^{\frac{3}{2}}$.

In the difference operators, we see the following limiting operator on $\mathbb{Z}_{\ge 0}$:

$$\psi(x) \mapsto \sqrt{x+1}\psi(x+1) + \sqrt{x}\psi(x) - s\psi(x)$$

So the kernel here is a projection onto a $[s, +\infty]$ part of the spectrum of the operator

$$\psi(x) \mapsto \sqrt{x+1}\psi(x+1) + \sqrt{x}\psi(x),$$

This difference operator has a continuous spectrum, and the kernel of the process is expressed through Hermite polynomials; but in the expression for the kernel the integer indices of the polynomials act as variables (cf. discrete Bessel kernel). In more detail see [BO07].

7.1.10 GUE; bulk and edge asymptotics (and the semicircle law): differential operators approach

[Similar treatment of the GUE asymptotics may be found in the T. Tao's blog: `http://goo.gl/XoLX4`].

7.2 Double contour integral kernels and their asymptotic analysis. Saddle points method

7.2.1 Asymptotics of the Plancherel measure

[Asymptotics of the Plancherel measure Pl_θ using double contour integrals are explained in detail at an accessible level in [Oko02], Section 3].

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

How to obtain contour integral representation of the discrete Bessel kernel

Suppose we have the discrete Bessel kernel

$$K(x,y)=\sqrt{\theta}\frac{J_{x-\frac{1}{2}}J_{y+\frac{1}{2}}-J_{x+\frac{1}{2}}J_{y-\frac{1}{2}}}{x-y},\qquad x,y\in\mathbb{Z}+\frac{1}{2}$$

where $J_n=J_n(2\sqrt{\theta})$ is the Bessel function. We also have the following summation expression:

$$K(x,y)=\sum_{s=\frac{1}{2},\frac{3}{2},\frac{5}{2},\dots}J_{x+s}J_{y+s},\qquad x,y\in\mathbb{Z}+\frac{1}{2}$$

One formula is equivalent to the other with the help of the three-term relations on the Bessel functions:

$$\sqrt{\theta}\,(J_{x+1}+J_{x-1})=xJ_x,\qquad x\in\mathbb{Z}.$$

Indeed, multiply by $(x-y)$, and apply the three-term relations to every term like

$$(x-y)J_{s+x}J_{s+y}.$$

The Bessel functions admit generating functions as

$$e^{\sqrt{\theta}(z-z^{-1})}=\sum_{n\in\mathbb{Z}}J_n(2\sqrt{\theta})z^n.$$

This allows to obtain the following double contour integral representation:

$$K(x,y)=\frac{1}{(2\pi\mathrm{i})^2}\oint\oint\frac{e^{\sqrt{\theta}(z-z^{-1}-w+w^{-1})}}{z-w}\frac{dzdw}{z^{x+\frac{1}{2}}w^{-y+\frac{1}{2}}}.$$

Here the integration is over $|z|>|w|$.

We write

$$\sum_{s=\frac{1}{2},\frac{3}{2},\frac{5}{2},\dots}J_{x+s}J_{y+s}=\sum_{s=\frac{1}{2},\frac{3}{2},\frac{5}{2},\dots}[z^{x+s}w^{-y+s}]e^{\sqrt{\theta}(z-z^{-1}-w+w^{-1})},$$

where $[\cdots]$ means coefficient by, and then we can sum the geometric series and get the answer.

An alternative way is described in [Oko02, §2.4].

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Acknowledgements

A part of these lecture notes is based on

- lecture notes of the course "Asymptotic representation theory" given at the Independent University of Moscow in 2009-10 by Grigori Olshanski. They are accessible at `https://lpetrov.cc/art/` (in Russian).
- the lecture notes of the course "Gibbs measures on branching graphs" given at the MIT in Fall 2011 by Alexei Borodin. They can be accessed at `http://math.berkeley.edu/~svs/borodin/` (typed by Steven Sam).

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

Bibliography

[AESW51] M. Aissen, A. Edrei, I. J. Schoenberg, and A. Whitney, *On the generating functions of totally positive sequences*, Proc. Nat. Acad. Sci. U. S. A. **37** (1951), 303–307.

[Ald85] D. Aldous, *Exchangeability and related topics*, École d'été de probabilités de Saint-Flour. XIII—1983, Lecture Notes in Mathematics, vol. 1117, Springer-Verlag, Berlin, 1985.

[ASW52] M. Aissen, I. J. Schoenberg, and A. Whitney, *On the generating functions of totally positive sequences I*, J. Analyse Math. **2** (1952), 93–103.

[BF08] A. Borodin and P. Ferrari, *Anisotropic growth of random surfaces in 2+1 dimensions*, arXiv:0804.3035 [math-ph].

[Bia01] P. Biane, *Approximate factorization and concentration for characters of symmetric groups*, International Mathematics Research Notices **2001** (2001), no. 4, 179–192, arXiv:math/0006111 [math.RT].

