Determinantal point processes

Maurice Duits

Contents

Chapter 1. Discrete determinantal point processes 5
1. Discrete point processes 5
2. Discrete determinantal point process 9
3. Determinants of (infinite) matrices 12

Chapter 2. Examples 15
1. Self-adjoint kernels 15
2. Self-adjoint projections 16
3. Translation invariant processes 17
4. L-ensembles 17
5. Eynard–Mehta Theorem 19
6. Biorthogonal Ensembles 21
7. Extended biorthogonal ensembles 21

Chapter 3. General Determinantal point processes 23
1. Random measures 23
2. Point processes 24
3. Factorial moments and correlation functions 25
4. Determinantal point processes 27
5. Convergence of determinantal point processes 30
6. Examples 30

Chapter 4. Biorthogonal ensembles 33
1. Definition 33
2. Orthogonal polynomial ensembles 35

Chapter 5. Non-intersecting path ensembles 37
1. Non-intersecting paths 37
2. Graphs obtained by gluing columns 39
3. Column-homogeneous parameters and Laurent transition matrices 45

Appendix A. Determinants and trace of kernels 49
1. Introduction 49
2. Preliminaries 50
3. Trace and determinant for finite rank operators 51
4. Continuous extensions 53
5. Fredholm determinants 57
6. A generalization 58
7. Trace class operators 59
8. Comparison between $\det_\Phi$, $\det_R$ and $\det_1$ 60

Appendix B. Laurent, Toeplitz, and Hankel operators 63

Appendix. Bibliography 67

CHAPTER 1

Discrete determinantal point processes

Roughly speaking, point process are random configurations of particles. Determinantal point processes are examples of simple point process, meaning that different particles can not occupy the same location. If the underlying space is continuous, a rigorous definition requires some effort. However, simple point processes on discrete spaces are easier to define and we choose to this before we come to the general situations in the next chapter.

1. Discrete point processes

In this chapter we will always consider point processes on a countable set $\mathbb{X}$. Such point processes can be constructed by considering the power set $2^{\mathbb{X}}$.

If $\mathbb{X}$ is finite, then so is $2^{\mathbb{X}}$.

If $\mathbb{X}$ is countably infinite, then we assume that $\mathbb{X}$ is a complete metric space with no accumulation points (that is, every bounded subset of $\mathbb{X}$ has finitely many points). We then equip $2^{\mathbb{X}}$ with the topology generated by the sets $\{X \in 2^{\mathbb{X}} \mid B \subset X\}$ for bounded sets B.

Definition 1. *A* **simple point process** *on a discrete $\mathbb{X}$ is a (Borel) probability measure on $2^{\mathbb{X}}$.*

1.1. Finite dimensional distributions.

An important notion to study point processes is that of counting statistics and finite dimensional distributions.

Definition 2. *Let $A \subset \mathbb{X}$. Then the counting statistic $N(A)$ on a set A for a point process on $\mathbb{X}$ is defined as*

$$N(A) = \sum_{x \in X} \chi_A(x)$$

where $\chi_A(x) = 1$ if $x \in A$ and $\chi_A(x) = 0$ otherwise.

Definition 3. *For $m \in \mathbb{N}_0$, bounded sets $A_1, \ldots, A_m$ and $k_1, \ldots, k_m \in \mathbb{N}_0$ we define the finite dimensional distribution as*

$$P_m(A_1, \ldots, A_k; n_1, \ldots, n_m) = \mathbb{P}(N(A_1) = k_1, \ \ldots N(A_k) = k_m).$$

The first important result is the following standard theorem.

Theorem 4. *A point process is uniquely determined by its finite dimensional distributions.*

Proof. See [**2**]. □

A natural question is to ask when a given collection of distributions $P_m(A_1, \ldots, A_k; n_1, \ldots, n_m)$ are the finite dimensional distributions for a point processes. This question has been studied in the literature and we mention the following characterization due to the Kolmogorov (?).

Theorem 5. *Let $P_m(A_1,\ldots,A_k;n_1,\ldots,n_m)$ for $m\in\mathbb{N}_0$, bounded sets $A_1,\ldots,A_m$ and $k_1,\ldots,k_m\in\mathbb{N}_0$, be collection of probability distributions. In order for them to be a the finite dimensional distributions of a point process it is necessary and sufficient that the following properties holds*

(1) *For any permutation $i_1,\ldots,i_m$ of $1,\ldots,m$ we have*

$$P_m(A_1,\ldots,A_k;n_1,\ldots,n_m)=P_m(A_{i_1},\ldots,A_{i_m};n_{i_1},\ldots,n_{i_m})$$

(2) $\sum_{r=0}^{\infty}P_{m+1}(A_1,\ldots,A_k,A_{m+1};n_1,\ldots,n_m,r)=P_m(A_1,\ldots,A_k;n_1,\ldots,n_m)$

(3) *For each disjoint pair of Borel sets A_1 nad A_2, we have that*

$$P_3(A_1,A_2,A_1\cup A_2;n_1,n_2,n_3)=0,$$

if $n_1+n_2\neq n_3$.

Proof. See [**2**, Theorem 9.2.X]. □

Since Theorems 4 and 36 only play a minor role for us and our text is not intended as a thorough treatment of the abstract theory of point processes, we only provide reference for these statements.

1.2. Correlation functions. An important role is played by the probability generating function defined as

$$\Phi(g)=\mathbb{E}\left[\prod_{x\in X}(1+g(x))\right] \tag{1}$$

where g is a function with bounded support in $\mathbb{X}$. By taking $g=(z-1)\chi_A$ for a bounded set A and $|z|\leq 1$ we find

$$\Phi(g)=\mathbb{E}\left[z^{N(A)}\right],$$

and, more generally, if $g=\sum_{j=1}^m(z_j-1)\chi_{A_j}$ for disjoint bounded sets A_j and $|z_j|\leq 1$ we have [1]

$$\Phi(g)=\mathbb{E}\left[\prod_{j=1}^m z_j^{N(A_j)}\right].$$

Note that this implies that the $\Phi(g)$ determine the finite dimensional distributions uniquely and therewith the point processes.

The correlation functions can be defined by expanding the product in (1). First we introduce some notation: For any $A\subset\mathbb{X}$ set

$$A^{(m)}=\{(x_1,\ldots,x_m)\in A^m\mid\forall_{j\neq k}\ :\ x_j\neq x_k\}. \tag{2}$$

Then we have the following theorem.

Theorem 6. *For any function g with finite support we have*

$$\Phi(g)=1+\sum_{m=1}^{\infty}\frac{1}{m!}\sum_{(x_1,\ldots,x_m)\subset\mathbb{X}^{(m)}}g(x_1)\cdots g(x_m)\rho_m(x_1,\ldots,x_m)$$

where

$$\rho_m(x_1,\ldots,x_m)=\mathbb{P}(\{x_1,\ldots,x_m\}\subset X).$$

[1]Since $\mathbb{X}$ is discrete, any function g with finite support, can be written in the form $g=\sum_{j=1}^k(z_j-1)\chi_{A_j}$.

*For $m \geq 1$, the function ρ_m is called the m-**point correlation function** for the point process.*

Proof. First note that

$$\prod_{x \in \mathbb{X}} (1 + f(x)) = \sum_{m=0}^{\infty} \sum_{S \subset \mathbb{X} : |S| = m} \prod_{x \in S} f(x),$$

for any f with finite support. Also note that the sum is finite since f has finite support and, therefore, the product vanishes if m exceeds the number of points in the support of f. An important step is now to administrate the terms in the sum in a different way by introducing an order on the set S by putting the elements in m-tuples. The entries of these m-tuples should all be different.

With the notation (2) we see that

$$\prod_{x \in S} f(x) = \frac{1}{m!} \sum_{(x_1, \ldots, x_m) \in S^{(m)}} f(x_1) \cdots f(x_m),$$

where the division by $m!$ is needed to account for all possible orderings of the elements in S. Moreover,

$$\prod_{x \in \mathbb{X}} (1 + f(x)) = \sum_{m=0}^{\infty} \frac{1}{m!} \sum_{(x_1, \ldots, x_m) \subset \mathbb{X}^{(m)}} f(x_1) \cdots f(x_m),$$

By setting $f(x) = g(x) \chi_X(x)$ we see that

$$\prod_{x \in X} (1 + g(x)) = \prod_{x \in \mathbb{X}} (1 + f(x)),$$

and thus

$$\Phi(g) = \mathbb{E}\left[\prod_{x \in X} (1 + g(x))\right] = \sum_{m=0}^{\infty} \frac{1}{m!} \sum_{(x_1, \ldots, x_m) \subset \mathbb{X}^{(m)}} g(x_1) \cdots g(x_m)\, \mathbb{E}[\chi_{\{x_1, \ldots, x_m\} \subset X}].$$

Since $\mathbb{E}[\chi_{\{x_1, \ldots, x_m\} \subset X}] = \mathbb{P}(\{x_1, \ldots, x_m\} \subset X)$ the statement follows. $\square$

The fact that the correlation functions are probabilities is special for discrete processes. As we will see later, if $\mathbb{X}$ is not discrete then there is still an analogue of Theorem 6, but the correlation functions are no longer probabilities and slighlty more elaborate to define. We will do this using the notion of factorial moments. [2]

1.3. Factorial moments. We recall the Pochhammer symbol: for $a \in \mathbb{R}$ and $\ell \in \mathbb{N}$ we set $(a)_\ell = a \cdots (a-1) \cdots (a - \ell + 1)$. Then, for $m \in \mathbb{N}$, we will construct a measure μ_m on $X^{(m)}$ as follows. Start with $1 \leq k \leq m$, pairwise disjoint sets $A_1, \ldots, A_k \subset \mathbb{X}$ and $\ell_j \in \mathbb{N}$ such that $\ell_1 + \ldots + \ell_m = m$. Then we define the **factorial moments** by

$$\mu_m \left(A_1^{(\ell_1)} \times \ldots \times A_k^{(\ell_k)}\right) = \mathbb{E}\left[\prod_{j=1}^{k} (N(A_j)_{\ell_j}\right].$$

[2]The next paragraph can be skipped on first reading

It can be shown that μ has a unique extension to a measure on $\mathbb{X}^m$, which we will denote by μ_m. [3] This measure has a density $\tilde{\rho}_m$ such that

$$\mu_m\left(A_1^{(\ell_1)} \times \ldots \times A_k^{(\ell_k)}\right) = \sum_{(x_1,x_2,\ldots,x_m)\in A_1^{(\ell_1)}\times\ldots\times A_k^{(\ell_k)}} \tilde{\rho}_m(x_1,\ldots,x_m).$$

By taking $A_j = \{x_j\}$ and $\ell_j = 1$ for $j = 1,\ldots,m$, we see that $\tilde{\rho}_m = \rho_m$. In other words, the correlation functions are densities for the factorial moments viewed as measures on $\mathbb{X}^{(m)}$.

That the factorial moments are natural objects can be seen from the following expansion

$$\begin{aligned}
(3)\quad \mathbb{E}\left[\prod_{j=1}^k z_j^{N(A_j)}\right] &= \mathbb{E}\left[\prod_j (1+(z_j-1))^{N(A_j)})\right] \\
&= \mathbb{E}\left[\prod_{j=1}^k \left(\sum_{\ell_j=0}^{N(A_j)} \frac{(N(A_j))_{\ell_j}}{\ell_j!}(z_j-1)^{\ell_j}\right)\right] \\
&= \mathbb{E}\left[\prod_{j=1}^k \left(\sum_{\ell_j=0}^{\infty} \frac{(N(A_j))_{\ell_j}}{\ell_j!}(z_j-1)^{\ell_j}\right)\right] \\
&= \sum_{\ell_1,\ldots,\ell_k=0}^{\infty} \frac{\prod_{j=1}^k (z_j-1)^{\ell_j}}{\ell_1!\cdots\ell_k!}\,\mathbb{E}\left[\prod_{j=1}^k (N(A_j))_{\ell_j}\right] \\
&= \sum_{m=0}^{\infty}\frac{1}{m!}\sum_{\ell_1+\ldots+\ell_k=m}\binom{m}{\ell_1\ \ell_2\ \cdots\ \ell_k}\prod_{j=1}^k (z_j-1)^{\ell_j}\,\mathbb{E}\left[\prod_{j=1}^k (N(A_j))_{\ell_j}\right].
\end{aligned}$$

We see that $\mathbb{E}\left[\prod_{j=1}^k z_j^{N(A_j)}\right]$ can be viewed as a generating function for the factorial moments.

In fact, the factorial moments can be used for an alternative derivation of Theorem 6. To this end, let $g = \sum_{j=1}^k (z_j-1)\chi_{A_j}$. Then

$$\begin{aligned}
&\prod_{j=1}^k (z_j-1)^{\ell_j}\,\mathbb{E}\left[\prod_{j=1}^k (N(A_j))_{\ell_j}\right] \\
&\quad= \sum_{(x_1,x_2,\ldots,x_m)\in A_1^{(\ell_1)}\times\ldots\times A_k^{(\ell_k)}} \prod_{j=1}^m (z_j-1)^{\ell_j}\rho_m(x_1,\ldots,x_m) \\
&\quad\quad= \sum_{(x_1,x_2,\ldots,x_m)\in A_1^{(\ell_1)}\times\ldots\times A_k^{(\ell_k)}} \prod_{j=1}^m g(x_j)\rho_m(x_1,\ldots,x_m) =
\end{aligned}$$

[3]The collection of sets $A_1^{(\ell_1)} \times \ldots A_k^{(\ell_k)}$ form a semi-ring. The extension to a measure is a standard exercise that is worth working out.

Now since $\prod_{j=1}^m g(x_j)\rho_m(x_1,\ldots,x_m)$ is symmetric we have

$$\sum_{\ell_1+\ldots+\ell_k=m} \binom{m}{\ell_1 \quad \ell_2 \quad\quad \ell_k} \sum_{(x_1,x_2,\ldots,x_m)\in A_1^{(\ell_1)}\times\ldots\times A_k^{(\ell_k)}} \prod_{j=1}^m g(x_j)\rho_m(x_1,\ldots,x_m)$$
$$= \sum_{(x_1,\ldots,x_m)\in\mathbb{X}^{(m)}} \prod_{j=1}^m g(x_j)\rho_m(x_1,\ldots,x_m)$$

By inserting this into (10) we find

$$\mathbb{E}\left[\prod_{x\in X}(1+g(x))\right] = \mathbb{E}\left[\prod_{j=1}^k z_j^{N(A_j)}\right]$$
$$= \sum_{m=0}^\infty \frac{1}{m!} \sum_{(x_1,\ldots,x_m)\in\mathbb{X}^{(m)}} \prod_{j=1}^m g(x_j)\rho_m(x_1,\ldots,x_m)$$

for $g=\sum_{j=1}^k (z_j-1)\chi_{A_j}$ and, since $\mathbb{X}$ is discrete, therewith for any function g with finite support. We thus arrive at the conclustion of Theorem 6. We will use this alternative derivation when introduction correlation functions for continuous point processes.

2. Discrete determinantal point process

Definition 7. *A* **determinantal point process** *on discrete set* $\mathbb{X}$ *is a simple point process on* $\mathbb{X}$ *such that there exists a* $K:\mathbb{X}\times\mathbb{X}\to\mathbb{C}$ *such that* [4]

$$\mathbb{P}(\{x_1,\ldots,x_k\}\subset X) = \det\left(K(x_i,x_j)\right)_{i,j=1}^k \tag{4}$$

for $k\in\mathbb{N}$*, distinct points* $x_1,\ldots,x_k\in\mathbb{X}$*. The function* K *is called a* **correlation kernel** *for the process.*

Before we continue we emphasize that the correlation kernel is not unique. Indeed, if K is a correlation kernel for a determinantal point process on a discrete space $\mathbb{X}$ and $G:\mathbb{X}\to\mathbb{C}\setminus\{0\}$ a nowhere vanishing function, then [5]

$$K_G(x_1,x_2) = \frac{G(x_1)}{G(x_2)}K(x_1,x_2), \qquad x_1,x_2\in\mathbb{X},$$

is also a correlation kernel for the same determinantal point process.

From the definition we have the special cases

$$\mathbb{P}(x\in X) = K(x,x)$$

and also, for $x\neq y$,

$$\mathbb{P}(x\in X, y\in X) = K(x,x)K(y,y)-K(x,y)K(y,x),$$

[4] The right-hand side is symmetric in $x_1,\ldots,x_n$

[5] This conjugation of the kernel is the easiest way to see that the kernel is not unique. It is far less clear if this conjugation is the only freedom that exits. This in fact, is an open problem.

which means [6]

$$\mathbb{P}(x \in X, y \in X) - \mathbb{P}(x \in X)\mathbb{P}(y \in X) = -K(x,y)K(y,x), \tag{5}$$

We thus see that $K(x,y)K(y,x)$ measures how strongly the two events $x \in X$ and $y \in X$ are correlated.

A natural statistic for point process is the number of points in a given interval.

Proposition 8. *Let K be a correlation kernel for a determinantal point process on a discrete space $\mathbb{X}$. Let $A \subset \mathbb{X}$ be bounded. Then*

(1) $\mathbb{E}[N(A))] = \sum_{x\in A} K(x,x)$
(2) $\mathbb{V}ar[N(A)] = \sum_{x\in A} K(x,x) - \sum_{x\in A}\sum_{y\in A} K(x,y)K(y,x)$

Proof. 1. The expectation is straightforward.
2. This is the result of a computation.

$$\begin{aligned}
\mathbb{E}[\sum_{x\in A}\sum_{y\in A}\chi_X(x)\chi_X(y)] &= \sum_{x\in A}\sum_{y\in A, y\neq x}\mathbb{E}[\chi_X(x)\chi_X(y)] + \sum_{x\in A}\mathbb{E}[\chi_X(x)] \\
&= \sum_{x\in A}\sum_{y\in A,y\neq x} \det\begin{pmatrix} K(x,x) & K(x,y) \\ K(y,x) & K(y,y)\end{pmatrix} + \sum_{x\in A} K(x,x) \\
&= \sum_{x\in A}\sum_{y\in A} \det\begin{pmatrix} K(x,x) & K(x,y) \\ K(y,x) & K(y,y)\end{pmatrix} + \sum_{x\in A} K(x,x) \\
&= \sum_{x\in A}\sum_{y\in A} (K(x,x)K(y,y) - K(x,y)K(y,x)) + \sum_{x\in A} K(x,x) \\
&= \sum_{x\in A} K(x,x) - \sum_{x\in A}\sum_{y\in A} K(x,y)K(y,x) + \left(\sum_{x\in A} K(x,x)\right)^2.
\end{aligned}$$

By subtracting $(\mathbb{E}[N(A)])^2 = \left(\sum_{x\in A} K(x,x)\right)^2$ the statement follows. □

It should be clear now that the computation in the proof can be extended to the higher moment of $N(A)$ and even to joint moments of $N(A_j)$. It follows then from Theorem 4 that the process is completely determined by a correlation kernel K. The results are rather complicated expressions. However, the Laplace transforms will have a a very elegant expression. Indeed, the $\Phi(g)$ can be expressed in terms of determinants involving the correlation kernel K:

[6] As we will see, in many interesting cases we have $K(y,x) = \overline{K(x,y)}$. If so, then the right-hand side of (5) is negative and thus such determinantal point processes are negatively correlated. Moreover, $K(x,y)$ can be solved from (5) up to a phase factor. However, requiring $K(y,x) = \overline{K(x,y)}$ would be too restrictive, since many interesting point process that arise in modern research do not obey this symmetry.

THEOREM 9. *Let K be the correlation kernel for a determinantal point process on a discrete space $\mathbb{X}$. For any $g : \mathbb{X} \to \mathbb{C}$ with finite support S, we have* [7]

$$\mathbb{E}[\prod_{x\in X}(1+g(x))] = \tag{6}$$

$$= 1 + \sum_{k=1}^{\infty} \frac{1}{k!} \sum_{x_1,x_2\dots,x_k\in\mathbb{X}} g(x_1)\cdots g(x_k) \det\left(K(x_i,x_j)\right)_{i,j=1}^{k})$$

$$= \det(I+gK),$$

where $\det(I+gK)$ *is the determinant of the* $|S|\times|S|$ *matrix* $(\delta_{x,y}+g(x)K(x,y))_{x,y\in S}$. [8]

PROOF. The first equality is a rather straightforward consequence of (4), Theorem 6 and the fact that the determinant vanishes whenever $x_i = x_j$ for $i \neq j$.

The fact that we can write this series as a determinant follows from the following standard identity for matrices:

$$\det(I+A) = 1 + \sum_{k=1}^{n} \frac{1}{k!} \sum_{i_1,i_2\dots,i_k\in\{1,\dots,n\}} \det\left(A_{i_j,i_\ell}\right)_{j,\ell=1}^{k}.$$

for any $n \times n$ matrix A.

□

COROLLARY 10. *(Gap probabilities) Let K be the correlation kernel for a determinantal point process on a discrete space $\mathbb{X}$. Then,*

$$\mathbb{P}(\{y_1,\dots,y_k\}\cap X = \emptyset) = \det\left(I - K(y_j,y_\ell)\right)_{j,\ell=1}^{k}$$

for any $k\in\mathbb{N}$ and distinct points $y_1,\dots,y_k \in \mathbb{X}$.

PROOF. Define the function $g(x) = -1$ if $x \in \{y_1,\dots,y_k\}$ and $g(x) = 0$ otherwise. Then

$$\prod_{x\in X}(1+g(x)) = \begin{cases} 1, & \text{if } \{y_1,\dots,y_k\}\cap X = \emptyset \\ 0, & \text{if } \{y_1,\dots,y_k\}\cap X \neq \emptyset \end{cases}$$

And thus,

$$\mathbb{E}\left[\prod_{x\in X}(1+g(x))\right] = \mathbb{P}(\{y_1,\dots,y_k\}\cap X = \emptyset)$$

The statement then follows from Theorem 9. □

There is an interesting consequence to this result.

COROLLARY 11. [9] *(Particle hole duality) Consider a determinantal point process on $\mathbb{X}$ that has K as a correlation kernel. We say that there is particle at X if there is a point at X. If there is no point, we say that there is hole at x. The holes also define a determinantal point process with kernel $I - K$.*

[7] Note that the determinant in the summand vanishes if $y_i = y_j$ for $i \neq j$ and hence we can drop the condition that the points need to be distinct.

[8] The infinite series is in fact only a sum with at most $|S|$ terms.

[9] Whereas the theorems and propositions here all of generalizations to continuous determinantal point processes, this particle/hole duality only makes sense for *discrete* point processes.

Finally, we mention the following result on the finite dimensional distribution of a determinantal point process.

THEOREM 12. *For bounded A_j and non-negative integers n_j we have*

$$P_k(A_1,\ldots,A_k;n_1,\ldots,n_k) = \frac{1}{(2\pi i)^k}\oint\cdots\oint \det\left(I+\left(z_1^{\chi_{A_1}}\cdots z_k^{\chi_{A_k}}-1\right)K\chi_{\cup_{j=1}^k A_j}\right)\frac{dz_1\cdots dz_k}{z_1^{n_1+1}\cdots z_k^{n_k+1}} \tag{7}$$

3. Determinants of (infinite) matrices

We will also sometimes need to use the expression (6), but for functions g with unbounded support. An example is the distribution of the largest point in a point process on $\mathbb{Z}$ wich can be respresentated as a gap probability for an infinite set. If g has unbouded support, the matrix gK is infinite and we have to be careful with dealing with (6). To this end, we will recall some theory about traces and determinants of infinite matrices.

Let R be a $\mathbb{X}\times\mathbb{X}$. We define the norm of R by

$$\|R\|_F = \max\left(\sum_{x\in\mathbb{X}}|R(x,x)|,\left(\sum_{x,y\in\mathbb{X}}|R(x,y)|^2\right)^{\frac{1}{2}}\right)$$

Following [**1**] [10] we call the space

$$\mathcal{D}=\{R\mid \|R\|_F<\infty\},$$

the Von Koch-Riesz algebra. This is a Banach algebra of infinite matrices that act as operators on $\ell_2(\mathbb{X})$. The finite rank operators are dense in this algebra.

DEFINITION 13. *For $R\in\mathcal{D}$ we define the* **trace** TrR *by*

$$\mathrm{Tr}R=\sum_{x\in\mathbb{X}}R(x,x)$$

and, the **determinant** *of $I+R$ by*

$$\det(I+R)=1+\sum_{k=1}^{\infty}\frac{1}{k!}\sum_{x_1,x_2\ldots,x_k\in\mathbb{X}}\det\left(R(x_i,x_j)\right)_{i,j=1}^k. \tag{8}$$

At this point, it is not clear that the determinant is well defined.

LEMMA 14. *Both $R\mapsto \mathrm{Tr}R$ and $R\mapsto\det(I+R)$ are well-defined. The map $R\mapsto\mathrm{Tr}R$ is Lipschitz continuous and $R\mapsto\det(I+R)$ is locally Lipschitz continuous with respect to $\|\cdot\|_F$.*

PROOF. The proof can be found in [**1**, Theorem II.7.1]. □

LEMMA 15. *The trace and determinant have the following properties*

(1) $\mathrm{Tr}(R_1+R_2)=\mathrm{Tr}R_1+\mathrm{Tr}R_2$
(2) $\det(1+R_1+R_2+R_1R_2)=\det(1+R_1)\det(I+R_2)$
(3) $\det(1+R_1R_2)=\det(1+R_2R_1)$

PROOF. This follows from the fact that the identies are standard for finite matrices and a continuity argument. □

[10] To be precise, [**1**] apply this with $\mathbb{X}=\mathbb{Z}$

THEOREM 16. *Suppose that K is the correlation kernel of a determinantal point process and $g : \mathbb{X} \to [-1, 0]$ is such that the $\mathbb{X} \times \mathbb{X}$ matrix with entries $g(x)K(x, y)$ is in the Von-Koch Riesz algebra, then*

$$\mathbb{E}[\prod_{x\in\mathbb{X}}(1 + g(x))] = \det(1 + gK)$$

PROOF. If X is finite then this is Theorem 9.

If X is infinite, then the product $\prod_{x\in\mathbb{X}}(1 + g(x))$ is an infinite product of numbers in $[0, 1]$ and therefore well-defined. Moreover,

$$\prod_{x\in X}(1 + g(x)) = \lim_{r\to\infty} \prod_{x\in X\cap B_{x_0,r}}(1 + g(x)) = \lim_{r\to\infty} \prod_{x\in X}(1 + \chi_{B_{x_0,r}})g(x))$$

and the limit is independent of the choice of x_0. By dominated convergence, we also have

$$\mathbb{E}\left[\prod_{x\in X}(1 + g(x))\right] = \lim_{r\to\infty} \mathbb{E}\left[\prod_{x\in X}(1 + \chi_{B_{x_0,r}})g(x))\right].$$

Now by Theorem 9 we have

$$\mathbb{E}\left[\prod_{x\in X}(1 + \chi_{B_{x_0,r}}(x)g(x))\right] = \det(1 + \chi_{B_{x_0,r}} gK).$$

The statement now follows by taking the limit $r \to \infty$ and using the continuity of the determinant. □

It is important to note that K does not have to be in the Von Koch- Riesz algebra! The function g can be used to mollify the kernel K. In the latter result, it only does it by acting on the first variable. Often we would need to mollify K in both arguments and use a more symmetric version:

$$K_g = sgn(g(x))\sqrt{|g(x)g(y)|}K(x, y)$$

If K is the Von Koch-Riesz algebra then $\det(I + gk) = \det(1 + K_g)$ but it can happen that K_g is in this algebra but not gK.

THEOREM 17. *Suppose that K is the correlation kernel of a determinantal point process and $g : \mathbb{X} \to [-1, 0]$ is such that the $\mathbb{X} \times \mathbb{X}$ matrix K_g is in the Von-Koch Riesz algebra, then*

$$\mathbb{E}[\prod_{x\in\mathbb{X}}(1 + g(x))] = \det(1 + K_g)$$

The fact that g takes values in $[0, 1]$ ensures that the infinite products are well-defined. There are of course other ways of ensuring this.

THEOREM 18. *Theorems 16 and 17 also hold for general $g : \mathbb{X} \to \mathbb{C}$ under the extra condition $\sum_{x\in\mathbb{X}} |g(x)| < \infty$, then*

CHAPTER 2

Examples

We now discuss several examples of determinantal point processes.

1. Self-adjoint kernels

Suppose that K is an $\mathbb{X}\times\mathbb{X}$ matrix that defines a self-adjoint bounded operator on $\ell_2(\mathbb{X})$. When is this operator the correlation kernel for a determinantal point process?

It is clear from (6) that the principal minors of K should be non-negative and thus the matrix needs to be positive semidefinite. By the particle–hole duality, also $I-K$ must be positive semidefinite. We thus need that $0 \le K \le I$. In fact, this is also sufficient.

THEOREM 19. *Let K be a self-adjoint bounded operator on $\ell_2(\mathbb{X})$. Then K is a correlation kernel for a determinantal point process if and only if $0 \le K \le I$.*

PROOF. As we already discussed that the conditions are necessary, it remains to prove that they are sufficient. Let $0 \le K \le I$ and consider the family of functions

$$\begin{aligned} (9)\quad & P_k(A_1,\dots,A_k;n_1,\dots,n_k) \\ &= \frac{1}{(2\pi i)^k}\oint\cdots\oint \det\Big(I+\big(z_1^{\chi_{A_1}}\cdots z_k^{\chi_{A_k}}-1\big)K\Big)\frac{dz_1\cdots dz_k}{z_1^{n_1+1}\cdots z_k^{n_k+1}}. \end{aligned}$$

We claim that this is a family of probability functions satisfying the criteria of Theorem 36 and thus these are the finite dimensional distributions of a point process. Once we have proved this claim, it is also straightforward that this point process is determinantal with K as a correlation kernel.

The most elaborate part of the work is to prove that the numbers in (9) are indeed non-negative. To this end, we replace K by $K_t = tK$ for $0 < t < 1$. Then, by continuity,

$$\begin{aligned} & P_k(A_1,\dots,A_k;n_1,\dots,n_k) \\ &= \lim_{t\uparrow 1}\frac{1}{(2\pi i)^k}\oint\cdots\oint \det\Big(I+\big(z_1^{\chi_{A_1}}\cdots z_k^{\chi_{A_k}}-1\big)K_t\Big)\frac{dz_1\cdots dz_k}{z_1^{n_1+1}\cdots z_k^{n_k+1}}. \end{aligned}$$

The benefit of working with K_t instead of K is that $I-K_t$ is invertible (if $K<I$, then we can take $t=1$ and this extra step is unnecessary). Thus

$$\begin{aligned} & \det\Big(I+\big(z_1^{\chi_{A_1}}\cdots z_k^{\chi_{A_k}}-1\big)K_t\Big) \\ &= \det(I-K_t)\det\Big(I+z_1^{\chi_{A_1}}\cdots z_k^{\chi_{A_k}}K_t(I-K_t)^{-1}\Big). \end{aligned}$$

It is therefore sufficient to prove that (9) with K replaced by K_t is non-negative.

Now $0 \le \det(I - K_t)$ and also $K_t(I - K_t)^{-1} \ge 0$. By expanding the second determinant on the right-hand side using (8), we see that the coefficients in front of the term $z_1^{n_1} \cdots z_k^{n_k}$ are non-negative, and thus the numbers in (9) are indeed non-negative.

That the expressions in (9) are probability distributions can be easily seen by summing over $n_1, \ldots, n_k$ from 0 to ∞ and then invoking the residue theorem from complex analysis k times.

Similar computations (the details are left to the reader) can be used to verify that the conditions of Theorem 36 are satisfied, and this finishes the proof. □

The latter result directly gives us a wide class of determinantal point processes, and many important examples are of this type. Having a correlation kernel that is a self-adjoint operator also has some important consequences, since properties of self-adjoint operators can be used to prove properties of the determinantal point processes that they generate.

PROPOSITION 20. *Let K be the correlation kernel for a determinantal point process on a discrete space $\mathbb{X}$. Assume that K is self-adjoint. The probability that the number of points in the random configuration is infinite is either* 0 *or* 1*. It is* 1 *if* $\mathrm{Tr} K = \sum_{x \in \mathbb{X}} K(x,x) = +\infty$ *and* 0 *if* $\mathrm{Tr} K = \sum_{x \in \mathbb{X}} K(x,x) < \infty$*.*

PROOF. If $\mathbb{E}[N(\mathbb{X})] = \mathrm{Tr} K = \sum_{x \in \mathbb{X}} K(x,x) < \infty$ then $\mathbb{P}(N(\mathbb{X}) = +\infty) = 0$.

On the other hand, assume that $\mathbb{E}[N(\mathbb{X})] = \mathrm{Tr} K = \sum_{x \in \mathbb{X}} K(x,x) = +\infty$. First we note that for any self-adjoint operator $0 \le B \le I$ we have

$$\det(I - B) \le e^{-\mathrm{Tr} B}.$$

Then, for any bounded set R and $k \in \mathbb{N}$ we have

$$\mathbb{P}(N(R) = k) \le 2^k \, \mathbb{E}\big[2^{-N(R)}\big] = 2^k \det\big(I - \tfrac{1}{2} K_R\big) \le \exp\big(-\tfrac{1}{2} \mathrm{Tr} K_R\big),$$

where K_R is the restriction of K to $R \times R$.

Now, by taking a sequence of bounded sets R_n such that $R_n \uparrow \mathbb{X}$ and $\mathrm{Tr} K_{R_n} \to +\infty$, we find

$$\mathbb{P}(N(\mathbb{X}) = k) = \lim_{n \to \infty} \mathbb{P}(N(R_n) = k) = 0.$$

Since this is true for every k, we must have $\mathbb{P}(N(\mathbb{X}) = +\infty) = 1$. □

2. Self-adjoint projections

Another important class of examples comes from orthogonal projections, that is, kernels K that satisfy $K^* = K$ and $K^2 = K$. Of special interest are the orthogonal projections onto n–dimensional subspaces, i.e., kernels of the form

$$K(x,y) = \sum_{j=1}^{n} f_j(x) f_j(y),$$

where $\{f_j\}_{j=1}^n$ is a set of orthonormal functions in $\ell_2(\mathbb{X})$.

PROPOSITION 21. *Let K be the correlation kernel for a determinantal point process on a discrete space $\mathbb{X}$. Assume that K is self-adjoint. Then the following are equivalent:*

(1) *The number of points in the random configuration equals n with probability* 1*.*
(2) *K is a self-adjoint projection onto an n-dimensional subspace.*

PROOF. $1 \Leftarrow 2$. Since $K^2 = K$ we know that $\mathrm{V}ar[N(\mathbb{X})] = 0$. Moreover, $\mathbb{E}[N(\mathbb{X})] = \mathrm{Tr}K = n$.

$1 \Rightarrow 2$. Since K is self-adjoint we know by Theorem 19 that $0 \le K \le I$ and hence $K^2 \le K$. Combining this with the assumption $\mathrm{V}ar[N(\mathbb{X})] = \mathrm{Tr}(K - K^2) = 0$ we see that $K^2 - K = 0$ and hence K is a projection. That it projects onto an n-dimensional subspace follows again from $\mathbb{E}[N(\mathbb{X})] = \mathrm{Tr}K = n$. □

3. Translation invariant processes

LEMMA 22. *Let $f : \mathbb{T} \to [0,1]$ be a measurable function on the unit circle and denote its Fourier coefficients by*

$$\hat{f}_k = \frac{1}{2\pi i} \oint f(z) \frac{dz}{z^{k+1}}.$$

Then $K(x,y) = \hat{f}_{x-y}$ for $x, y \in \mathbb{Z}$ defines a determinantal point process on $\mathbb{Z}$.

The process is **translation invariant**. *That is, if g is a function of bounded support and $g_y(x) = g(x-y)$ for $y \in \mathbb{Z}$, then*

$$\Phi(g) = \Phi(g_y).$$

EXAMPLE 23. *If $f(z) = \rho$ with $0 < \rho < 1$, then the determinantal point process is the (product) Bernoulli point process on $\mathbb{Z}$ with constant intensity ρ. In this case,*

$$\det\big(K(x_i, x_j)\big)_{i,j=1}^n = \prod_{i=1}^n K(x_i, x_i) = \rho^n.$$

EXAMPLE 24. *Take $f(z) = \frac{1}{4}(2 + z + 1/z)$.*

EXAMPLE 25. *Take $f = \chi_A$, the characteristic function of a measurable set A. In this case, the kernel has the additional property that $K^2 = K$ and thus K is a self-adjoint projection.*

In the special case $A = \{e^{it} \mid -c \le t \le c\}$, we have

$$K(x,y) = \frac{\sin c(x-y)}{\pi(x-y)},$$

and the determinantal point process with this kernel is called the **discrete sine process**.

4. L-ensembles

We start with some notation. Let $\mathbb{X}$ be a discrete space and $(L(x,y))_{x,y\in\mathbb{X}}$ be a $|\mathbb{X}| \times |\mathbb{X}|$ matrix. For any $X \subset \mathbb{X}$ the matrix L_X is defined as the $|X| \times |X|$ matrix

$$L_X = (L(x,y))_{x,y\in X}.$$

We will also use the notation $\overline{Y}$ for the complement of $Y \subset \mathbb{X}$.

DEFINITION 26. *Let $\mathbb{X}$ be a discrete space and $(L(x,y))_{x,y\in\mathbb{X}}$ be a $|\mathbb{X}| \times |\mathbb{X}|$ matrix with non-negative principal minors. Then, the L-ensemble corresponding to L is the point process on $\mathbb{X}$ defined by*

$$\mathbb{P}(X) = \frac{\det L_X}{\det(I + L)}.$$

THEOREM 27. *Let $\mathbb{X}$ be a discrete space and $(L(x,y))_{x,y\in\mathbb{X}}$ be a $|\mathbb{X}|\times|\mathbb{X}|$ matrix with non-negative principal minors. Then the associated L-ensemble is a determinantal point process with kernel*

$$K = L(I+L)^{-1} = I - (I+L)^{-1}.$$

PROOF. This follows by a computation. Let $Y \subset \mathbb{X}$. Then

$$\begin{aligned}\mathbb{P}(Y \subset X) &= \sum_{X\supset Y} \mathbb{P}(X) = \sum_{X\supset Y} \frac{\det L_X}{\det(I+L)} \\ &= \frac{\det(I_{\overline{Y}}+L)}{\det(I+L)}.\end{aligned}$$

Then we invoke some linear algebra:

$$\begin{aligned}\frac{\det(I_{\overline{Y}}+L)}{\det(I+L)} &= \det\big((I_{\overline{Y}}+L)(I+L)^{-1}\big) \\ &= \det\big(I - I_Y(I+L)^{-1}\big) = \det K_Y.\end{aligned}$$

This proves the statement. □

Not every determinantal point process is an L-ensemble. For instance, in an L-ensemble the number of points is not necessarily fixed.

There is however a modification of L-ensembles that will be useful to us.

Let $\mathbb{X}$ be a discrete space and $\mathbb{S} \subset \mathbb{X}$. Then, given an L-ensemble on $\mathbb{X}$ we can define a point process on $\mathbb{S}$ by requiring the random point-configuration for the L-ensemble to contain $\overline{\mathbb{S}}$.

DEFINITION 28. *The* **conditional L-ensemble** *is the point process on $\mathbb{S}$ defined by*

$$\mathbb{P}(X) = \frac{\det L_{X\cup\overline{\mathbb{S}}}}{\det(I_{\mathbb{S}}+L)}.$$

THEOREM 29. *The conditional L-ensemble is a determinantal point process with kernel*

$$K = I_{\mathbb{S}} - \big((I_{\mathbb{S}}+L)^{-1}\big)_{\mathbb{S}}.$$

PROOF. The proof is the same as above, with only minor modifications.

Let $Y \subset \mathbb{X}$. Then

$$\begin{aligned}\mathbb{P}(Y \subset X) &= \sum_{X\supset Y} \mathbb{P}(X) = \sum_{X\supset Y} \frac{\det L_{\bar{\mathbb{S}}\cup X}}{\det(I_{\mathbb{S}}+L)} \\ &= \frac{\det(I_{\mathbb{S}\setminus Y}+L)}{\det(I_{\mathbb{S}}+L)}.\end{aligned}$$

Then, again, we invoke some linear algebra:

$$\begin{aligned}\frac{\det(I_{\mathbb{S}\setminus Y)}+L)}{\det(I_{\mathbb{S}}+L)} &= \det\big((I_{\mathbb{S}\setminus Y}+L)(I_{\mathbb{S}}+L)^{-1}\big) \\ &= \det\big(I \; - I_Y(I_{\mathbb{S}}+L)^{-1}\big) \\ &= \det\big(I_{\mathbb{S}} \; - I_Y(I_{\mathbb{S}}+L)^{-1}\big) = \det K_Y.\end{aligned}$$

□

5. Eynard–Mehta Theorem

Let $\mathbb{X}$ be a measurable space, and fix integers $n \geq 1$ and $N \geq 0$. We consider a probability measure of the following form.