[BK08] A. Borodin and J. Kuan, *Asymptotics of Plancherel measures for the infinite-dimensional unitary group*, Adv. Math. **219** (2008), no. 3, 894–931, arXiv:0712.1848 [math.RT].

[BO00] A. Borodin and G. Olshanski, *Harmonic functions on multiplicative graphs and interpolation polynomials*, Electronic Journal of Combinatorics **7** (2000), R28, arXiv:math/9912124 [math.CO].

[BO06] ———, *Stochastic dynamics related to Plancherel measure on partitions*, Representation Theory, Dynamical Systems, and Asymptotic Combinatorics (V. Kaimanovich and A. Lodkin, eds.), 2, vol. 217, Transl. AMS, 2006, pp. 9–22, arXiv:math–ph/0402064.

[BO07] ———, *Asymptotics of Plancherel-type random partitions*, Journal of Algebra **313** (2007), no. 1, 40–60, arXiv:math/0610240.

[BO12] ———, *The boundary of the Gelfand-Tsetlin graph: A new approach*, Adv. Math. **230** (2012), 1738–1779, arXiv:1109.1412 [math.CO].

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

[BO13] ______, *The Young bouquet and its boundary*, Moscow Mathematical Journal **13** (2013), no. 2, 193–232, arXiv:1110.4458 [math.RT].

[BOO00] A. Borodin, A. Okounkov, and G. Olshanski, *Asymptotics of Plancherel measures for symmetric groups*, J. Amer. Math. Soc. **13** (2000), no. 3, 481–515, arXiv:math/9905032 [math.CO].

[Bor11a] A. Borodin, *Determinantal point processes*, Oxford Handbook of Random Matrix Theory (G. Akemann, J. Baik, and P. Di Francesco, eds.), Oxford University Press, 2011, arXiv:0911.1153 [math.PR].

[Bor11b] ______, *Gibbs measures on branching graphs*, 2011, Lecture notes. Typed by Steven Sam.

[Boy83] R. Boyer, *Infinite traces of AF-algebras and characters of* $U(\infty)$, J. Operator Theory **9** (1983), 205–236.

[Edr53] A. Edrei, *On the generating function of a doubly infinite, totally positive sequence*, Trans. Amer. Math. Soc. **74** (1953), 367–383.

[Fel71] W. Feller, *An Introduction to Probability Theory and Its Applications. Vol. 1-2*, Wiley, 1968, 1971.

[FH91] W. Fulton and J. Harris, *Representation Theory: A First Course*, Springer, 1991.

[FZ00] S. Fomin and A. Zelevinsky, *Total positivity: Tests and parametrizations*, The Mathematical Intelligencer **22** (2000), no. 1, 23–33, arXiv:math/9912128 [math.RA].

[GK02] F.R. Gantmacher and M.G. Krein, *Oscillation matrices and kernels and small vibrations of mechanical systems*, AMS Chelsea Publishing, 2002.

[GO06] A. Gnedin and G. Olshanski, *The boundary of Eulerian number triangle*, Moscow Mathematical Journal **6** (2006), no. 3, 461–475, arXiv:math/0602610.

[GO09] ______, *A q-analogue of de Finetti's theorem*, The electronic journal of combinatorics **16** (2009), R16, arXiv:0905.0367 [math.PR].

[GP05] A. Gnedin and J. Pitman, *Exchangeable Gibbs partitions and Stirling triangles*, Zapiski Nauchn. Semin. POMI **325** (2005), 83–102, reproduced in J. Math. Sci., New York 138 (2006), No. 3, 5674–5685; arXiv:math/0412494.

[GV85] I. Gessel and G. Viennot, *Binomial determinants, paths, and hook length formulae*, Advances in mathematics **58** (1985), no. 3, 300–321.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

[HKPV06] J.B. Hough, M. Krishnapur, Y. Peres, and B. Virág, *Determinantal processes and independence*, Probability Surveys **3** (2006), 206–229, arXiv:math/0503110 [math.PR].

[ITZ10] M. Ishikawa, H. Tagawa, and J. Zeng, *A q-analogue of catalan hankel determinants*, arXiv:1009.2004 [math.CO].

[Joh01] K. Johansson, *Discrete orthogonal polynomial ensembles and the Plancherel measure*, Annals of Mathematics **153** (2001), no. 1, 259–296, arXiv:math/9906120 [math.CO].

[Kar68] S. Karlin, *Total positivity, Vol.1*, Stanford, 1968.

[Ker03] S. Kerov, *Asymptotic representation theory of the symmetric group and its applications in analysis*, vol. 219, AMS, Translations of Mathematical Monographs, 2003.

[KM59] S. Karlin and J. McGregor, *Coincidence probabilities*, Pacific J. Math. **9** (1959), 1141–1164.

[KOO98] S. Kerov, A. Okounkov, and G. Olshanski, *The boundary of Young graph with Jack edge multiplicities*, Intern. Math. Research Notices **4** (1998), 173–199, arXiv:q-alg/9703037.