For each level $m = 1, \ldots, N$, let

$$x^{(m)} = \big(x_1^{(m)}, \ldots, x_n^{(m)}\big) \in \mathbb{X}^n.$$

Let $\{\phi_j\}_{j=1}^n$ and $\{\psi_j\}_{j=1}^n$ be functions on $\mathbb{X}$, and for $m = 1, \ldots, N-1$ let

$$T_{m,m+1} : \mathbb{X} \times \mathbb{X} \to \mathbb{C}.$$

We assume that

$$\det\big[\phi_j(x_k^{(1)})\big]_{j,k=1}^n \prod_{m=1}^{N-1} \det\big[T_{m,m+1}(x_j^{(m)}, x_k^{(m+1)})\big]_{j,k=1}^n \det\big[\psi_j(x_k^{(N)})\big]_{j,k=1}^n$$

defines a probability measure on $(\mathbb{X}^n)^N$.

As we will see later, measures of this type occur frequently in random matrix theory and in dimer models. We will show below that this measure can be interpreted as a conditional L-ensemble, and we will derive an explicit expression for the associated correlation kernel.

Consider the $\{1, \ldots, n\} \cup \mathbb{X}^N$ indexed matrix L, defined in block form by

$$L = \begin{pmatrix} 0 & \Phi & 0 & \cdots & \cdots & 0 \\ 0 & 0 & -T_{1,2} & 0 & \cdots & 0 \\ 0 & 0 & 0 & -T_{2,3} & & 0 \\ \vdots & & & & \ddots & \vdots \\ 0 & & & & 0 & -T_{N-1,N} \\ \Psi & 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}.$$

Here Φ is a $\{1, \ldots, n\} \times \mathbb{X}$ matrix and Ψ is a $\mathbb{X} \times \{1, \ldots, n\}$ matrix, defined by

$$\Phi_{j,x} = \phi_j(x), \qquad \Psi_{x,j} = \psi_j(x).$$

LEMMA 30. *The measure defined above is a conditional L-ensemble induced by the matrix L and the conditioning set $\mathbb{S} = \mathbb{X}^n$.*

PROOF. (To be completed.) □

We now compute the correlation kernel using the general theory of conditional L-ensembles. For this purpose, we will need the following standard linear algebra result.

LEMMA 31 (Schur complement inverse formula). *Let*

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

where $A \in \mathbb{C}^{m \times m}$, $D \in \mathbb{C}^{n \times n}$, and the block sizes are compatible. Assume that D is invertible and define the Schur complement

$$T := A - BD^{-1}C.$$

If T is invertible, then M is invertible and

$$M^{-1} = \begin{pmatrix} T^{-1} & -T^{-1}BD^{-1} \\ -D^{-1}CT^{-1} & D^{-1} + D^{-1}CT^{-1}BD^{-1} \end{pmatrix}.$$

PROOF. We use the block factorization

$$M = \begin{pmatrix} I & BD^{-1} \\ 0 & I \end{pmatrix} \begin{pmatrix} T & 0 \\ C & D \end{pmatrix}, \qquad T = A - BD^{-1}C.$$

Both factors are invertible under the stated assumptions. Their inverses are given by

$$\begin{pmatrix} I & BD^{-1} \\ 0 & I \end{pmatrix}^{-1} = \begin{pmatrix} I & -BD^{-1} \\ 0 & I \end{pmatrix}, \qquad \begin{pmatrix} T & 0 \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} T^{-1} & 0 \\ -D^{-1}CT^{-1} & D^{-1} \end{pmatrix}.$$

Multiplying these two matrices yields the stated formula for M^{-1}. □

We introduce the notation

$$T_{m_1,m_2} = \prod_{k=m_1}^{m_2-1} T_{k,k+1},$$

for $m_2 > m_1$.

THEOREM 32 (Eynard-Mehta Theorem). *The point process defined* (**??**) *defines a determinatnal point process with kernel*

$$K((m_1,x_1),(m_2,x_2)) = -\chi_{m_1<m_2} T_{m_1,m_2}(x_1,x_2) + \sum_{i,j=1}^{n} (T_{m_1,N}\Psi\ )(x_1,j)(G^{-1})_{j,i}\,(\Phi T_{1,m_2})\,(i,x_2)$$

PROOF. (sketch) We need to compute the inverse of

$$\begin{pmatrix} 0 & \Phi & 0 & \cdots & \cdots & 0 \\ 0 & I & -T_{1,2} & 0 & \cdots & 0 \\ 0 & 0 & I & -T_{2,3} & & 0 \\ \vdots & & & & \ddots & \vdots \\ 0 & & & & I & -T_{N-1,N} \\ \Psi & 0 & \cdots & \cdots & \cdots & I \end{pmatrix}.$$

and we write this in block form using

$$A = 0, B = \begin{pmatrix} \Phi & 0 & \cdots & \cdots & 0 \end{pmatrix}, C = \begin{pmatrix} 0 & \cdots & \cdots & 0 & \Psi \end{pmatrix}^T$$

and

$$D = \begin{pmatrix} I & -T_{1,2} & 0 & \cdots & 0 \\ 0 & I & -T_{2,3} & & 0 \\ \vdots & & & \ddots & \vdots \\ & & & I & -T_{N-1,N} \\ 0 & \cdots & \cdots & \cdots & I \end{pmatrix}.$$

Clearly, D is invertible and

$$D^{-1} = \begin{pmatrix} I & T_{1,2} & T_{1,3} & \cdots & T_{1,N} \\ 0 & I & T_{2,3} & & T_{2,N} \\ \vdots & & & \ddots & \vdots \\ & & & I & T_{N-1,N} \\ 0 & \cdots & \cdots & \cdots & I \end{pmatrix}.$$

Now $BD^{-1}C = \Phi T_{1,N}\Psi = G$ and

$$D^{-1}C = \begin{pmatrix} T_{1,N}\Psi \\ T_{2,N}\Psi \\ \vdots \\ T_{N-1,N}\Psi \\ \Psi \end{pmatrix}$$

and

$$BD^{-1} = \begin{pmatrix} \Phi & \Phi T_{1,2} & \Phi T_{1,3} & \cdots & \Phi T_{1,N} \end{pmatrix}$$

so that

$$D^{-1}C(BD^{-1}C)^{-1}BD^{-1}$$

is the matrix (in block form)

$$\left((T_{m_1,N}\Psi G^{-1}\Phi T_{1,m_2})_{m_1,m_2=1}^{N}\right)$$

where $T_{m,m} = I$.

□

6. Biorthogonal Ensembles

7. Extended biorthogonal ensembles

CHAPTER 3

General Determinantal point processes

1. Random measures

Let E be a complete separable metric space. Denote by $\mathcal{M}(E)$ the space of all non-negative Borel measures on E, and by $\mathcal{M}^{\#}(E)$ the subset of all measures that are finite on bounded subsets of E. Such measures are called *boundedly finite*.

We equip $\mathcal{M}^{\#}(E)$ with the topology of vague convergence: $\mu_n \to \mu$ if and only if

$$\int f \, d\mu_n \longrightarrow \int f \, d\mu$$

for all bounded continuous functions f with bounded support. The corresponding Borel σ-algebra is the smallest σ-algebra for which $\mu \mapsto \mu(A)$ is measurable for all bounded Borel sets A. With this topology, $\mathcal{M}^{\#}(E)$ is itself a complete separable metric space (see [**2**, Proposition 9.1.IV]).

A *random measure* on E is a random element of $\mathcal{M}^{\#}(E)$, i.e. a measurable mapping from a probability space into $\mathcal{M}^{\#}(E)$.

There are various ways to characterize random measures. Any random measure μ gives rise to a family of real-valued random variables, indexed by $A \in \mathcal{B}(E)$, via

$$\mu(A) : \Omega \to [0, \infty).$$

The distribution function of $\mu(A)$ is

$$F(A; x) = \mathbb{P}\big(\mu(A) \le x\big), \qquad x \in [0, \infty).$$

In fact, the law of μ is completely determined by the joint distributions

$$F(A_1, \dots, A_k; x_1, \dots, x_k) = \mathbb{P}\big(\mu(A_1) \le x_1, \dots, \mu(A_k) \le x_k\big),$$

where $k \in \mathbb{N}$ and $A_1, \dots, A_k$ are disjoint sets taken from any semi-ring of bounded sets that generates the Borel σ-algebra.

Another important characterization uses linear functionals. For any Borel function $f : E \to \mathbb{R}$ with bounded support, the integral

$$\int f(x) \, d\mu(x)$$

is finite almost surely. For such f we consider the characteristic functional

$$\mathcal{L}_f = \mathbb{E}\left[\exp\left(i \int f(x) \, d\mu(x)\right)\right].$$

Since μ is boundedly finite, these expectations are well-defined for all bounded f with bounded support. By taking f of the form $f = \sum_{j=1}^{k} t_j \chi_{A_j}$, where $t_j \in \mathbb{R}$ and the sets A_j are disjoint and bounded, we see that the collection of all $\mathcal{L}_f$ (over such f) determines the distribution of the random measure.

2. Point processes

Let $\mathcal{N}^{\#}(E)$ be the subset of $\mathcal{M}^{\#}(E)$ consisting of those measures ν such that

$$\nu(\{x\}) \in \{0,1\} \quad \text{for all } x \in E.$$

Thus $\nu \in \mathcal{N}^{\#}(E)$ if and only if there exists a discrete subset $X \subset E$ such that

$$\nu = \sum_{x \in X} \delta_x.$$

Since ν is boundedly finite, we have $|X \cap A| < \infty$ for any bounded Borel set A, and therefore X has no accumulation points in E.

For such a measure $\nu = \sum_{x \in X} \delta_x$ and any function f with bounded support, we have

$$\int f(x)\, d\nu(x) = \sum_{x \in X \cap \operatorname{supp}(f)} f(x),$$

and the sum on the right-hand side has only finitely many terms.

Definition 33. *A* simple point process *on E is a probability measure on $\mathcal{M}^{\#}(E)$ such that*

$$\mathbb{P}\big(\nu \in \mathcal{N}^{\#}(E)\big) = 1.$$

Equivalently, it is a random element of $\mathcal{N}^{\#}(E)$.

Definition 34. *For a simple point process $\nu = \sum_{x \in X} \delta_x$ and a Borel set $A \subset E$, the* counting statistic $N(A)$ *is defined by*

$$N(A) = \nu(A) = |X \cap A|.$$

Note that $N(A)$ is almost surely finite for bounded A, but it can take any value in $\mathbb{N}_0$, and there is no a priori bound.

Proposition 35. *A point process is completely determined by its finite-dimensional distributions*

$$P_m(A_1, \dots, A_m; n_1, \dots, n_m) = \mathbb{P}\big(N(A_j) = n_j,\ j = 1, \dots, m\big),$$

for $m \in \mathbb{N}$, $n_1, \dots, n_m \in \mathbb{N}_0$, and bounded sets $A_1, \dots, A_m$.

Proof. See [**2**]. □

Theorem 36 (Kolmogorov–type consistency). *Let*

$$P_m(A_1, \dots, A_m; n_1, \dots, n_m),$$

for $m \in \mathbb{N}_0$, bounded sets $A_1, \dots, A_m$, and $n_1, \dots, n_m \in \mathbb{N}_0$, be a collection of probability distributions. These are the finite-dimensional distributions of a point process if and only if the following properties hold:

(1) ***Symmetry:*** *For every permutation $(i_1, \dots, i_m)$ of $\{1, \dots, m\}$,*

$$P_m(A_1, \dots, A_m; n_1, \dots, n_m) = P_m(A_{i_1}, \dots, A_{i_m}; n_{i_1}, \dots, n_{i_m}).$$

(2) ***Consistency:*** *For all $m \geq 0$ and bounded $A_1, \dots, A_{m+1}$,*

$$\sum_{r=0}^{\infty} P_{m+1}(A_1, \dots, A_m, A_{m+1}; n_1, \dots, n_m, r) = P_m(A_1, \dots, A_m; n_1, \dots, n_m).$$

(3) ***Additivity:*** *For each pair of disjoint Borel sets A_1 and A_2,*

$$P_3(A_1, A_2, A_1 \cup A_2; n_1, n_2, n_3) = 0 \quad \text{whenever } n_1 + n_2 \neq n_3.$$

(4) ***Continuity from above:*** *For $n \in \mathbb{N}$ and any sequence of bounded sets A_N such that $A_N \downarrow \emptyset$,*

$$\lim_{N\to\infty} P_1(A_N;0) = 1.$$

PROOF. See [**2**, Theorem 9.2.X]. [1] □

3. Factorial moments and correlation functions

We recall the Pochhammer symbol: for $a \in \mathbb{R}$ and $\ell \in \mathbb{N}$,

$$(a)_\ell = a(a-1)\cdots(a-\ell+1).$$

For $m \in \mathbb{N}$, we will construct a measure μ_m on E^m as follows. Let $1 \le k \le m$, let $A_1,\dots,A_k \subset E$ be pairwise disjoint bounded Borel sets, and let $\ell_1,\dots,\ell_k \in \mathbb{N}$ satisfy

$$\ell_1 + \ldots + \ell_k = m.$$

We set

$$A_j^{(\ell_j)} = \big\{(x_1^{(j)},\dots,x_{\ell_j}^{(j)}) \in A_j^{\ell_j} : x_p^{(j)} \neq x_q^{(j)} \text{ for } p \neq q\big\}.$$

Then we define the *factorial moments* by

$$\mu_m\big(A_1^{(\ell_1)} \times \cdots \times A_k^{(\ell_k)}\big) = \mathbb{E}\left[\prod_{j=1}^k \big(N(A_j)\big)_{\ell_j}\right].$$

It can be shown that μ_m has a unique extension to a measure on E^m. [2]

The factorial moments arise naturally from the following expansion. Let $A_1,\dots,A_k$ be bounded Borel sets and $z_1,\dots,z_k \in \mathbb{C}$. Then

$$\begin{aligned}
(10)\quad \mathbb{E}\left[\prod_{j=1}^k z_j^{N(A_j)}\right] &= \mathbb{E}\left[\prod_{j=1}^k (1+(z_j-1))^{N(A_j)}\right] \\
&= \mathbb{E}\left[\prod_{j=1}^k \left(\sum_{\ell_j=0}^{N(A_j)} \frac{\big(N(A_j)\big)_{\ell_j}}{\ell_j!}(z_j-1)^{\ell_j}\right)\right] \\
&= \mathbb{E}\left[\prod_{j=1}^k \left(\sum_{\ell_j=0}^{\infty} \frac{\big(N(A_j)\big)_{\ell_j}}{\ell_j!}(z_j-1)^{\ell_j}\right)\right] \\
&\overset{*}{=} \sum_{\ell_1,\dots,\ell_k=0}^{\infty} \frac{\prod_{j=1}^k (z_j-1)^{\ell_j}}{\ell_1!\cdots\ell_k!}\, \mathbb{E}\left[\prod_{j=1}^k \big(N(A_j)\big)_{\ell_j}\right] \\
&= \sum_{m=0}^\infty \frac{1}{m!} \sum_{\ell_1+\ldots+\ell_k=m} \binom{m}{\ell_1 \quad \ell_2 \quad \ldots \quad \ell_k} \Big(\prod_{j=1}^k (z_j-1)^{\ell_j}\Big) \mathbb{E}\left[\prod_{j=1}^k \big(N(A_j)\big)_{\ell_j}\right].
\end{aligned}$$

Thus $\mathbb{E}\big[\prod_{j=1}^k z_j^{N(A_j)}\big]$ can be viewed as a generating function for the factorial moments.

---

[1] In the discrete case, one of the conditions is redundant; here we state the general version.

[2] The collection of sets $A_1^{(\ell_1)} \times \cdots \times A_k^{(\ell_k)}$ forms a semi-ring. The extension to a measure is a standard exercise that is worth working out.

The step indicated by $\stackrel{*}{=}$ requires justification, since we interchange an infinite sum and an expectation. For this we need an extra assumption.

LEMMA 37. *Assume that, for every bounded $A \subset E$, we have*

$$\frac{\mu_{m+1}\big(A^{(m+1)}\big)}{m\,\mu_m\big(A^{(m)}\big)} \longrightarrow 0 \qquad \text{as } m \to \infty,$$

where $A^{(m)}$ denotes the subset of A^m of m-tuples of distinct points. Then, for bounded $A_1, \ldots, A_k$ and $z_1, \ldots, z_k \in \mathbb{C}$, the expectation

$$\mathbb{E}\left[\prod_{j=1}^{k} z_j^{N(A_j)}\right]$$

is well-defined and finite, and

$$\mathbb{E}\left[\prod_{j=1}^{k} z_j^{N(A_j)}\right] = \sum_{m=0}^{\infty} \frac{1}{m!} \sum_{\ell_1+\ldots+\ell_k=m} \binom{m}{\ell_1 \quad \ell_2 \quad \ldots \quad \ell_k} \Big(\prod_{j=1}^{k} (z_j-1)^{\ell_j}\Big) \times \mu_m\big(A_1^{(\ell_1)} \times \cdots \times A_k^{(\ell_k)}\big).$$

The series converges uniformly for $(z_1, \ldots, z_k)$ in compact subsets of $\mathbb{C}^k$.

DEFINITION 38. *We say that a point process has m-*point correlation functions (with respect to a reference measure μ_0) *if there exists a measure μ_0 on E and symmetric functions $\rho_m : E^m \to [0, \infty)$, $m = 1, 2, \ldots$, such that*

$$d\mu_m(x_1, \ldots, x_m) = \rho_m(x_1, \ldots, x_m)\, d\mu_0(x_1) \cdots d\mu_0(x_m), \quad m = 1, 2, \ldots.$$

*In this case, the functions ρ_m are called the m-*point correlation functions *of the point process, with respect to the reference measure μ_0.*

In the discrete case, the correlation functions ρ_m are probabilities (see the previous chapter). If E is not discrete, this is no longer true, but the ρ_m still have a natural probabilistic interpretation.

PROPOSITION 39. *If the m-point correlation functions exist, then for distinct points $x_1, \ldots, x_m \in E$,*

$$\rho_m(x_1, \ldots, x_m) = \lim_{\Delta \downarrow 0} \frac{\mathbb{P}(\text{each neighbourhood } [x_j, x_j+\Delta] \text{ contains at least one point})}{\prod_{j=1}^{m} \mu_0([x_j, x_j+\Delta])},$$

whenever the limit exists and the intervals are chosen in a suitable local coordinate system.

PROOF. Here comes a proof. □

THEOREM 40. *Suppose that the m-point correlation functions exist with μ_0 as a boundedly finite reference measure, and that, for every bounded $A \subset E$,*

$$\frac{\mu_{m+1}\big(A^{(m+1)}\big)}{m\,\mu_m\big(A^{(m)}\big)} \longrightarrow 0 \qquad \text{as } m \to \infty. \tag{11}$$

Then for any bounded function $g : E \to \mathbb{C}$ with bounded support, we have

$$\mathbb{E}\left[\prod_{x \in X} (1+g(x))\right] = \sum_{m=0}^{\infty} \frac{1}{m!} \int_{E^m} \prod_{j=1}^{m} g(x_j)\, \rho_m(x_1, \ldots, x_m)\, d\mu_0(x_1) \cdots d\mu_0(x_m).$$

The series converges absolutely and locally uniformly in g.

PROOF. It suffices (why?) to prove the statement for functions of the form

$$g = \sum_{j=1}^{k} (z_j - 1)\chi_{A_j}$$

with bounded disjoint sets $A_1, \ldots, A_k$ and complex numbers $z_1, \ldots, z_k$. Then

$$\Big(\prod_{j=1}^{k} (z_j - 1)^{\ell_j}\Big)\mathbb{E}\left[\prod_{j=1}^{k} \big(N(A_j)\big)_{\ell_j}\right] \\ = \int_{A_1^{(\ell_1)}\times\cdots\times A_k^{(\ell_k)}} \prod_{j=1}^{m} (z_{c(j)} - 1)\, \rho_m(x_1, \ldots, x_m)\, d\mu_0(x_1)\cdots d\mu_0(x_m),$$

where $c(j)$ indicates to which block $A_{c(j)}$ the variable x_j belongs. Since $g(x) = z_j - 1$ for $x \in A_j$, this becomes

$$\int_{A_1^{(\ell_1)}\times\cdots\times A_k^{(\ell_k)}} \prod_{j=1}^{m} g(x_j)\, \rho_m(x_1, \ldots, x_m)\, d\mu_0(x_1)\cdots d\mu_0(x_m).$$

Now the integrand $\prod_{j=1}^{m} g(x_j)\rho_m(x_1, \ldots, x_m)$ is symmetric in $(x_1, \ldots, x_m)$, so

$$\sum_{\ell_1+\ldots+\ell_k=m} \binom{m}{\ell_1 \quad \ell_2 \quad \ldots \quad \ell_k} \int_{A_1^{(\ell_1)}\times\cdots\times A_k^{(\ell_k)}} \prod_{j=1}^{m} g(x_j)\, \rho_m(x_1, \ldots, x_m)\, d\mu_0^m \\ = \int_{E^{(m)}} \prod_{j=1}^{m} g(x_j)\, \rho_m(x_1, \ldots, x_m)\, d\mu_0^m,$$

where $E^{(m)}$ denotes the set of m-tuples of distinct points in E. Inserting this into (10) yields the desired expansion for such g, and by approximation for all bounded g with bounded support. □

4. Determinantal point processes

DEFINITION 41. *A* determinantal point process *is a simple point process for which the m-point correlation functions exist for all $m \geq 1$ and are given by*

$$\rho_m(x_1, \ldots, x_m) = \det\big(K(x_i, x_j)\big)_{i,j=1}^{m}, \qquad m = 1, 2, \ldots,$$

for some function $K : E \times E \to \mathbb{C}$. We call K a correlation kernel *for the determinantal point process.*

If K is a correlation kernel for a determinantal point process, then it is clear from the definition that all information about the process is encoded in K. However, the correlation kernel is not unique. Indeed, for any nowhere vanishing function $G : E \to \mathbb{C} \setminus \{0\}$,

$$\widetilde{K}(x, y) = K(x, y)\, \frac{G(x)}{G(y)}$$

is also a correlation kernel for the same determinantal point process.