[Kra99] C. Krattenthaler, *Advanced determinant calculus*, Séminaire Lotharingien Combin **42** (1999), B42q.

[Len73] A. Lenard, *Correlation functions and the uniqueness of the state in classical statistical mechanics*, Communications in Mathematical Physics **30** (1973), no. 1, 35–44.

[Lin73] B. Lindström, *On the vector representations of induced matroids*, Bulletin of the London Mathematical Society **5** (1973), no. 1, 85–90.

[LS77] BF Logan and L.A. Shepp, *A variational problem for random Young tableaux*, Advances in Mathematics **26** (1977), no. 2, 206–222.

[Mac95] I.G. Macdonald, *Symmetric functions and Hall polynomials*, 2nd ed., Oxford University Press, 1995.

[NSU91] A.F. Nikiforov, S.K. Suslov, and V.B. Uvarov, *Classical orthogonal polynomials of a discrete variable*, Springer-Verlag Berlin, 1991.

[Oko97] Andrei Okounkov, *Log-Concavity of Multiplicities with Application to Characters of $U(\infty)$*, Advances in Mathematics **127** (1997), no. 2, 258–282.

[Oko00] A. Okounkov, *Random matrices and random permutations*, International Mathematics Research Notices **2000** (2000), no. 20, 1043–1095, arXiv:math/9903176 [math.CO].

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

[Oko02] ______, *Symmetric functions and random partitions*, Symmetric functions 2001: Surveys of Developments and Perspectives (S. Fomin, ed.), Kluwer Academic Publishers, 2002, arXiv:math/0309074 [math.CO].

[Ols08] G. Olshanski, *Difference operators and determinantal point processes*, Functional Analysis and Its Applications **42** (2008), no. 4, 317–329, arXiv:0810.3751 [math.PR].

[OO98] A. Okounkov and G. Olshanski, *Asymptotics of Jack polynomials as the number of variables goes to infinity* , Int. Math. Res. Notices **1998** (1998), no. 13, 641–682, arXiv:q-alg/9709011.

[Pet12] L. Petrov, *The Boundary of the Gelfand-Tsetlin Graph: New Proof of Borodin-Olshanski's Formula, and its q-analogue*, arXiv:1208.3443 [math.CO]. To appear in Moscow Math. J.

[Phe66] R.R. Phelps, *Lectures on Choquet's theorems*, Van Nostrand, 1966.

[PS02] M. Prähofer and H. Spohn, *Scale invariance of the PNG droplet and the Airy process*, J. Stat. Phys. **108** (2002), 1071–1106, arXiv:math.PR/0105240.

[Shi95] A.N. Shiryaev, *Probability*, Springer, 1995.

[Sos00] A. Soshnikov, *Determinantal random point fields*, Russian Mathematical Surveys **55** (2000), no. 5, 923–975, arXiv:math/0002099 [math.PR].

[Sta01] R. Stanley, *Enumerative Combinatorics. Vol. 2*, Cambridge University Press, Cambridge, 2001, With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin.

[Tho64] E. Thoma, *Die unzerlegbaren, positive-definiten Klassenfunktionen der abzählbar unendlichen, symmetrischen Gruppe*, Math. Zeitschr **85** (1964), 40–61.

[Ver74] A. Vershik, *A description of invariant measures for actions of certain infinite-dimensional groups*, Soviet Math. Dokl. **15** (1974), no. 5, 1396–1400.

[VK77] A.M. Vershik and S.V. Kerov, *Asymptotics of the plancherel measure of the symmetric group and the limiting form of young tableaux*, Doklady AN SSSR **233** (1977), no. 6, 1024–1027, English translation: Soviet Mathematics Doklady **18** (1977), 527—531.

[VK81a] A. Vershik and S. Kerov, *Asymptotic theory of the characters of the symmetric group*, Funktsional. Anal. i Prilozhen. **15** (1981), no. 4, 15–27, 96.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38

[VK81b] ______, *Characters and factor representations of the infinite symmetric group*, Dokl. Akad. Nauk SSSR **257** (1981), no. 5, 1037–1040.

[VK82] ______, *Characters and factor-representations of the infinite unitary group*, Dokl. Akad. Nauk SSSR **267** (1982), no. 2, 272–276.

[VK87] ______, *Locally semisimple algebras. Combinatorial theory and the K_0-functor*, Journal of Mathematical Sciences **38** (1987), no. 2, 1701–1733.

[Voi76] D. Voiculescu, *Representations factorielles de type II_1 de $U(\infty)$*, J. Math. Pures Appl. **55** (1976), 1–20.

[Wey97] H. Weyl, *The Classical Groups. Their Invariants and Representations*, Princeton University Press, 1997.

AMS Open Math Notes: Works in Progress; Reference # OMN:202001.110813; Last Revised: 2020-01-10 09:59:38