To express the quantity

$$\Phi(g) = \mathbb{E}\left[\prod_{x\in X}(1+g(x))\right]$$

in terms of a determinant, we need additional assumptions on K. These assumptions will be imposed on the restrictions

$$K_A(x,y) = \chi_A(x)\,K(x,y)\,\chi_A(y), \qquad x,y\in E,$$

for bounded sets $A\subset E$.

Assumption 1. *The simplest (but strong) assumption is that K is continuous on $E\times E$ when E is not discrete. In that case, many technical subtleties disappear, and most examples of interest to us satisfy this condition. For a fully general theory of determinantal point processes, however, this assumption is too restrictive.*

Assumption 2. *A common assumption in the literature is that K defines a locally trace class operator. That is, for each bounded $A\subset E$, the operator $\mathcal{K}_A$ on $L_2(A,\mu_0)$,*

$$\mathcal{K}_A f(x) = \int_A K(x,y)\,f(y)\,d\mu_0(y)$$

is trace class. This allows one to use the well-developed theory of trace class operators and their determinants and traces. The drawback is that, if the diagonal $\{(x,x) : x\in E\}$ has μ_0-measure zero in $E\times E$, then the operator is insensitive to the values of K on the diagonal, while for point processes the diagonal values $K(x,x)$ encode the local intensities and are therefore very important. Although there are ways to handle this issue, we find that the purely operator-theoretic approach becomes unnecessarily complicated for our purposes.

Assumption 3. *We now introduce a more flexible assumption.*

For a measurable function $R : E\times E\to\mathbb{C}$, define

$$\|R\|_S = \max\left(\left(\iint_{E\times E}|R(x,y)|^2\,d\mu_0(x)d\mu_0(y)\right)^{1/2},\ \int_E |R(x,x)|\,d\mu_0(x)\right).$$

Strictly speaking, to make this into a norm we pass to equivalence classes: declare $R_1\equiv R_2$ if

$$R_1(x,y) = R_2(x,y) \quad \text{for } \mu_0\times\mu_0\text{-a.e. } (x,y)$$

and

$$R_1(x,x) = R_2(x,x) \quad \text{for } \mu_0\text{-a.e. } x.$$

The space of equivalence classes with finite $\|\cdot\|_S$ is a normed vector space, which we denote by $\mathcal{B}$. It is a Banach space; in fact it is naturally isomorphic to a subspace of $L_2(E\times E,\mu_0\otimes\mu_0)\oplus L_1(E,\mu_0)$.

Our assumption on the correlation kernel is that $\|K_A\|_S<\infty$ for each bounded $A\subset E$.

Remark 42. *A function R with $\|R\|_S<\infty$ need not define a trace class operator. There exist continuous kernels K for which the associated integral operator is not trace class. We refer to the appendix for further discussion. From the viewpoint of determinantal point processes, this is not a defect: the diagonal values $K(x,x)$ are crucial and should be part of the data.*

In what follows we will mostly work under the first assumption, namely that the correlation kernel K is continuous.

THEOREM 43. *Let X be a determinantal point process with continuous correlation kernel K and reference measure μ_0. Then, for any bounded function $g : E \to \mathbb{C}$ with bounded support,*

$$\mathbb{E}\left[\prod_{x\in X}(1+g(x))\right] = \sum_{m=0}^{\infty}\frac{1}{m!}\int_{E^m}\prod_{j=1}^{m} g(x_j)\det\bigl(K(x_i,x_j)\bigr)_{i,j=1}^{m}\,d\mu_0(x_1)\cdots d\mu_0(x_m),$$

and the series converges absolutely.

PROOF. We only need to verify the assumption (11) from Theorem 40. This can be done using Hadamard's inequality for determinants:

$$\left|\det(K(x_i,x_j))_{i,j=1}^{m}\right| \le \prod_{i=1}^{m}\left(\sum_{j=1}^{m}|K(x_i,x_j)|^2\right)^{1/2}.$$

Using this inequality and the boundedness of K on bounded sets, one shows that

$$\begin{aligned}\mu_m(A^{(m)}) = \int_{A^m}\det\bigl(K(x_i,x_j)\bigr)_{i,j=1}^{m} d\mu_0(x_1)\cdots d\mu_0(x_m)&\\ \le m^{m/2}\,\|K_A\|_{\Phi}^{m}\,(\mu_0(A))^{m/2}&,\end{aligned}$$

where $\|\cdot\|_\Phi$ is the uniform norm used in the Fredholm theory (see the appendix). From this bound one checks that (11) holds, and the theorem follows from Theorem 40. □

Under the continuity assumption, the right-hand side is precisely the Fredholm determinant of the integral operator with kernel $g(x)K(x,y)$, acting on $L_2(S_g,\mu_0)$, where S_g is the support of g. We write

$$\mathbb{E}\left[\prod_{x\in X}(1+g(x))\right] = \det\nolimits_\Phi\bigl(I+gK_{S_g}\bigr),$$

or, with a slight abuse of notation, simply $\det_\Phi(I+gK)$.

We refer to the appendix for a rigorous treatment of the Fredholm determinant and its properties. For now, we only use that the Fredholm determinant enjoys the usual identities one expects from determinants of finite matrices.

An important class of observables are the *gap probabilities*.

DEFINITION 44. *For $A \subset E$ we define the* gap probability *as*

$$\mathbb{P}(N(A)=0).$$

PROPOSITION 45. *For any bounded Borel set $A \subset E$,*

$$\mathbb{P}(N(A)=0) = \det\nolimits_\Phi(I-K_A).$$

PROOF. Apply Theorem 43 with the choice $g(x) = -\chi_A(x)$. □

We can also define a trace for continuous kernels $R : E\times E \to \mathbb{C}$ by

$$\mathrm{Tr}_\Phi R = \int_E R(x,x)\,d\mu_0(x),$$

whenever the integral is absolutely convergent (e.g. if $R(x,x)$ has bounded support).

Proposition 46. *For the counting statistic $N(A)$ we have:*

- $\mathbb{E}[N(A)] = \mathrm{Tr}_\Phi K_A = \int_A K(x,x)\, d\mu_0(x)$,
- $\mathrm{V}ar[N(A)] = \mathrm{Tr}_\Phi K_A - \mathrm{Tr}_\Phi K_A^2 = \int_A K(x,x)\, d\mu_0(x) - \int_A \int_A K(x,y)K(y,x)\, d\mu_0(x) d\mu_0(y)$.

Proof. First note that

$$\mathbb{E}[N(A)] = \frac{d}{dz}\bigg|_{z=1} \mathbb{E}[z^{N(A)}].$$

For $|z-1|$ small, we have the expansion

$$\det_\Phi\big(I + (z-1)K_A\big) = \exp\left(\sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j} (z-1)^j \mathrm{Tr}_\Phi K_A^j\right),$$

and differentiation at $z = 1$ gives the first identity. The variance follows similarly by differentiating twice. □

5. Convergence of determinantal point processes

(Here one can discuss convergence of kernels and the corresponding convergence in distribution of the associated determinantal point processes; we omit the details.)

6. Examples

We now discuss some examples of determinantal point processes. A particularly important class comes from Hermitian kernels. [3] Let K be a continuous Hermitian kernel on $E \times E$, i.e.

$$K(y,x) = \overline{K(x,y)} \quad \text{for all } x, y \in E.$$

Theorem 47. *Let K be a continuous Hermitian kernel such that the associated integral operator $\mathcal{K}$ on $L_2(E,\mu_0)$ is bounded. Then K is a correlation kernel for a determinantal point process if and only if the spectrum $\sigma(\mathcal{K})$ is contained in $[0,1]$.*

Proof. Here comes a proof (compare with the discrete case in the previous chapter; the general case is based on the same ideas). □

Theorem 48. *Let K be a continuous Hermitian kernel such that the associated integral operator $\mathcal{K}$ is bounded. Then the total number of points $N(E)$ is almost surely either finite or infinite, and*

$$\mathbb{P}\big(N(E) = \infty\big) = 0 \quad \text{if } \mathrm{Tr}_\Phi K = \int_E K(x,x)\, d\mu_0(x) < \infty,$$

and

$$\mathbb{P}\big(N(E) = \infty\big) = 1 \quad \text{if } \mathrm{Tr}_\Phi K = \int_E K(x,x)\, d\mu_0(x) = \infty.$$

Proof. Here comes a proof (see the discrete case in the previous chapter). □

Theorem 49. *Let K be a continuous Hermitian kernel such that the associated integral operator $\mathcal{K}$ is bounded. Then the following statements are equivalent:*

(1) *There exists $n \in \mathbb{N}$ such that $N(E) = n$ almost surely.*

[3] There are also important examples for which no Hermitian correlation kernel exists.

(2) *The operator $\mathcal{K}$ is an orthogonal projection on an n-dimensional subspace, i.e. $\mathcal{K}^2 = \mathcal{K}$ and $\operatorname{rank}\mathcal{K} = n$.*

PROOF. Here comes a proof (see the discrete case in the previous chapter). □

THEOREM 50. *Let $E = \mathbb{R}$ and let $K : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a continuous Hermitian kernel such that $\sigma(\mathcal{K}) \subset [0, 1]$. Assume that, for some $s \in \mathbb{R}$,*

$$\int_s^\infty K(x,x)\, d\mu_0(x) < \infty.$$

Then the point process has a largest point $\xi_{\max}$ with probability 1, and

$$\mathbb{P}(\xi_{\max} \le s) = \det_\Phi\big(I - K_{[s,\infty)}\big).$$

SKETCH OF PROOF. Since

$$\mathbb{E}\big[N([s,\infty))\big] = \int_s^\infty K(x,x)\, d\mu_0(x) < \infty,$$

the probability of having infinitely many points in $[s, \infty)$ is zero. Thus there must be a largest point. Note that

$$\mathbb{P}(\xi_{\max} \le s) = \mathbb{P}(N([s,\infty)) = 0) = \mathbb{E}\left[\prod_{x\in X}(1 + g(x))\right],$$

where $g(x) = -1$ for $x \ge s$ and $g(x) = 0$ otherwise. The statement then follows from Theorem 43, after justifying the passage to the limit for the unbounded set $[s, \infty)$ via an exhaustion by bounded intervals. □

6.1. The sine process. The sine process is the determinantal point process on $\mathbb{R}$ with correlation kernel

$$K_{\text{sine}}(x,y) = \frac{\sin \pi(x-y)}{\pi(x-y)} = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{it(x-y)}\, dt.$$

The corresponding operator on $L_2(\mathbb{R})$ is convolution with the sinc function. Its Fourier transform is the indicator function of $[-\pi, \pi]$, so K_{sine} is the kernel of an orthogonal projection. Hence it defines a determinantal point process.

The sine process is translation invariant.

6.2. The Airy point process. The Airy function $\operatorname{Ai}(x)$ is the unique solution of

$$y''(x) = x\, y(x),$$

satisfying the asymptotics

$$y(x) \sim \frac{1}{2\sqrt{\pi}}\, x^{-1/4} \exp\left(-\frac{2}{3}x^{3/2}\right), \qquad x \to +\infty.$$

It can be shown that

$$\operatorname{Ai}'(x) = \mathcal{O}\big(x^{1/4}\exp(-\tfrac{2}{3}x^{3/2})\big), \qquad \operatorname{Ai}''(x) = \mathcal{O}\big(x^{3/4}\exp(-\tfrac{2}{3}x^{3/2})\big)$$

as $x \to +\infty$.

Define the Airy kernel

$$K_{\text{Airy}}(x,y) = \begin{cases} \dfrac{\operatorname{Ai}(x)\operatorname{Ai}'(y) - \operatorname{Ai}(y)\operatorname{Ai}'(x)}{x-y}, & x \ne y, \\ (\operatorname{Ai}'(x))^2 + \operatorname{Ai}(x)\operatorname{Ai}''(x), & x = y. \end{cases}$$

One can show that K_{Airy} is the kernel of an orthogonal projection, hence it defines a determinantal point process, called the *Airy point process*.

Since

$$\int_s^\infty K_{\text{Airy}}(x,x)\,dx < \infty \quad \text{for each } s \in \mathbb{R},$$

the corresponding determinantal point process has a largest point $\xi_{\max}$ with probability one. The distribution

$$\mathbb{P}(\xi_{\max} \le s) = \det_\Phi\big(I - K_{\text{Airy}}\big)_{L_2(s,\infty)}$$

is called the *Tracy–Widom distribution* and is usually denoted by $F_{TW}(s)$.

6.3. Orthogonal polynomial ensembles. Let μ be a measure on $\mathbb{R}$ with finite moments:

$$\int |x|^k\,d\mu(x) < \infty \quad \text{for all } k \in \mathbb{N}_0.$$

We define the family of orthonormal polynomials $\{p_k\}_{k\ge 0}$ by the conditions:

- p_k is a polynomial of degree k with positive leading coefficient;
- $\displaystyle\int p_k(x)p_j(x)\,d\mu(x) = \delta_{jk}$ for $j = 0, \ldots, k$.

Then the reproducing kernel

$$K_n(x,y) = \sum_{k=0}^{n-1} p_k(x)\,p_k(y)$$

(together with the reference measure μ) defines a determinantal point process on $\mathbb{R}$. Since K_n is the kernel of a self-adjoint projection of rank n, this determinantal point process has exactly n points almost surely.

CHAPTER 4

Biorthogonal ensembles

1. Definition

Definition 51. *Let $\mathbb{X}$ be a complete metric space and μ a Borel measure on $\mathbb{X}$. A probability measure $\mathbb{P}_n$ on $\mathbb{X}^n$ is called a* **biorthogonal ensemble** *of size n (with reference measure μ) if and only if there exist*

$$\phi_1, \dots, \phi_n \in L_2(\mathbb{X}, \mu), \qquad \psi_1, \dots, \psi_n \in L_2(\mathbb{X}, \mu),$$

such that

$$d\mathbb{P}_n(x_1, \dots, x_n) = \frac{1}{Z_n} \det\big(\phi_j(x_k)\big)_{j,k=1}^n \det\big(\psi_j(x_k)\big)_{j,k=1}^n \prod_{k=1}^n d\mu(x_k),$$

for some normalizing constant $Z_n > 0$.

Note that the families of functions $\{\phi_j\}_{j=1}^n$ and $\{\psi_j\}_{j=1}^n$ in the definition are not unique. Indeed, by performing row operations in the determinants one can replace each function ϕ_k (resp. ψ_k) with a linear combination of the ϕ_j's (resp. ψ_j's), possibly changing the normalization constant, as long as all the new functions remain linearly independent. This means that a biorthogonal ensemble is determined by the n–dimensional subspaces

$$V_\phi = \mathrm{span}\{\phi_1, \dots, \phi_n\}, \qquad V_\psi = \mathrm{span}\{\psi_1, \dots, \psi_n\}.$$

The functions ϕ_j and ψ_k in Definition 51 can be replaced by any other systems of functions $\tilde{\phi}_1, \dots, \tilde{\phi}_n$ and $\tilde{\psi}_1, \dots, \tilde{\psi}_n$ that span the spaces V_ϕ and V_ψ respectively.

Lemma 52. *With the notation above,*

$$Z_n = n! \det G, \qquad G_{ij} = \int_{\mathbb{X}} \phi_i(x)\, \psi_j(x)\, d\mu(x).$$

Proof. This follows directly from Andreiéf's identity stated in Lemma 65. □

A biorthogonal ensemble is a probability measure on $\mathbb{X}^n$ and hence it defines a point process on $\mathbb{X}$ with exactly n points.

Theorem 53. *Let $\mathbb{P}_n$ be a biorthogonal ensemble given by*

$$d\mathbb{P}_n(x_1, \dots, x_n) = \frac{1}{Z_n} \det\big(\phi_j(x_k)\big)_{j,k=1}^n \det\big(\psi_j(x_k)\big)_{j,k=1}^n \prod_{k=1}^n d\mu(x_k),$$

and let

$$G_{ij} = \int_{\mathbb{X}} \phi_i(x)\, \psi_j(x)\, d\mu(x), \qquad i, j = 1, \dots, n.$$

Assume that G is invertible. Then this point process is determinantal with correlation kernel

$$K(x,y) = \sum_{i,j=1}^{n} \phi_i(x)\,(G^{-1})_{ij}\,\psi_j(y). \tag{12}$$

PROOF. Let g be a function with bounded support on $\mathbb{X}$. Then

$$\mathbb{E}\Big[\prod_{k=1}^{n}(1+g(x_k))\Big] = \frac{1}{Z_n}\int_{\mathbb{X}}\cdots\int_{\mathbb{X}}\prod_{k=1}^{n}(1+g(x_k))\det\big(\phi_j(x_k)\big)_{j,k=1}^{n}\det\big(\psi_j(x_k)\big)_{j,k=1}^{n}\prod_{k=1}^{n}d\mu(x_k).$$

By Andreiéf's identity (Lemma 65) with

$$f_j(x) = (1+g(x))\,\phi_j(x), \qquad g_j(x) = \psi_j(x),$$

we obtain

$$\begin{aligned}\mathbb{E}\Big[\prod_{k=1}^{n}(1+g(x_k))\Big] &= \frac{1}{Z_n}\,n!\det\Big(\int_{\mathbb{X}}(1+g(x))\,\phi_i(x)\,\psi_j(x)\,d\mu(x)\Big)_{i,j=1}^{n}\\ &= \frac{1}{\det G}\det\big(G+H\big),\end{aligned}$$

where

$$H_{ij} = \int_{\mathbb{X}} g(x)\,\phi_i(x)\,\psi_j(x)\,d\mu(x).$$

Thus

$$\mathbb{E}\Big[\prod_{k=1}^{n}(1+g(x_k))\Big] = \det\big(I+G^{-1}H\big).$$

By Lemma 69, this is the Fredholm determinant of an operator with kernel

$$K_g(x,y) = g(x)\,K(x,y), \qquad K(x,y) = \sum_{i,j=1}^{n}\phi_i(x)\,(G^{-1})_{ij}\,\psi_j(y).$$

Comparing with the general determinantal formula for generating functionals (see Theorem 9), we conclude that K is indeed the correlation kernel, proving (12). □

LEMMA 54. *The kernel K of a biorthogonal ensemble is the integral kernel of a projection (not necessarily orthogonal), i.e. the associated operator $\mathcal{K}$ satisfies $\mathcal{K}^2 = \mathcal{K}$.*

PROOF. The operator $\mathcal{K}$ has finite rank at most n and its range coincides with the span of the functions $\phi_1,\ldots,\phi_n$. From the representation

$$K(x,y) = \sum_{i,j=1}^{n}\phi_i(x)\,(G^{-1})_{ij}\,\psi_j(y)$$

and the definition of G, a direct computation shows that $\mathcal{K}^2 = \mathcal{K}$. □

To compute the correlation kernel for a biorthogonal ensemble, we need, in principle, to invert the Gram matrix G. Note that G is an $n\times n$ matrix and in applications n is often very large, so inverting G explicitly can be difficult.

Instead of trying to invert G, we can use the observation at the beginning of this section and choose different systems $\{\tilde{\phi}_j\}_{j=1}^n$ and $\{\tilde{\psi}_j\}_{j=1}^n$ such that

$$\tilde{G}_{jk} = \int_{\mathbb{X}} \tilde{\phi}_j(x)\, \tilde{\psi}_k(x)\, d\mu(x) = \delta_{jk}.$$

We say that $\{\tilde{\phi}_j\}$ and $\{\tilde{\psi}_j\}$ are *biorthogonal*. In this basis the Gram matrix is the identity and hence trivial to invert. This is best seen through examples.

REMARK 55. *It is always possible to find such biorthogonal functions. Indeed, since G is invertible we can take*

$$\tilde{\phi}_j = \sum_{k=1}^n (G^{-1})_{jk}\, \phi_k, \qquad \tilde{\psi}_j = \psi_j,$$

which yields $\int \tilde{\phi}_j(x)\, \tilde{\psi}_k(x)\, d\mu(x) = \delta_{jk}$.

We also note that there is no unique way of constructing biorthogonal systems. Different choices may have different advantages.

2. Orthogonal polynomial ensembles

Let μ be a measure on $\mathbb{R}$ with all moments finite, i.e.

$$\int_{\mathbb{R}} |x|^m\, d\mu(x) < \infty, \qquad m = 0, 1, 2, \ldots$$

Denote by $S(\mu) = \operatorname{supp} \mu$ the support of μ. Consider the probability measure on $S(\mu)^n$ defined by

$$\frac{1}{Z_n} \prod_{1 \le i < j \le n} (x_i - x_j)^2 \prod_{j=1}^n d\mu(x_j).$$

The Vandermonde determinant can be written as

$$\prod_{1 \le i < j \le n} (x_i - x_j) = \det\bigl(x_i^{j-1}\bigr)_{i,j=1}^n,$$

so the density can be written as

$$\frac{1}{Z_n} \det\bigl(x_i^{j-1}\bigr)_{i,j=1}^n \det\bigl(x_i^{j-1}\bigr)_{i,j=1}^n \prod_{j=1}^n d\mu(x_j).$$

In other words, this is a biorthogonal ensemble with

$$\phi_j(x) = x^{j-1}, \qquad \psi_j(x) = x^{j-1}, \qquad j = 1, \ldots, n.$$

Now, by performing column operations on the matrices inside the determinants, we can rewrite this measure as

$$\frac{1}{Z_n} \det\bigl(p_{j-1}(x_i)\bigr)_{i,j=1}^n \det\bigl(p_{j-1}(x_i)\bigr)_{i,j=1}^n \prod_{j=1}^n d\mu(x_j),$$

where p_j are the orthonormal polynomials with respect to μ, so that

$$G_{ij} = \int_{\mathbb{R}} p_i(x)\, p_j(x)\, d\mu(x) = \delta_{ij}.$$

Applying Theorem 53 we therefore obtain a determinantal point process with kernel

$$K_n(x, y) = \sum_{j=0}^{n-1} p_j(x)\, p_j(y),$$

and this determinantal point process is precisely the orthogonal polynomial ensemble introduced earlier.

CHAPTER 5

Non-intersecting path ensembles

1. Non-intersecting paths

Let $G = (V, E)$ be a directed graph. Throughout we assume that G is acyclic, so that no directed path can visit the same vertex more than once. Each edge $e \in E$ is assigned a strictly positive weight $w(e) \in (0, \infty)$.

A (directed) path π in G is a finite sequence of vertices

$$\pi = (v_0, v_1, \dots, v_k),$$

such that $(v_{i-1}, v_i) \in E$ for all $i = 1, \dots, k$. The weight of a path is defined multiplicatively by

$$w(\pi) = \prod_{i=1}^{k} w(v_{i-1}, v_i).$$

Fix an integer $n \geq 1$ and choose two ordered n-tuples of distinct vertices

$$P_1, \dots, P_n \qquad \text{and} \qquad Q_1, \dots, Q_n$$

in V, interpreted as starting points and ending points. For $P, Q \in V$, we denote by $\Pi(P, Q)$ the set of all directed paths in G starting at P and ending at Q.

A path ensemble is an n-tuple $\boldsymbol{\pi} = (\pi_1, \dots, \pi_n)$ with $\pi_j \in \Pi(P_j, Q_j)$. The ensemble is called *non-intersecting* if no two paths share a vertex, that is,

$$\pi_i \cap \pi_j = \varnothing \qquad \text{for all } i \neq j,$$

where intersections are understood at the level of vertices. We write

$$\Pi_{\text{ni}}\big((P_1, \dots, P_n), (Q_1, \dots, Q_n)\big)$$

for the set of all non-intersecting path ensembles connecting the prescribed starting points to the prescribed ending points.

The weight of an ensemble $\boldsymbol{\pi} = (\pi_1, \dots, \pi_n)$ is defined by

$$w(\boldsymbol{\pi}) = \prod_{j=1}^{n} w(\pi_j).$$

Assuming that the total weight of all non-intersecting ensembles is finite, we introduce the partition function

$$Z = \sum_{\boldsymbol{\pi} \in \Pi_{\text{ni}}} w(\boldsymbol{\pi}),$$

and define a probability measure on Π_{ni} by

$$\mathbb{P}(\boldsymbol{\pi}) = \frac{1}{Z} w(\boldsymbol{\pi}).$$

A key role in the analysis of non-intersecting paths is played by the single-path transition function g defined by

$$g(P,Q) = \sum_{\pi\in\Pi(P,Q)} w(\pi), \qquad P,Q\in V,$$

which record the total weight of all directed paths between two vertices. The fundamental connection between single-path and multi-path quantities is provided by the Lindström–Gessel–Viennot theorem.

THEOREM 56 (Lindström–Gessel–Viennot). *Let $G=(V,E)$ be a directed acyclic graph with edge-weights $w\colon E\to(0,\infty)$. Fix vertices $P_1,\dots,P_n$ and $Q_1,\dots,Q_n$ in V. Then*

$$\det\big[\, g(P_i,Q_j)\,\big]_{i,j=1}^{n} = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \sum_{\boldsymbol{\pi}\in\Pi_{\mathrm{ni}}\left((P_1,\dots,P_n),(Q_{\sigma(1)},\dots,Q_{\sigma(n)})\right)} w(\boldsymbol{\pi}). \tag{13}$$

PROOF. Expanding the determinant gives

$$\det\big[\, g(P_i,Q_j)\,\big]_{i,j=1}^{n} = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\prod_{i=1}^{n} g(P_i,Q_{\sigma(i)}).$$

Using the definition of g and distributivity, we obtain

$$\prod_{i=1}^{n} g(P_i,Q_{\sigma(i)}) = \sum_{(\pi_1,\dots,\pi_n)} \prod_{i=1}^{n} w(\pi_i),$$

where the sum runs over all n-tuples with $\pi_i\in\Pi(P_i,Q_{\sigma(i)})$. Thus the determinant equals a signed sum over all such path families. Note that this include both families of paths that are non-intersecting and intersecting, and it remains to show that the sum over intersecting paths vanishes. To this end we construct a sign-reversing involution on the set of intersecting families.

Since G is acyclic, fix a topological ordering of its vertices. Given an intersecting family, let v be the earliest vertex (in this order) that lies on at least two paths, and let $a<b$ be the smallest indices such that $v\in\pi_a\cap\pi_b$. Write

$$\pi_a = \alpha_a\cdot\beta_a, \qquad \pi_b = \alpha_b\cdot\beta_b,$$

where α_a,α_b are the initial segments ending at v and β_a,β_b are the remaining tails. Define new paths by swapping the tails,

$$\pi'_a = \alpha_a\cdot\beta_b, \qquad \pi'_b = \alpha_b\cdot\beta_a,$$

and let $\sigma' = \sigma\circ(a\,b)$. This preserves total weight and reverses the sign, while remaining an involution. Hence all intersecting families cancel pairwise, and only non-intersecting ensembles contribute, yielding (13). □

COROLLARY 57 (Partition function). *Assume that any non-intersecting path ensemble connecting $\{P_1,\dots,P_n\}$ to $\{Q_1,\dots,Q_n\}$ necessarily matches P_i to Q_i for each i. Then*

$$Z = \det\big[\, g(P_i,Q_j)\,\big]_{i,j=1}^{n}.$$

PROOF. Under the stated assumption, all terms in (13) corresponding to non-trivial permutations vanish, and the remaining term is precisely the partition function Z. □

2. Graphs obtained by gluing columns

Fix an integer $N \geq 1$. We consider directed graphs whose vertex set is the full integer strip

$$V = \{0, 1, \ldots, N\} \times \mathbb{Z}.$$

Edges are added column-by-column. More precisely, for each $x \in \{0, 1, \ldots, N-1\}$ we specify a set of directed edges between the slices $\{x\} \times \mathbb{Z}$ and $\{x+1\} \times \mathbb{Z}$, and possibly also vertical edges inside the right slice $\{x+1\} \times \mathbb{Z}$. Gluing the columns means that the slice $\{x+1\} \times \mathbb{Z}$ serves simultaneously as the right boundary of column x and the left boundary of column $x+1$.

We will use four elementary column types. In all cases $y \in \mathbb{Z}$.

Type I (Bernoulli up-step). The only edges in column x are

$$(x, y) \to (x+1, y), \qquad (x, y) \to (x+1, y+1).$$

Thus a path crossing the column either goes straight or steps up by one.

Type II (Bernoulli down-step). The only edges in column x are

$$(x, y) \to (x+1, y), \qquad (x, y) \to (x+1, y-1).$$

Thus a path crossing the column either goes straight or steps down by one.

Type III (horizontal + vertical down at the right slice). The column contains all horizontal edges

$$(x, y) \to (x+1, y),$$

and, in addition, all vertical edges at the right slice pointing down,

$$(x+1, y) \to (x+1, y-1).$$

Equivalently, after entering the slice $x+1$, a path may move down by an arbitrary number of steps before proceeding further to the right in subsequent columns.

Type IV (horizontal + vertical up at the right slice). The column contains all horizontal edges

$$(x, y) \to (x+1, y),$$

and, in addition, all vertical edges at the right slice pointing up,

$$(x+1, y) \to (x+1, y+1).$$

Equivalently, after entering the slice $x+1$, a path may move up by an arbitrary number of steps before proceeding further to the right.

By choosing for each $x \in \{0, \ldots, N-1\}$ one of the four types above, we obtain a directed graph on $\{0, \ldots, N\} \times \mathbb{Z}$ built by gluing columns. Since every edge either increases x by one (horizontal/diagonal) or keeps x fixed while changing y (vertical edges within a slice), and since vertical edges are oriented consistently (up or down within a given slice), the resulting graph is directed and acyclic.

EXAMPLE 58 (Alternating Type I and Type III columns). *As a concrete illustration of the column-gluing construction, consider a graph on*

$$V = \{0, 1, \ldots, N\} \times \mathbb{Z}$$

obtained by gluing columns in an alternating pattern: for even x we use Type I on the strip from x to $x+1$, and for odd x we use Type III.

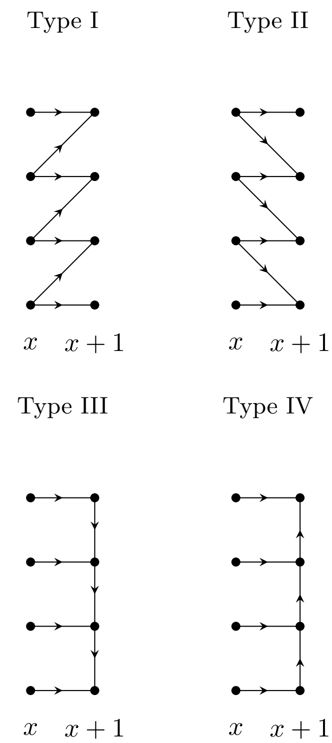

FIGURE 1. Four elementary column types used to build directed graphs on $\{0, \ldots, N\} \times \mathbb{Z}$.

EXAMPLE 59 (N columns of Type IV followed by N columns of Type III). *Fix an integer $N \geq 1$ and consider the directed graph on*

$$V = \{0, 1, \ldots, 2N\} \times \mathbb{Z}$$

obtained by gluing $2N$ columns as follows: for $x = 0, 1, \ldots, N-1$ the column between x and $x+1$ is of Type IV, and for $x = N, N+1, \ldots, 2N-1$ the column between x and $x+1$ is of Type III. In other words, we use a block of N Type IV columns followed by a block of N Type III columns. This produces a directed acyclic graph with a natural left-to-right direction and a single interface at $x = N$.

2.1. Edge weights. We now equip the glued graph with edge weights. Throughout, we impose the convention that *all horizontal edges have weight* 1. Only the remaining (non-horizontal) edges carry nontrivial weights, which may depend on the position (x, y) and on the column type.

Consider a column between the slices x and $x+1$, and let $y \in \mathbb{Z}$.

Type I. The diagonal up-edges

$$(x, y) \to (x+1, y+1)$$

are assigned weights

$$w\big((x, y) \to (x+1, y+1)\big) = \alpha_{x,y} \in (0, \infty),$$

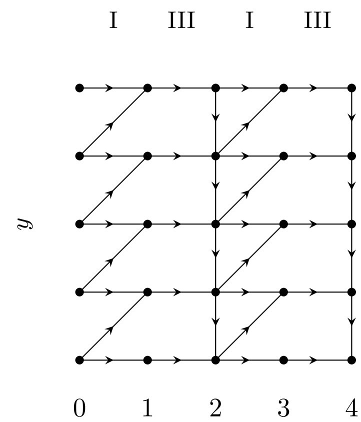

FIGURE 2. An example obtained by gluing columns in the alternating pattern I–III–I–III. Type I columns allow $(x, y) \to (x+1, y)$ and $(x, y) \to (x+1, y+1)$, while Type III columns allow $(x, y) \to (x+1, y)$ and vertical down moves $(x+1, y) \to (x+1, y-1)$ within the right slice.

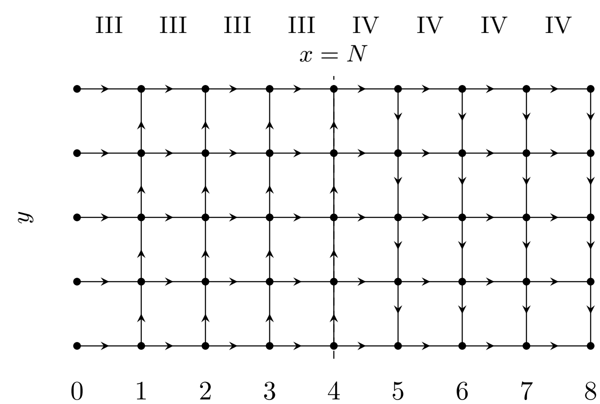

FIGURE 3. Example with $N = 4$: a block of four Type III columns followed by four Type IV columns.

while the horizontal edges $(x, y) \to (x + 1, y)$ have weight 1.

Type II. The diagonal down-edges

$$(x, y) \to (x + 1, y - 1)$$

are assigned weights

$$w\big((x, y) \to (x + 1, y - 1)\big) = \beta_{x,y} \in (0, \infty),$$

while the horizontal edges $(x, y) \to (x + 1, y)$ have weight 1.

Type III. The vertical up-edges at the right slice,

$$(x + 1, y) \to (x + 1, y + 1),$$

are assigned weights

$$w\big((x+1,y)\to(x+1,y+1)\big)=\gamma_{x+1,y}\in(0,\infty),$$

while the horizontal edges $(x,y)\to(x+1,y)$ have weight 1.

Type IV. The vertical down-edges at the right slice,

$$(x+1,y)\to(x+1,y-1),$$

are assigned weights

$$w\big((x+1,y)\to(x+1,y-1)\big)=\delta_{x+1,y}\in(0,\infty),$$

while the horizontal edges $(x,y)\to(x+1,y)$ have weight 1.

With these conventions, the weight of a path is the product of the weights of its edges, and the weight of a path ensemble is the product of the weights of the individual paths. Since all horizontal edges have unit weight, only diagonal and vertical moves contribute nontrivially to the total weight.

2.2. Non-interesting path ensembles and a point process. We impose boundary conditions

$$P_j=(0,-j),\qquad Q_j=(2N,-j),\qquad j=1,\dots,n,$$

and consider n-tuples of non-intersecting directed paths connecting P_j to Q_j.

The paths naturally induce a point process as follows. Let $\boldsymbol{\pi}=(\pi_1,\dots,\pi_n)$ be a non-intersecting path ensemble. At each vertical section $x\in\{0,1,\dots,2N\}$ we associate an ordered n-tuple of heights

$$\boldsymbol{\eta}^{(x)}=(\eta_1^{(x)}<\cdots<\eta_n^{(x)})\in\mathbb{Z}^n,$$

defined according to the type of column immediately to the left of x.

- If the column between $x-1$ and x is of **Type III**, then $\eta_j^{(x)}$ is the *highest* height reached by the jth path at slice x.
- If the column between $x-1$ and x is of **Type IV**, then $\eta_j^{(x)}$ is the *lowest* height reached by the jth path at slice x.
- If the column between $x-1$ and x is of **Type I or II**, then no vertical motion occurs at slice x, and $\eta_j^{(x)}$ is the unique height at which the jth path visits that slice.

By non-intersection, the ordering

$$\eta_1^{(x)}<\cdots<\eta_n^{(x)}$$

is preserved for all x. The boundary conditions read

$$\boldsymbol{\eta}^{(0)}=(-n,\dots,-1),\qquad \boldsymbol{\eta}^{(2N)}=(-n,\dots,-1).$$

The associated point process. The non-intersecting path ensemble induces a space–time point configuration

$$\eta:=\{(x,\eta_j^{(x)}):\ x=0,\dots,2N,\ j=1,\dots,n\}\subset\{0,\dots,2N\}\times\mathbb{Z}.$$

We refer to η as the point process associated with the non-intersecting paths. This is the main observable of interest.

Transition matrices. For each $x \in \{0, 1, \dots, 2N-1\}$ we define a (generally infinite) transition matrix

$$T_{x,x+1} = \big(T_{x,x+1}(y, y')\big)_{y,y' \in \mathbb{Z}},$$

where $T_{x,x+1}(y, y')$ is the total weight of all single-path segments whose *recorded height* changes from y at slice x to y' at slice $x+1$ while crossing the xth column.

For $0 \le x < x' \le 2N$ we define the multi-step transition matrix by

$$T_{x,x'} := T_{x,x+1} T_{x+1,x+2} \cdots T_{x'-1,x'}.$$

Determinantal distribution of the point process.

THEOREM 60 (Product form of the path measure). *For any choice of strictly increasing n-tuples*

$$\boldsymbol{\eta}^{(1)}, \dots, \boldsymbol{\eta}^{(2N-1)} \in \mathbb{Z}^n,$$

the probability that the associated point process η satisfies $\boldsymbol{\eta}^{(x)} = (\eta_1^{(x)}, \dots, \eta_n^{(x)})$ for all $x = 0, \dots, 2N$ is given by

$$\mathbb{P}\Big(\boldsymbol{\eta}^{(1)}, \dots, \boldsymbol{\eta}^{(2N-1)}\Big) = \frac{1}{Z} \prod_{x=0}^{2N-1} \det\big[T_{x,x+1}\big(\eta_i^{(x)}, \eta_j^{(x+1)}\big)\big]_{i,j=1}^{n},$$

where the partition function is

$$Z = \det\big[\, T_{0,2N}(-i, -j)\,\big]_{i,j=1}^{n}.$$

COROLLARY 61 (Single-slice marginal of the point process). *Fix $m \in \{0, 1, \dots, 2N\}$. For any strictly increasing n-tuple $\mathbf{y} = (y_1 < \cdots < y_n)$,*

$$\mathbb{P}\big(\boldsymbol{\eta}^{(m)} = \mathbf{y}\big) = \frac{1}{Z} \det\big[\, T_{0,m}(-i, y_j)\,\big]_{i,j=1}^{n} \det\big[\, T_{m,2N}(y_i, -j)\,\big]_{i,j=1}^{n}.$$

2.3. Correlation functions. We now turn to correlation functions of the point process

$$\eta = \{(x, \eta_j^{(x)}) : \ x = 0, \dots, 2N, \ j = 1, \dots, n\} \subset \{0, \dots, 2N\} \times \mathbb{Z}.$$

For a fixed slice $m \in \{0, \dots, 2N\}$ we write

$$\eta^{(m)} := \{\eta_1^{(m)}, \dots, \eta_n^{(m)}\} \subset \mathbb{Z},$$

so that $\eta = \bigcup_{m=0}^{2N} \{m\} \times \eta^{(m)}$.

Throughout we use the transition matrices $T_{x,x'}$ defined earlier, and we set

$$G := \big(G_{i,j}\big)_{i,j=1}^{n}, \qquad G_{i,j} := T_{0,2N}(-i, -j).$$

By the LGV theorem and our boundary conditions, the partition function equals

$$Z = \det G.$$

In particular, we assume $\det G \neq 0$, so that G^{-1} exists.

Correlation functions. For $k \geq 1$, the k-point correlation function of the slice process $\eta^{(m)}$ is defined by

$$\rho_k^{(m)}(y_1, \dots, y_k) := \mathbb{P}\big(\{y_1, \dots, y_k\} \subset \eta^{(m)}\big), \qquad y_1, \dots, y_k \in \mathbb{Z} \text{ distinct.}$$

Similarly, for the full space–time process η we define

$$\rho_k\big((x_1, y_1), \dots, (x_k, y_k)\big) := \mathbb{P}\big(\{(x_1, y_1), \dots, (x_k, y_k)\} \subset \eta\big),$$

for distinct space–time points $(x_a, y_a) \in \{0, \dots, 2N\} \times \mathbb{Z}$.

THEOREM 62 (Determinantal correlations on a fixed slice). *Fix $m \in \{0, 1, \dots, 2N\}$ and define a kernel on $\mathbb{Z} \times \mathbb{Z}$ by*

$$K_n^{(m)}(y, y') := \sum_{i,j=1}^{n} T_{m,2N}(y, -j)\,(G^{-1})_{j,i}\,T_{0,m}(-i, y'). \tag{14}$$

Then the point process $\eta^{(m)}$ on $\mathbb{Z}$ is determinantal with kernel $K^{(m)}$, i.e. for any $k \geq 1$ and any distinct $y_1, \dots, y_k \in \mathbb{Z}$,

$$\rho_k^{(m)}(y_1, \dots, y_k) = \det\big[\, K_N^{(m)}(y_a, y_b)\,\big]_{a,b=1}^{k}.$$

PROOF. The weight-sum formula for the single-slice marginal (proved earlier) reads

$$\mathbb{P}\big(\eta^{(m)} = \{y_1 < \cdots < y_n\}\big) = \frac{1}{\det G} \det\big[\, T_{0,m}(-i, y_j)\,\big]_{i,j=1}^{n} \det\big[\, T_{m,2N}(y_i, -j)\,\big]_{i,j=1}^{n}.$$

This is an instance of the Eynard–Mehta (or Cauchy–Binet) structure: a probability measure on n-point subsets of $\mathbb{Z}$ proportional to a product of two determinants. A standard determinantal calculation (expanding minors and using Cauchy–Binet with the normalization $\det A$) shows that all inclusion probabilities are given by determinants with kernel (14). ☐

The previous theorem describes the correlations within a single slice. We next state the determinantal structure of the full space–time process η.

THEOREM 63 (Extended determinantal point process). *Define a kernel on $\big(\{0, \dots, 2N\} \times \mathbb{Z}\big)^2$ by*

$$K_n\big((x, y), (x', y')\big) := -\mathbf{1}_{\{x<x'\}}\,T_{x,x'}(y, y') + \sum_{i,j=1}^{n} T_{x,2N}(y, -j)\,(G^{-1})_{j,i}\,T_{0,x'}(-i, y'). \tag{15}$$

Then η is a determinantal point process with correlation kernel K, i.e. for any $k \geq 1$ and any distinct space–time points $(x_1, y_1), \dots, (x_k, y_k)$,

$$\rho_k\big((x_1, y_1), \dots, (x_k, y_k)\big) = \det\big[\, K_n\big((x_a, y_a), (x_b, y_b)\big)\,\big]_{a,b=1}^{k}.$$

Moreover, the single-slice kernel of Theorem 62 is obtained by restriction: for $x = x' = m$ one has

$$K_n\big((m, y), (m, y')\big) = K_n^{(m)}(y, y').$$

PROOF. This is the standard Eynard–Mehta theorem for non-intersecting path ensembles with fixed boundary data, applied to the product-of-determinants representation of the path measure in terms of the one-step matrices $T_{x,x+1}$. The indicator term $-\mathbf{1}_{\{x<x'\}}T_{x,x'}$ encodes the directed time-ordering of the paths, while the second term imposes the boundary conditions through the matrix G^{-1}. The

restriction to a single slice follows immediately from (15) because the indicator term vanishes when $x = x'$. □

3. Column-homogeneous parameters and Laurent transition matrices

We now specialize to the case in which all parameters are constant within each column. More precisely, for each column index r we assume that the non-horizontal edge weights do not depend on the vertical coordinate y, but only on r and on the column type. Thus we write

$$\alpha_{r,y} \equiv \alpha_r, \qquad \beta_{r,y} \equiv \beta_r, \qquad \gamma_{r,y} \equiv \gamma_r, \qquad \delta_{r,y} \equiv \delta_r,$$

whenever the corresponding edge type is present in column r. Horizontal edges continue to have weight 1.

Under this assumption, all transition matrices $T_{r,r+1}$ are translation invariant in the vertical direction. Consequently, each $T_{r,r+1}$ is a Laurent matrix, i.e. a bi-infinite matrix whose entries depend only on the difference $y' - y$.

One-step transition matrices. Writing $k = y' - y$, the matrix elements take the following form.

- **Type I:**
$$T_{r,r+1}(y, y') = \delta_{k,0} + \alpha_r\, \delta_{k,1}.$$
- **Type II:**
$$T_{r,r+1}(y, y') = \delta_{k,0} + \beta_r\, \delta_{k,-1}.$$
- **Type III:**
$$T_{r,r+1}(y, y') = \begin{cases} \gamma_r^{y'-y}, & k \geq 0, \\ 0, & k < 0. \end{cases}$$
- **Type IV:**
$$T_{r,r+1}(y, y') = \begin{cases} \delta_r^{y-y'}, & k \leq 0, \\ 0, & k > 0. \end{cases}$$

In each case $T_{r,r+1}(y, y')$ depends only on $k = y' - y$, confirming that $T_{r,r+1}$ is a Laurent matrix.

Symbol representation. We then introduce the symbols:

$$a_r(z) := \sum_{k\in\mathbb{Z}} T_{r,r+1}(0, k)\, z^k, \qquad z \in \mathbb{C}^\times.$$

For the four column types this gives

- **Type I:**
$$a_r(z) = 1 + \alpha_r z.$$
- **Type II:**
$$a_r(z) = 1 + \beta_r/z.$$
- **Type III:**
$$a_r(z) = \sum_{k\geq 0} \delta_r^k z^k = \frac{1}{1 - \gamma_r z},.$$
- **Type IV:**
$$a_r(z) = \sum_{k\geq 0} \gamma_r^k z^{-k} = \frac{1}{1 - \delta_r/z}.$$

Multi-step transitions. Since products of Laurent matrices are again Laurent matrices, all multi-step transition matrices

$$T_{x,x'} = T_{x,x+1} T_{x+1,x+2} \cdots T_{x'-1,x'}$$

are Laurent matrices as well. Their symbols are given by

$$a_{x,x'}(z) = \prod_{r=x}^{x'-1} a_r(z),$$

and thus

$$T_{x,x'} = L(\tilde{a}_{x,x'}),$$

where $L(a)$ stands for the Laurent matrices with symbol a, and we also used the notation $\tilde{a}(z) = a(1/z)$-

Note also, for $i,j = 1,\dots,n$ we have

$$G_{i,j} = (L(\tilde{a}_{x,x'}))_{-i,-j} = (L(a_{x,x'}))_{i,j}$$

and thus

$$G_{i,j} = T_n(a).$$

This Laurent-symbol representation will be used extensively in the analysis of correlation kernels and scaling limits.

3.1. Taking the limit $n \to \infty$. For finite n the model is still quite involved, since the correlation kernel contains the inverse of a finite Toeplitz matrix. In general, inverting such matrices is a difficult task. A major simplification occurs, however, when we take the limit $n \to \infty$. In this regime, the finite Toeplitz matrix is replaced by a Toeplitz operator, whose inversion is, remarkably, much more tractable.

What follows requires justification—which will be provided later—but for the moment we proceed formally. We take the limit $n \to \infty$ and simply replace $T_n(a)^{-1}$ by $T(a)^{-1}$. This leads to the following result.

THEOREM 64. *Suppose that $a_{0,N}$ admits a Wiener–Hopf factorization*

$$a_{0,N} = a_+ a_-.$$

Then the limit $n \to \infty$ of the point process exists, is determinantal, and has correlation kernel

$$K\big((x,y),(x',y')\big) = -\mathbf{1}_{\{x<x'\}} \frac{1}{2\pi i} \oint_{|z|=1-} a_{x,x'}(z) \frac{dz}{z^{y'-y+1}} + \frac{1}{(2\pi i)^2} \oint_{|z|=1-} dz \oint_{|w|=1+} dw\, a_{x,2N}(z)\, a_+^{-1}(z)\, a_-^{-1}(w)\, a_{0,x'}(w) \frac{1}{z(w-z)} \frac{z^y}{w^{y'}}. \tag{16}$$

PROOF. By definition,

$$K\big((x,y),(x',y')\big) = -\mathbf{1}_{\{x<x'\}} T_{x,x'}(y,y') + \sum_{i,j=0}^{\infty} T_{x,2N}(y,-j)\,(T(a)^{-1})_{j,i}\, T_{0,x'}(-i,y'). \tag{17}$$

Using the Wiener–Hopf inversion formula

$$T(a)^{-1} = T(a_+^{-1})\, T(a_-^{-1}),$$

together with

$$T_{x,2N}(y,-j) = \big(T(a_{x,2N})\big)_{-y,j}, \qquad T_{0,x'}(-i,y') = \big(T(a_{0,x'})\big)_{i,-y'},$$

we obtain

(18)
$$K\big((x,y),(x',y')\big) = -\mathbf{1}_{\{x<x'\}}\, T_{x,x'}(y,y') + \big(T(a_{x,2N})\, T(a_+^{-1})\, T(a_-^{-1})\, T(a_{0,x'})\big)_{-y,-y'}.$$

Since a_+^{-1} (resp. a_-^{-1}) has only nonnegative (resp. nonpositive) Fourier modes, this simplifies to

$$K\big((x,y),(x',y')\big) = -\mathbf{1}_{\{x<x'\}}\, T_{x,x'}(y,y') + \big(T(a_{x,2N}a_+^{-1})\, T(a_-^{-1}a_{0,x'})\big)_{-y,-y'}. \tag{19}$$

Writing out the Toeplitz operators gives

$$\sum_{\ell=0}^{\infty} \frac{1}{(2\pi i)^2} \oint dz \oint dw\, a_{x,2N}(z)\, a_+^{-1}(z)\, a_-^{-1}(w)\, a_{0,x'}(w) \frac{1}{z^{-y-\ell+1}} \frac{1}{w^{\ell+y'+1}}.$$

For $|w| > |z|$,

$$\sum_{\ell=0}^{\infty} \frac{z^\ell}{w^{\ell+1}} = \frac{1}{w-z},$$

which yields

$$\begin{aligned} K\big((x,y),(x',y')\big) &= -\mathbf{1}_{\{x<x'\}}\, T_{x,x'}(y,y') \\ &+ \frac{1}{(2\pi i)^2} \oint_{|z|=1-} dz \oint_{|w|=1+} dw\, a_{x,2N}(z)\, a_+^{-1}(z)\, a_-^{-1}(w)\, a_{0,x'}(w) \frac{1}{z(w-z)} \frac{z^y}{w^{y'}}. \end{aligned} \tag{20}$$

Finally,

$$T_{x,x'}(y,y') = \frac{1}{2\pi i} \oint_{|z|=1-} a_{x,x'}(z)\, \frac{dz}{z^{y'-y+1}},$$

which completes the formal computation. □

3.2. Example: Aztec diamond. Consider now the case

$$a_{2x} = 1 + az, \qquad x = 0, \ldots, N-1,$$

and

$$a_{2x+1} = \frac{1}{1 - a/z}, \qquad x = 0, \ldots, N-1,$$

where $a \in (0,1)$. This corresponds to alternating columns of type I and type IV.

Configurations of the associated non-intersecting paths encode domino tilings of the Aztec diamond. Thus, we obtain a probability measure on the set of all such tilings. Vertical dominoes have weight a, while horizontal dominoes have weight 1; this model is known as the biased Aztec diamond.

The resulting point process is determinantal, with correlation kernel

(21)
$$\begin{aligned} K\big((2p,y),(2q,y')\big) = -\mathbf{1}_{\{p<q\}} \frac{1}{2\pi i} \oint_{|z|=1-} (1+az)^{\,q-p} \left(1 - \frac{a}{z}\right)^{-(q-p)} \frac{dz}{z^{\,y'-y+1}} \\ + \frac{1}{(2\pi i)^2} \oint_{|z|=1-} dz \oint_{|w|=1+} dw\; (1+az)^{-p} \left(1 - \frac{a}{z}\right)^{-(N-p)} \\ \times (1+aw)^q \left(1 - \frac{a}{w}\right)^{N-q} \frac{1}{z(w-z)} \frac{z^y}{w^{y'}}. \end{aligned}$$

By taking the limit $a \to 1$, we obtain the correlation kernel for the determinantal point process corresponding to the uniform measure on all possible tilings.

3.3. Edrei–Thoma theory and natural Toeplitz building blocks. In the column-homogeneous setting, each one-step transition matrix $T_{x,x+1}$ is a Toeplitz (Laurent) matrix, so single-path weights are translation invariant in the vertical direction. In order for the Lindström–Gessel–Viennot determinants to define nonnegative weights for non-intersecting path ensembles, it is natural to impose a stronger structural condition on these transitions, namely *total nonnegativity*: all minors of the Toeplitz kernel $T_{x,x+1}(y, y')$ are nonnegative. In the Toeplitz case this property is equivalent to the classical Pólya frequency condition PF_∞.

A fundamental classification theorem due to Edrei and Thoma describes all such totally nonnegative Toeplitz kernels in terms of their symbols. In the formulation relevant for our purposes, if

$$a(z) = \sum_{k\in\mathbb{Z}} a_k z^k$$

is the symbol of a Toeplitz kernel with (a_k) totally nonnegative, then a admits a canonical factorization of the form

$$a(z) = C\, z^\kappa\, \exp\bigl(t(z + z^{-1})\bigr) \prod_{m\geq 1} \frac{1 + \alpha_m z}{1 - \gamma_m z} \prod_{m\geq 1} \frac{1 + \beta_m z^{-1}}{1 - \delta_m z^{-1}},$$

where $C \geq 0$, $\kappa \in \mathbb{Z}$, $t \geq 0$, and the parameters $\alpha_m, \beta_m, \gamma_m, \delta_m \geq 0$ satisfy appropriate summability conditions ensuring convergence on an annulus containing the unit circle.

Each factor in this representation has a direct interpretation in terms of elementary graph structures. The linear factors

$$1 + \alpha z \quad \text{and} \quad 1 + \beta z^{-1}$$

correspond to Bernoulli up- and down-steps (Types I and II), while the geometric factors

$$(1 - \delta z)^{-1} \quad \text{and} \quad (1 - \gamma z^{-1})^{-1}$$

correspond to stacks of vertical moves at a fixed slice (Types IV and III). In this way, the four elementary column types introduced earlier arise as the atomic factors in the Edrei–Thoma classification.

From this perspective, the column-gluing construction is not ad hoc: in the translation-invariant setting, any homogeneous one-step transition compatible with total nonnegativity — and hence with a positive non-intersecting path measure — can be assembled, up to harmless shifts and exponential factors, from these elementary building blocks. The exponential term $\exp\bigl(t(z + z^{-1})\bigr)$ may be interpreted as a continuous-time nearest-neighbor walk and can be approximated by suitable limits of alternating Bernoulli columns.

Consequently, measures on non-intersecting paths constructed from products of Toeplitz determinants and satisfying the natural positivity and consistency requirements are expected to arise from graphs built out of these elementary column types. The Edrei–Thoma theorem thus provides a structural justification for the graphical framework used throughout these notes.

APPENDIX A

Determinants and trace of kernels

1. Introduction

Let μ be a finite Borel measure on a metric space S and let $K : S \times S \to \mathbb{C}$ be a $\mu \times \mu$–measurable function satisfying

$$\iint_{S\times S} |K(x,y)|^2 \, d\mu(x)\, d\mu(y) < \infty. \tag{22}$$

Note that such measurable functions form a vector space, and an algebra when equipped with the product

$$(K_1 \cdot K_2)(x,y) = \int_S K_1(x,z)\, K_2(z,y)\, d\mu(z).$$

We will denote this algebra by $\mathcal{B}$.

Special examples of interest are the kernels that are finite sums of separable kernels:

$$\mathcal{F} = \left\{ \sum_{j=1}^{n} f_j(x)\, g_j(y) \;\middle|\; n \in \mathbb{N},\ f_j, g_j \in L_2(S,\mu),\ j = 1, \dots, n \right\}.$$

At various places (but certainly not always) we will identify K with the integral operator $\mathcal{K}$ on $L_2(S,\mu)$ given by

$$\mathcal{K}f(x) = \int_S K(x,y)\, f(y)\, d\mu(y),$$

for $x \in S$. The condition (22) ensures that $\mathcal{K}$ is a bounded operator. In fact, it makes the operator Hilbert–Schmidt, as we will see later on. The elements of $\mathcal{F}$ correspond to the finite rank operators on $L_2(S,\mu)$.

For integral operators with continuous kernel one can naturally extend the notions of trace and determinant. The idea is the following:

(1) Define $\mathrm{Tr}\mathcal{K}$ and $\det(I+\mathcal{K})$ for $K \in \mathcal{F}$ by viewing them as finite rank operators.
(2) Introduce a norm on $\mathcal{F}$ such that $K \mapsto \mathrm{Tr}\mathcal{K}$ and $K \mapsto \det(I+\mathcal{K})$ are locally Lipschitz continuous.
(3) Extend the definition of determinant and trace by continuity.

There are various different norms that one can define such that the maps $K \mapsto \mathrm{Tr}\mathcal{K}$ and $K \mapsto \det(I+\mathcal{K})$ are locally Lipschitz continuous. Interestingly, these definitions (and their values!) may depend on the chosen norm. We will discuss three norms:

- uniform norm for continuous kernels,
- trace norm for trace class operators,
- a new norm that is specially adapted for determinantal point processes.

The first two are classical ways of defining trace and determinants. The third is new to the best of our knowledge. An important feature of this third norm is that it depends on the values on the diagonal $K(x,x)$. As a consequence we should no longer identify the kernels with their integral operators (indeed, for continuous measures, the integral operator does not depend on the diagonal $K(x,x)$, as it is a set of $\mu\times\mu$–measure zero).

Finally, we remark that there are subtle differences between these theories. We will illustrate this in a separate section.

2. Preliminaries

We first recall three well-known results that will be important to us.

LEMMA 65 (Andreiéf's identity). *Let (X,ν) be a measure space. Then*

$$\int_X\cdots\int_X \det\big(f_j(x_k)\big)_{j,k=1}^n \det\big(g_j(x_k)\big)_{j,k=1}^n\,d\nu(x_1)\cdots d\nu(x_n) = n!\det\left(\int_X f_j(x)\,g_k(x)\,d\nu(x)\right)_{j,k=1}^n,$$

for any $n\in\mathbb{N}$ and square integrable functions $f_1,\dots,f_n,g_1,\dots,g_n$.

PROOF. The proof follows from a straightforward computation:

$$\begin{aligned}
&\int_X\cdots\int_X \det\big(f_j(x_k)\big)_{j,k=1}^n \det\big(g_j(x_k)\big)_{j,k=1}^n\,d\nu(x_1)\cdots d\nu(x_n)\\
&=\int_X\cdots\int_X \sum_{\sigma\in S_n}\sum_{\tau\in S_n}\operatorname{sign}(\sigma)\operatorname{sign}(\tau)\prod_{j=1}^n f_{\sigma(j)}(x_j)\prod_{j=1}^n g_{\tau(j)}(x_j)\,d\nu(x_1)\cdots d\nu(x_n)\\
&\qquad=\sum_{\sigma\in S_n}\sum_{\tau\in S_n}\operatorname{sign}(\sigma)\operatorname{sign}(\tau)\prod_{j=1}^n\left(\int_X f_{\sigma(j)}(x)\,g_{\tau(j)}(x)\,d\nu(x)\right)\\
&\qquad\quad=\sum_{\sigma\in S_n}\sum_{\tau\in S_n}\operatorname{sign}(\sigma^{-1})\operatorname{sign}(\tau)\prod_{j=1}^n\left(\int_X f_j(x)\,g_{\tau\circ\sigma^{-1}(j)}(x)\,d\nu(x)\right),
\end{aligned}$$

where the last equality holds by reordering the product according to the permutation σ^{-1}.

Since $\operatorname{sign}(\sigma^{-1})\operatorname{sign}(\tau)=\operatorname{sign}(\tau\circ\sigma^{-1})$, the double sum simplifies to a sum over $\sigma'=\tau\circ\sigma^{-1}\in S_n$ and we thus arrive at

$$\int_X\cdots\int_X \det\big(f_j(x_k)\big)_{j,k=1}^n \det\big(g_j(x_k)\big)_{j,k=1}^n\,d\nu(x_1)\cdots d\nu(x_n) = n!\sum_{\sigma'\in S_n}\operatorname{sign}(\sigma')\prod_{j=1}^n\left(\int_X f_j(x)\,g_{\sigma'(j)}(x)\,d\nu(x)\right),$$

and thus the statement follows. □

If ν is the counting measure on $\{1,\dots,m\}$, then by setting $C_{jk}=f_j(x_k)$ and $D_{jk}=g_j(x_k)$, this is the Cauchy–Binet identity for the determinant of the product of two rectangular matrices. If in addition $n=m$ then it follows (after a symmetrization) that $\det(CD)=\det C\det D$ for square matrices C and D.

LEMMA 66 (Sylvester's identity). *For any $A \in \mathbb{C}^{k\times m}$ and $B \in \mathbb{C}^{m\times k}$ we have*

$$\det(I_k + AB) = \det(I_m + BA).$$

PROOF. The statement follows from the following trick:

$$\begin{aligned}
\det(I_k + AB) &= \det\begin{pmatrix} I_k + AB & A \\ 0 & I_m \end{pmatrix} \\
&= \det\begin{pmatrix} I_k & A \\ -B & I_m \end{pmatrix} \det\begin{pmatrix} I_k & 0 \\ B & I_m \end{pmatrix} \\
&= \det\begin{pmatrix} I_k & 0 \\ B & I_m \end{pmatrix} \det\begin{pmatrix} I_k & A \\ -B & I_m \end{pmatrix} \\
&= \det\left(\begin{pmatrix} I_k & 0 \\ B & I_m \end{pmatrix} \begin{pmatrix} I_k & A \\ -B & I_m \end{pmatrix}\right) \\
&= \det\begin{pmatrix} I_k & A \\ 0 & I_m + BA \end{pmatrix} = \det(I_m + BA).
\end{aligned}$$

Here we used that $\det(CD) = \det C \det D$ for any square matrices C and D. □

LEMMA 67 (von Koch form of the determinant). *Let A be an $n \times n$ matrix. Then*

$$\det(I + A) = 1 + \sum_{k=1}^{n} \frac{1}{k!} \sum_{i_1,\dots,i_k \in \{1,\dots,n\}} \det\big(A_{i_j,i_\ell}\big)_{j,\ell=1}^{k}.$$

PROOF. A proof can be found in standard texts on determinants; we omit the details. □

3. Trace and determinant for finite rank operators

For finite rank operators the obvious way of defining the trace and determinant is via their matrix representation. Let $\mathcal{K}$ be a finite rank operator and let M be a subspace of $L_2(S,\mu)$ such that both the range and the cokernel of $\mathcal{K}$ are subspaces of M. In other words, $\mathcal{K}$ acts trivially on $M^\perp$. Then we take an orthonormal basis $\{e_1,\dots,e_N\}$ for M (where $N = \dim M$) and define

$$\mathrm{Tr}\mathcal{K} = \sum_{i=1}^{N} (\mathcal{K}e_i, e_i)$$

and

$$\det(I + \mathcal{K}) = \det\big(\delta_{jk} + (\mathcal{K}e_k, e_j)\big)_{j,k=1}^{N}.$$

Here $(f,g) = \int_S f(x)\,\overline{g(x)}\,d\mu(x)$ denotes the $L_2(S,\mu)$ inner product between f and g.

The definitions of $\mathrm{Tr}\mathcal{K}$ and $\det(I+\mathcal{K})$ do not depend on the chosen basis, which can be seen from the following lemmas.

LEMMA 68. *Let $K(x,y) = \sum_{j=1}^{r} f_j(x)\, g_j(y)$ be the kernel for the finite rank operator $\mathcal{K}$. Then the following hold:*

(1) $\mathrm{Tr}\mathcal{K} = \sum_{j=1}^{r} \lambda_j$, *where λ_j are the eigenvalues of $\mathcal{K}$ (counted with algebraic multiplicity);*

(2) $\mathrm{Tr}\mathcal{K} = \int_S K(x,x)\,d\mu(x) = \sum_{j=1}^{r}(f_j, g_j)$.

PROOF. The first property is standard. The second property is a straightforward consequence of Plancherel's theorem (or simply of the definition of the trace in terms of an orthonormal basis). □

LEMMA 69. *Let $K(x,y) = \sum_{j=1}^{r} f_j(x)\,g_j(y)$ be the kernel for the finite rank operator $\mathcal{K}$. Then the following hold:*

(1) $\det(I+\mathcal{K}) = \prod_{j=1}^{r}(1+\lambda_j)$;
(2) $\det(I+\mathcal{K}) = \det\big(\delta_{jk} + (f_j, g_k)\big)_{j,k=1}^{r}$.

PROOF. The first property is again standard.
For the second property, note that

$$(\mathcal{K}e_i, e_j) = \sum_{\ell=1}^{r}(g_\ell, e_i)\,(f_\ell, e_j).$$

Define the matrices

$$\begin{aligned} A_{i\ell} &= \langle g_\ell, e_i\rangle, \qquad i=1,\dots,N,\ \ell=1,\dots,r,\\ B_{\ell j} &= \langle f_\ell, e_j\rangle, \qquad \ell=1,\dots,r,\ j=1,\dots,N. \end{aligned}$$

Then $(\mathcal{K}e_i, e_j) = (AB)_{ij}$ and thus the determinant can be written as $\det(I_N + AB)$. By Sylvester's identity (Lemma 66) we get

$$\det(I_N+\mathcal{K}) = \det(I_N + AB) = \det(I_r + BA).$$

A direct computation shows that $(BA)_{jk} = \int_S f_j(x)\,g_k(x)\,d\mu(x)$, proving the second property. □

For the next important form of the determinant we use Andreiéf's identity.

LEMMA 70. *For a finite rank operator $\mathcal{K}$ with kernel K we have*

$$\det(I+\mathcal{K}) = \sum_{m=0}^{\infty}\frac{1}{m!}\int_S\cdots\int_S \det\big(K(x_i,x_j)\big)_{i,j=1}^{m}\,d\mu(x_1)\cdots d\mu(x_m), \tag{23}$$

and the series truncates at $m = r$, where r is the rank of $\mathcal{K}$.

PROOF. Write $K(x,y) = \sum_{\ell=1}^{r} f_\ell(x)\,g_\ell(y)$. By applying Andreiéf's identity twice we obtain

$$\begin{aligned} &\int_S\cdots\int_S \det\big(K(x_i,x_j)\big)_{i,j=1}^{m}\,d\mu(x_1)\cdots d\mu(x_m)\\ &= \frac{1}{m!}\sum_{k_1=1}^{r}\cdots\sum_{k_m=1}^{r}\int_S\cdots\int_S \det\big(f_{k_j}(x_i)\big)_{i,j=1}^{m}\det\big(g_{k_j}(x_i)\big)_{i,j=1}^{m}\,d\mu(x_1)\cdots d\mu(x_m)\\ &\qquad\qquad = \sum_{k_1=1}^{r}\cdots\sum_{k_m=1}^{r}\det\big((f_{k_j}, g_{k_i})\big)_{i,j=1}^{m}. \end{aligned}$$

Thus

$$\sum_{m=0}^{r} \frac{1}{m!} \int_S \cdots \int_S \det\big(K(x_i, x_j)\big)_{i,j=1}^m \, d\mu(x_1) \cdots d\mu(x_m)$$
$$= \sum_{m=0}^{r} \frac{1}{m!} \sum_{k_1=1}^{r} \cdots \sum_{k_m=1}^{r} \det\big((f_{k_j}, g_{k_i})\big)_{i,j=1}^m .$$

Hence the statement follows by applying Lemma 67 and part (2) of Lemma 69. □

The trace and determinant have many important properties. We will especially use the ones in the following lemma.

LEMMA 71. *The determinant has the following properties for finite rank operators:*

(1) $\det(I + K_1)\det(I + K_2) = \det(I + K_1 + K_2 + K_1K_2)$.
(2) $\log\det(I+zK) = \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j} z^j \,\mathrm{Tr} K^j$, *valid for* $|z|$ *sufficiently small (for example,* $|z| \le 1/\max_j |\lambda_j(K)|$*).*

4. Continuous extensions

Let $\mathcal{B}_*$ be a subalgebra of $\mathcal{B}$ that is equipped with a norm $\|\cdot\|_*$ such that:

(1) $\mathcal{B}_*$ is a Banach algebra, i.e., $\mathcal{B}_*$ is complete and
$$\|K_1K_2\|_* \le \|K_1\|_* \, \|K_2\|_*.$$
(2) The finite rank operators (i.e. kernels in $\mathcal{F}$) form a dense subspace of $\mathcal{B}_*$; we denote this dense subspace by $\mathcal{B}_F$.
(3) The map $K \mapsto \mathrm{Tr}\mathcal{K}$ is a continuous linear functional on $\mathcal{B}_F$ with respect to $\|\cdot\|_*$.

Strictly speaking, in our examples $\|\cdot\|_*$ will be a seminorm that will be turned into a norm by considering equivalence classes in a standard fashion: we say $K_1 \sim K_2$ if and only if $\|K_1 - K_2\|_* = 0$. The continuity of the trace then shows that $\|K_1 - K_2\|_* = 0$ implies $\mathrm{Tr}\mathcal{K}_1 = \mathrm{Tr}\mathcal{K}_2$, making the trace a well-defined function on the equivalence classes. This is not a trivial point, since our equivalence classes will be different in the upcoming examples.

The fact that $\mathcal{B}_*$ is a Banach algebra is a useful property that we will exploit at several places.

LEMMA 72. *Let* $K \in \mathcal{B}_*$. *Then*

(1) $K^j \in \mathcal{B}_*$ *and* $\|K^j\|_* \le \|K\|_*^j$ *for* $j = 1, 2, \ldots$;
(2) $\exp(K) = \sum_{j=0}^{\infty} \frac{K^j}{j!} \in \mathcal{B}_*$ *and* $\|\exp K\|_* \le \exp(\|K\|_*)$.

If $\|K\|_* < 1$, *then*

(3) $(I - K)^{-1} = \sum_{j=0}^{\infty} K^j \in \mathcal{B}_*$ *and* $\|(I - K)^{-1}\|_* \le (1 - \|K\|_*)^{-1}$;
(4) $\log(I - K) = -\sum_{j=1}^{\infty} \frac{K^j}{j} \in \mathcal{B}_*$ *and* $\|\log(I - K)\|_* \le -\log(1 - \|K\|_*)$.

PROOF. The proof is standard: it relies on the completeness of $\mathcal{B}_*$ and the submultiplicative property of $\|\cdot\|_*$. □

All the arguments in the following paragraph are copied from [**1**].

Theorem 73. *The map $K \mapsto \det(I+\mathcal{K})$ on $\mathcal{B}_F$ is locally Lipschitz continuous with respect to $\|\cdot\|_*$ and thus has a continuous extension to the whole of $\mathcal{B}_*$.*

Proof. We first prove that $K \mapsto \det(I+\mathcal{K})$ is continuous on the ball in $\mathcal{B}_F$ of radius $0<r<1$. Indeed, for any $F, G \in \mathcal{B}_F$ such that $\|F\|_*, \|G\|_* < r$ we have

$$\begin{aligned}|\det(I+F)-\det(I+G)| &= \left|\exp\Big(\sum_{j=1}^{\infty}\frac{(-1)^{j+1}}{j}\mathrm{Tr}F^j\Big)-\exp\Big(\sum_{j=1}^{\infty}\frac{(-1)^{j+1}}{j}\mathrm{Tr}G^j\Big)\right| \\ &\le c\left|\sum_{j=1}^{\infty}\frac{(-1)^{j+1}}{j}\mathrm{Tr}F^j-\sum_{j=1}^{\infty}\frac{(-1)^{j+1}}{j}\mathrm{Tr}G^j\right| \\ &\le c_1\sum_{j=1}^{\infty}\frac{1}{j}\left|\mathrm{Tr}F^j-\mathrm{Tr}G^j\right| \le c_1c_2\sum_{j=1}^{\infty}\frac{1}{j}\left\|F^j-G^j\right\|_*,\end{aligned}$$

for some constants $c, c_1, c_2 > 0$. Using that $\mathcal{B}_*$ is a Banach algebra we find

$$\left\|F^j-G^j\right\|_* = \left\|\sum_{\ell=0}^{j-1}F^\ell(F-G)G^{j-\ell-1}\right\|_* \le \sum_{\ell=0}^{j-1}\|F\|_*^\ell\,\|F-G\|_*\,\|G\|_*^{j-\ell-1} \le jr^{j-1}\|F-G\|_*,$$

and thus

$$|\det(I+F)-\det(I+G)| \le \frac{c_1c_2}{1-r}\,\|F-G\|_*,$$

which proves that $F\mapsto\det(I+F)$ is continuous (in fact Lipschitz continuous) in the ball with radius $r<1$.

Next we prove that $K\mapsto\det(I+\mathcal{K})$ is continuous on all of $\mathcal{B}_F$. To this end, fix $F\in\mathcal{B}_F$ and note that

$$\det(I+F)-\det(I+G) = \det(I+G)\Big(\det\big(I+(I+G)^{-1}(F-G)\big)-1\Big),$$

for any $F,G\in\mathcal{B}_F$. Then by the argument above there exist $\rho>0$ and $c_3>0$ such that

$$|\det(I+F)-\det(I+G)| \le c_3|\det(I+G)|\,\|(I+G)^{-1}\|_*\|F-G\|_*$$

whenever $\|F-G\|_*<\rho$. This proves the local Lipschitz continuity. □

We are now ready to define the trace and determinant by continuous extension.

Definition 74. *We denote the continuous extensions of $K\mapsto\mathrm{Tr}\mathcal{K}$ and $K\mapsto\det(I+\mathcal{K})$ from $\mathcal{B}_F$ to $\mathcal{B}_*$ by Tr_* and $\det_*$. That is, for $K\in\mathcal{B}_*$ we set*

$$\mathrm{Tr}_*K=\lim_{n\to\infty}\mathrm{Tr}F_n,\qquad \det_*(I+K)=\lim_{n\to\infty}\det(I+F_n),$$

where $\{F_n\}_n\subset\mathcal{B}_F$ is any sequence converging to K in $\mathcal{B}_*$.

Many of the properties of the trace and determinant for finite matrices transfer directly (by continuity) to elements of $\mathcal{B}_*$. We mention two important identities.

Lemma 75. *For $K_1, K_2, K\in\mathcal{B}_*$ we have*

(1) $\det_*(I+K_1)\det_*(I+K_2)=\det_*(I+K_1+K_2+K_1K_2)$;

(2) $\log\det_*(I+zK)=\sum_{j=1}^{\infty}\frac{(-1)^{j+1}}{j}z^j\,\mathrm{Tr}_*K^j$, *valid for $|z|$ sufficiently small (depending on K and $\|\cdot\|_*$).*

Proof. The first property follows directly from continuity and the fact that the identity is correct in $\mathcal{B}_F$.

For the second property, first recall that the identity holds if $K \in \mathcal{B}_F$ and z is sufficiently small. Since the right-hand side is a well-defined analytic function for z in a neighborhood of the origin, the identity extends to that neighborhood. For a general $K \in \mathcal{B}_*$, take $F_n \in \mathcal{B}_F$ such that $F_n \to K$ as $n \to \infty$:

$$\log \det_*(I + zK) = \lim_{n\to\infty} \log \det(I + zF_n) = \lim_{n\to\infty} \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j} z^j \operatorname{Tr} F_n^j.$$

What is left to prove is that we can take the limit term by term. This follows from the estimate

$$|\operatorname{Tr} F_n^j| \le c_1 \|F_n\|_*^j,$$

and thus the series converges uniformly in n for z in a sufficiently small disk. We can therefore interchange limits. Then, by continuity of the trace we find $\operatorname{Tr} F_n^j \to \operatorname{Tr}_* K^j$ and this proves the statement. □

A consequence of the latter lemma is that $z \mapsto \det_*(I + zK)$ is analytic in a sufficiently small disk around the origin. In fact, it is an entire function, as we now prove.

Theorem 76. *For $K \in \mathcal{B}_*$ the map $z \mapsto \det_*(I + zK)$ is an entire function.*

Proof. Let $\{F_n\}_n \subset \mathcal{B}_F$ be such that $F_n \to K$ in $\mathcal{B}_*$ as $n \to \infty$. We will prove that $z \mapsto \det_*(I + zK)$ is analytic on a disk around the origin of arbitrary radius $R > 0$. Fix $R > 0$ and choose N such that $\|K - F_n\|_* \le 1/R$ for all $n \ge N$. Then, for $n \ge N$ we write

$$\det_*(I + zK) = \det_*\big(I + z(K - F_n)\big) \det_*\Big(I + z\big(I + z(K - F_n)\big)^{-1} F_n\Big). \tag{24}$$

By Lemma 75 we know that $\det_*\big(I + z(K - F_n)\big)$ is analytic for $|z| < R$. It remains to show that the second factor is also analytic. To this end, write

$$\det_*\Big(I + z\big(I + z(K - F_n)\big)^{-1} F_n\Big) = \lim_{M\to\infty} \det_*\Big(I + z\sum_{j=0}^{M} (-z)^j (K - F_n)^j F_n\Big),$$

where the convergence is uniform for z in compact subsets of the disk $|z| < R$. Now $\mathcal{B}_F$ is an ideal in $\mathcal{B}_*$ and thus

$$T_{n,M}(z) := z\sum_{j=0}^{M} (-z)^j (K - F_n)^j F_n \in \mathcal{B}_F,$$

and hence

$$\det_*\big(I + T_{n,M}(z)\big) = \det\big(I + T_{n,M}(z)\big)$$

is a polynomial in z. Moreover, by uniform convergence, these polynomials converge to an analytic function for $|z| < R$.

We thus see that both terms on the right-hand side of (24) are analytic for $|z| < R$ and thus so is $\det_*(I + zK)$. Since R is arbitrary the statement follows. □

Corollary 77. *For every $K \in \mathcal{B}_*$ we have an entire series expansion*

$$\det_*(I + zK) = 1 + \sum_{k=1}^{\infty} a_k(K)\, z^k,$$

with

$$\lim_{k\to\infty}\left|a_k(K)\right|^{1/k}=0.$$

THEOREM 78. *Let $K\in\mathcal{B}_*$ and $\{F_n\}_n\subset\mathcal{B}_F$ such that $F_n\to K$ as $n\to\infty$. Then*

$$\det{}_*(I+K)=1+\sum_{m=1}^{\infty}\frac{1}{m!}\lim_{n\to\infty}\int_S\cdots\int_S\det\bigl(F_n(x_i,x_j)\bigr)_{i,j=1}^{m}\,d\mu(x_1)\cdots d\mu(x_m).$$

It is tempting to try to interchange the limit with the integral and simply replace the integrand by $\det\bigl(K(x_i,x_j)\bigr)_{i,j=1}^{m}$, but this is not necessarily correct (for instance, for general trace class operators this does not necessarily hold). However, for certain explicit choices for $\mathcal{B}_*$ this can be justified, as we will see later on.

Another theorem that is useful to us is the following.

THEOREM 79. *Suppose that for some monotone increasing function $G:[0,\infty)\to[0,\infty)$ we have that*

$$|\det(I+\mathcal{K})|\le G(\|\mathcal{K}\|_*)$$

for all $\mathcal{K}\in\mathcal{B}_F$. Then, for any $\mathcal{K}_1,\mathcal{K}_2\in\mathcal{B}_*$,

$$|\det{}_*(I+\mathcal{K}_1)-\det{}_*(I+\mathcal{K}_2)|\le\|\mathcal{K}_1-\mathcal{K}_2\|_*\,G\bigl(\|\mathcal{K}_1\|_*+\|\mathcal{K}_2\|_*+1\bigr).$$

PROOF. For the trace, the corresponding statement is immediate from linearity and continuity.

For the determinant, fix $\mathcal{K}_1,\mathcal{K}_2\in\mathcal{B}_*$ and define

$$f(z)=\det{}_*\Bigl(I+\tfrac{1}{2}(\mathcal{K}_1+\mathcal{K}_2)+z(\mathcal{K}_2-\mathcal{K}_1)\Bigr).$$

Then f is entire in z and $f(-\frac{1}{2})=\det_*(I+\mathcal{K}_1)$ and $f(\frac{1}{2})=\det_*(I+\mathcal{K}_2)$.

Moreover,

$$f\left(\tfrac{1}{2}\right)-f\left(-\tfrac{1}{2}\right)=f'(\xi)$$

for some $\xi\in(-\frac{1}{2},\frac{1}{2})$ by the mean value theorem. Since f is entire, we can use Cauchy's inequality to deduce

$$\sup_{|\xi|\le 1/2}|f'(\xi)|\le\frac{1}{R}\sup_{|z|=R+1/2}|f(z)|,$$

for any $R>0$. Take $R=\|\mathcal{K}_1-\mathcal{K}_2\|_*^{-1}$ (if $\mathcal{K}_1\ne\mathcal{K}_2$; otherwise there is nothing to prove) and invoke the assumed bound on the determinant to obtain

$$\begin{aligned}\left|f\left(\tfrac{1}{2}\right)-f\left(-\tfrac{1}{2}\right)\right|&\le\|\mathcal{K}_1-\mathcal{K}_2\|_*\sup_{|z|=R+1/2}|f(z)|\\&\le\|\mathcal{K}_1-\mathcal{K}_2\|_*\sup_{|z|=R+1/2}G\left(\left\|\tfrac{1}{2}(\mathcal{K}_1+\mathcal{K}_2)+z(\mathcal{K}_2-\mathcal{K}_1)\right\|_*\right).\end{aligned}$$

By using the fact that G is monotone increasing we further deduce that

$$\sup_{|z|=R+1/2}G\left(\left\|\tfrac{1}{2}(\mathcal{K}_1+\mathcal{K}_2)+z(\mathcal{K}_2-\mathcal{K}_1)\right\|_*\right)\le G\bigl(\|\mathcal{K}_1\|_*+\|\mathcal{K}_2\|_*+1\bigr),$$

and this gives the stated inequality. □

5. Fredholm determinants

Let us now assume that $S \subset \mathbb{R}^d$ is compact and the kernel K is continuous. We denote the space of all integral operators with continuous kernel by $\mathcal{B}_\Phi$ and endow this space with the norm

$$\|\mathcal{K}\|_\Phi = \mu(S) \max_{x,y \in S} |K(x,y)|.$$

Equipped with this norm, $\mathcal{B}_\Phi$ is a Banach algebra, i.e., it is both a Banach space and an algebra, and

$$\|K_1 K_2\|_\Phi \leq \|K_1\|_\Phi \|K_2\|_\Phi.$$

Note also that $\|K\|_\infty \leq \|\mathcal{K}\|_\Phi$.

Lemma 80. *The subspace $\mathcal{F}$ of finite rank operators in $\mathcal{B}_\Phi$ forms a dense subalgebra.*

Proof. That $\mathcal{F}$ is a subalgebra is immediate. That it is dense follows, for instance, by Féjer's theorem from Fourier analysis. Indeed, using a multivariate version of Féjer's theorem we see that any continuous K can be uniformly approximated by sums of products of trigonometric polynomials. Since these sums give finite rank operators, we obtain the statement. □

Proposition 81. *The trace and determinant on $\mathcal{B}_\Phi \cap \mathcal{B}_F$ are continuous with respect to the norm $\|\cdot\|_\Phi$ and can therefore be continuously extended to $\mathcal{B}_\Phi$. These extensions will be denoted by* Tr_Φ *and* $\det_\Phi$.

The trace $\mathrm{Tr}_\Phi \mathcal{K}$ *satisfies*

$$\mathrm{Tr}_\Phi \mathcal{K} = \int_S K(x,x)\, d\mu(x). \tag{25}$$

The Fredholm determinant $\det_\Phi(I + \mathcal{K})$ *satisfies*

$$\det_\Phi(I+\mathcal{K}) = \sum_{m=0}^{\infty} \frac{1}{m!} \int_S \cdots \int_S \det\big(K(x_i,x_j)\big)_{i,j=1}^m \, d\mu(x_1)\cdots d\mu(x_m), \tag{26}$$

and the right-hand side is absolutely convergent.

Proof. It follows from Lemma 68 (second property) that

$$|\mathrm{Tr}_\Phi \mathcal{K}_1 - \mathrm{Tr}_\Phi \mathcal{K}_2| \leq \|\mathcal{K}_1 - \mathcal{K}_2\|_\Phi,$$

for any $\mathcal{K}_1, \mathcal{K}_2 \in \mathcal{B}_\Phi \cap \mathcal{B}_F$. This means that we can indeed extend the definition of the trace and determinant continuously to all of $\mathcal{B}_\Phi$.

It remains to check that the formulas (25) and (26) are indeed correct. For $\mathcal{K} \in \mathcal{B}_\Phi$, let $\mathcal{K}_n \in \mathcal{B}_\Phi \cap \mathcal{B}_F$ be a sequence converging to $\mathcal{K}$ in $\|\cdot\|_\Phi$. Then

$$\mathrm{Tr}_\Phi \mathcal{K} = \lim_{n\to\infty} \mathrm{Tr} \mathcal{K}_n = \lim_{n\to\infty} \int_S K_n(x,x)\, d\mu(x),$$

and

$$\det_\Phi(I+\mathcal{K}) = \lim_{n\to\infty} \sum_{m=0}^{\infty} \frac{1}{m!} \int_S \cdots \int_S \det\big(K_n(x_i,x_j)\big)_{i,j=1}^m \, d\mu(x_1)\cdots d\mu(x_m).$$

By uniform convergence of $K_n \to K$, we can take the limit under the integrals, which proves the statement. □

LEMMA 82. *Let* $G(x) = \sum_{m=0}^{\infty} \frac{m^{m/2} x^m}{m!}$. *Then*

$$|\det_\Phi(I + \mathcal{K})| \le G(\|\mathcal{K}\|_\Phi).$$

PROOF. By Hadamard's inequality we have

$$\left|\det\big(K(x_i, x_j)\big)_{i,j=1}^m\right| \le m^{m/2} \, \|\mathcal{K}\|_\Phi^m.$$

Together with Lemma 70 this implies that

$$|\det(I + \mathcal{K})| \le \sum_{m=0}^{\infty} \frac{m^{m/2} \, \|\mathcal{K}\|_\Phi^m}{m!}, \tag{27}$$

proving the statement. □

We list the most important properties.

LEMMA 83. *For* $K_1, K_2, K \in \mathcal{B}_\Phi$ *we have*

(1) $|\mathrm{Tr}_\Phi \mathcal{K}_1 - \mathrm{Tr}_\Phi \mathcal{K}_2| \le \|\mathcal{K}_1 - \mathcal{K}_2\|_\Phi$;
(2) $|\det_\Phi(I + \mathcal{K}_1) - \det_\Phi(I + \mathcal{K}_2)| \le \|\mathcal{K}_1 - \mathcal{K}_2\|_\Phi \, G\big(\|\mathcal{K}_1\|_\Phi + \|\mathcal{K}_2\|_\Phi + 1\big)$;
(3) $\det_\Phi(I + \mathcal{K}_1)\det_\Phi(I + \mathcal{K}_2) = \det_\Phi(I + \mathcal{K}_1 + \mathcal{K}_2 + \mathcal{K}_1\mathcal{K}_2)$;
(4) $\log\det_\Phi(I + zK) = \displaystyle\sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j} z^j \, \mathrm{Tr}_\Phi K^j$, *valid for* $|z|$ *sufficiently small (depending on* K *and* $\|\cdot\|_\Phi$*).*

Here $G(x) = \sum_{m=0}^{\infty} \frac{m^{m/2} x^m}{m!}$.

6. A generalization

We now consider a generalization of the Fredholm determinant.

Consider the space $\mathcal{B}_R$ of kernels $K : S \times S \to \mathbb{C}$ such that

$$\|K\|_R = \max\left(\left(\iint_{S\times S} |K(x,y)|^2 \, d\mu(x)\, d\mu(y)\right)^{1/2}, \int_S |K(x,x)| \, d\mu(x)\right) < \infty.$$

After introducing equivalence classes by saying that $K_1 = K_2$ if and only if

(1) $K_1(x, y) = K_2(x, y)$ for $\mu \times \mu$–almost every $(x, y) \in S \times S$, and
(2) $K_1(x, x) = K_2(x, x)$ for μ–almost every $x \in S$,

we see that $\mathcal{B}_R$ is a Banach space.

We stress that, in general, one should not consider the elements of $\mathcal{B}_R$ as integral kernels for bounded operators. Indeed, if the diagonal $D = \{(x, x) \mid x \in S\}$ is a set of measure 0 then the values of K on D do not contribute to the operator. But as an element of $\mathcal{B}_R$ these values are highly relevant. For instance, if we take $K(x, y) = 0$ for $x \ne y$ and assume that D has measure 0, then the corresponding operator is trivial, but its norm in $\mathcal{B}_R$ is positive if $K(x, x) \ne 0$ for x in a set of positive measure.

This is not a problem. In fact, it will turn out to be very useful for us. Indeed, for determinantal point processes the values along the diagonal are mean densities of points, and thus highly relevant. From this point of view it is somewhat unfortunate to view the kernel only as an operator.

THEOREM 84. *The set*

$$\mathcal{B}_F = \left\{ \sum_{j=1}^{m} f_j(x)\, g_j(y) \;\middle|\; m \in \mathbb{N},\ f_1, \ldots, f_m, g_1, \ldots, g_m \in L_2(S, \mu) \right\}$$

is dense in $\mathcal{B}_R$.

THEOREM 85. *The trace and determinant are continuous with respect to the norm $\|\cdot\|_R$ and can therefore be continuously extended to $\mathcal{B}_R$. These extensions will be denoted by Tr_R and $\det_R$.*

The trace $\mathrm{Tr}_R \mathcal{K}$ satisfies

$$\mathrm{Tr}_R \mathcal{K} = \int_S K(x,x)\, d\mu(x). \tag{28}$$

The determinant $\det_R(I + \mathcal{K})$ satisfies

$$\det{}_R(I+\mathcal{K}) = \sum_{m=0}^{\infty} \frac{1}{m!} \int_S \cdots \int_S \det\bigl(K(x_i, x_j)\bigr)_{i,j=1}^{m} d\mu(x_1) \cdots d\mu(x_m), \tag{29}$$

and the right-hand side is absolutely convergent.

7. Trace class operators

In the next discussion, we return to the more general situation where K is a compact operator on $L_2(S, \mu)$ and S is a general measure space. If K is compact, then K^*K is both compact and self-adjoint. By the spectral theorem K^*K has eigenvalues $\sigma_j(K)^2$ that we order in decreasing order

$$\sigma_1(K)^2 \geq \sigma_2(K)^2 \geq \ldots \geq 0.$$

The positive square roots $\sigma_j(K)$ are called the singular values of K. The p–Schatten class of operators is then defined as the set of all operators K such that

$$\sum_{j=1}^{\infty} \sigma_j(K)^p < \infty,$$

and we define

$$\|K\|_p = \left(\sum_{j=1}^{\infty} \sigma_j(K)^p \right)^{1/p}.$$

We also define

$$\mathcal{B}_p = \{K \text{ compact } \mid \|K\|_p < \infty\}.$$

In case $p = 1$, the operators in $\mathcal{B}_1$ are called trace class operators and $\|\cdot\|_1$ is called the trace norm. In case $p = 2$, the operators in $\mathcal{B}_2$ are called Hilbert–Schmidt operators and $\|\cdot\|_2$ is called the Hilbert–Schmidt norm. It can be shown that

$$\|K\|_2^2 = \iint_{S \times S} |K(x,y)|^2 \, d\mu(x)\, d\mu(y),$$

but no such simple expression for the trace norm is known. In fact, it is often rather difficult to verify that a certain operator is trace class. A useful idea is to show that $K = AB$ where A, B are Hilbert–Schmidt operators. Indeed, in that case K is trace class since

$$\|K\|_1 = \|AB\|_1 \leq \|A\|_2 \|B\|_2.$$

This implies in particular that if K is Hilbert–Schmidt, then K^j is trace class for all $j \geq 2$.

The finite rank operators on $L_2(S, \mu)$ are dense in the trace class operators. Moreover, it can be shown that if we order the eigenvalues of K as $|\lambda_1| \geq |\lambda_2| \geq \dots$, then $\sigma_j(K) \geq |\lambda_j(K)|$. This implies that for finite rank operators K we have

$$|\mathrm{Tr} K| \leq \|K\|_1.$$

The trace is thus Lipschitz continuous with respect to the trace norm and we can continuously extend the definition of the trace to all trace class operators. This extension will be denoted by Tr_1.

As for the determinant, if $\mathcal{K} \in \mathcal{B}_1$ has eigenvalues λ_j, we have

$$|\det(I+\mathcal{K})| = \left|\prod_{j=1}^{\infty}(1+\lambda_j)\right| \leq \exp\Big(\sum_{j=1}^{\infty}|\lambda_j|\Big) \leq \exp\Big(\sum_{j=1}^{\infty}\sigma_j(K)\Big) \leq \exp\big(\|\mathcal{K}\|_1\big).$$

As a consequence we also see that the determinant can be extended to a locally Lipschitz continuous function on $\mathcal{B}_1$ and this extension will be denoted by $\det_1$.

Lemma 86. *For $K_1, K_2, K \in \mathcal{B}_1$ we have*

(1) $|\mathrm{Tr}_1 K_1 - \mathrm{Tr}_1 K_2| \leq \|K_1 - K_2\|_1$;
(2) $|\det_1(I+K_1) - \det_1(I+K_2)| \leq \|K_1 - K_2\|_1 \exp\big(\|K_1\|_1 + \|K_2\|_1 + 1\big)$;
(3) $\det_1(I+K_1)\det_1(I+K_2) = \det_1(I+K_1+K_2+K_1K_2)$;
(4) $\log\det_1(I+zK) = \sum_{j=1}^{\infty}\frac{(-1)^{j+1}}{j}z^j\mathrm{Tr}_1 K^j$, *valid for z in a sufficiently small neighborhood of* 0.

8. Comparison between $\det_\Phi$, $\det_R$ and $\det_1$

It is important to note that $\det_\Phi$, $\det_R$ and $\det_1$ are not necessarily the same. To start with, they are not necessarily defined on the same sets. Clearly, not all trace class operators have continuous kernels. Moreover, Carleman showed the existence of a continuous K with $S = [0,1]$ and Lebesgue measure μ, such that $\mathcal{K}$ is not trace class. For such operators $\det_\Phi = \det_R$ is well defined but $\det_1$ is not. But even when all determinants and traces are well defined, they do not necessarily take the same value.

Proposition 87. *There exists $K \in \mathcal{B}_R$ such that $\mathcal{K}$ is trace class, but $\mathrm{Tr}_R K \neq \mathrm{Tr}_1 K$ and $\det_R(I+K) \neq \det_1(I+K)$.*

It is important to note that such examples are not pathological, but are relevant in our discussion of determinantal point processes.

We end this section with some positive results.

Proposition 88. *If K is continuous and S is compact, then $K \in \mathcal{B}_R$ and*

$$\mathrm{Tr}_R K = \mathrm{Tr}_\Phi K, \qquad \det_R(I+K) = \det_\Phi(I+K).$$

Proposition 89. *If K is continuous and Hermitian non-negative, then $\det_\Phi = \det_1 = \det_*$ and $\mathrm{Tr}_\Phi = \mathrm{Tr}_1 = \mathrm{Tr}_*$ (whenever these are defined).*

Proposition 90. *If K is continuous and $\mathcal{K}$ is trace class, then $\det_\Phi = \det_1 = \det_*$ and $\mathrm{Tr}_\Phi = \mathrm{Tr}_1 = \mathrm{Tr}_*$.*

PROPOSITION 91. *If K is continuous and Hermitian non-negative, then $\mathcal{K}$ is trace class.*

The fact that K is Hermitian is important here, as the proof relies on Mercer's theorem.

APPENDIX B

Laurent, Toeplitz, and Hankel operators

0.1. Laurent operators and symbols. Let $\ell^2(\mathbb{Z})$ be the Hilbert space of square-summable sequences $f = (f_k)_{k\in\mathbb{Z}}$. A (bi-infinite) matrix $L = (L_{ij})_{i,j\in\mathbb{Z}}$ is called a *Laurent matrix* if L_{ij} depends only on $i - j$, i.e.

$$L_{ij} = a_{i-j} \qquad (i, j \in \mathbb{Z})$$

for some sequence $(a_k)_{k\in\mathbb{Z}}$. The corresponding bounded operator on $\ell^2(\mathbb{Z})$ (when it exists) is called a *Laurent operator* and is denoted $L(a)$, where a is the *symbol*

$$a(z) = \sum_{k\in\mathbb{Z}} a_k z^k, \qquad z \in \mathbb{T}.$$

When $(a_k) \in \ell^1(\mathbb{Z})$, the series converges absolutely on $\mathbb{T}$ and $L(a)$ is bounded, acting by convolution:

$$(L(a)f)_i = \sum_{k\in\mathbb{Z}} a_{i-k} f_k.$$

In this case one has the algebra property

$$L(a)\,L(b) = L(ab),$$

reflecting the fact that Laurent operators diagonalize under the Fourier transform.

A convenient Banach algebra for symbols is the *Wiener algebra*

$$\mathcal{W} := \Big\{a(z) = \sum_{k\in\mathbb{Z}} a_k z^k : \sum_{k\in\mathbb{Z}} |a_k| < \infty\Big\}, \qquad \|a\|_{\mathcal{W}} = \sum_{k\in\mathbb{Z}} |a_k|.$$

If $a \in \mathcal{W}$ and $a(z) \neq 0$ for all $z \in \mathbb{T}$, then $1/a \in \mathcal{W}$ (Wiener's lemma). Consequently,

$$L(a)^{-1} = L(1/a)$$

on $\ell^2(\mathbb{Z})$. This is one reason why operator-level inverses are often simpler than finite-matrix inverses.

0.2. Toeplitz and Hankel operators. Let $\ell^2(\mathbb{Z}_{\geq 0})$ be the Hardy-type half-space, and let $P_+ : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ be the orthogonal projection onto indices ≥ 0. For a Laurent operator $L(a)$, the associated *Toeplitz operator* is the compression

$$T(a) := P_+\,L(a)\,P_+\Big|_{\ell^2(\mathbb{Z}_{\geq 0})}.$$

In matrix form, $T(a)$ has entries

$$T(a)_{ij} = a_{i-j} \qquad (i, j \geq 0),$$

so it is constant along diagonals. The *finite Toeplitz matrix* of size n is the further compression

$$T_n(a) := P_n\,T(a)\,P_n,$$

where P_n projects onto $\{0, 1, \ldots, n-1\}$.

The *Hankel operator* associated with a is defined by

$$H(a) := P_+\, L(a)\, J\, P_+\Big|_{\ell^2(\mathbb{Z}_{\geq 0})},$$

where J is the flip $(Jf)_k = f_{-k-1}$ on $\ell^2(\mathbb{Z})$. In matrix form,

$$H(a)_{ij} = a_{i+j+1} \qquad (i, j \geq 0),$$

so it is constant along anti-diagonals.

A fundamental algebraic identity relating these objects is

$$T(a)\,T(b) = T(ab) \;+\; H(a)\,H(\widetilde{b}), \qquad \widetilde{b}(z) := b(z^{-1}), \tag{30}$$

valid under mild summability assumptions (e.g. $a, b \in \mathcal{W}$). The additional Hankel term measures the failure of the Toeplitz compression to be multiplicative.

0.3. Wiener–Hopf factorization and invertibility of Toeplitz operators. Write the Fourier decomposition

$$a(z) = a_-(z)\, a_+(z),$$

where

$$a_+(z) = \sum_{k\geq 0} a_k^{(+)} z^k, \qquad a_-(z) = \sum_{k\leq 0} a_k^{(-)} z^k,$$

and both a_+ and a_- are nonvanishing on $\mathbb{T}$. Such a factorization is called a *Wiener–Hopf factorization.* In the Wiener algebra $\mathcal{W}$ one can formulate it as

$$a \in \mathcal{W},\ a(z) \neq 0\ \forall z \in \mathbb{T},\ \mathrm{wind}(a, 0) = 0 \quad \Longrightarrow \quad a = a_- a_+, \quad a_\pm^{\pm 1} \in \mathcal{W},$$

where $\mathrm{wind}(a, 0)$ is the winding number of $a(\mathbb{T})$ around 0. When $\mathrm{wind}(a, 0) \neq 0$, a canonical factorization includes an extra monomial z^κ.

When $a = a_- a_+$ with $a_\pm^{\pm 1} \in \mathcal{W}$ and $\mathrm{wind}(a, 0) = 0$, the Toeplitz operator $T(a)$ is invertible on $\ell^2(\mathbb{Z}_{\geq 0})$, and one has the operator identity

$$T(a)^{-1} = T(a_+^{-1})\, T(a_-^{-1}) \quad \text{modulo a compact correction (in fact Hankel-type).} \tag{31}$$

More precisely, using (30) one finds that $T(a_+^{-1})T(a_-^{-1})$ is a two-sided inverse of $T(a)$ up to a correction expressed through Hankel operators built from $a_\pm^{\pm 1}$. This is the analytic core of the Wiener–Hopf method for Toeplitz operators.

0.4. Finite Toeplitz matrices and approximate inverses as $n \to \infty$. Even when $T(a)$ is easy to invert at the operator level, the explicit inverse of the finite matrix $T_n(a)$ can be difficult. A useful viewpoint is that $T_n(a)$ is a finite section of $T(a)$, and $T_n(a)^{-1}$ should be close to the corresponding finite section of $T(a)^{-1}$, up to boundary effects near 0 and $n - 1$.

Assume $a \in \mathcal{W}$ is nonvanishing on $\mathbb{T}$ with $\mathrm{wind}(a, 0) = 0$, and choose a Wiener–Hopf factorization $a = a_- a_+$ with $a_\pm^{\pm 1} \in \mathcal{W}$. Then the following principles are standard.

(1) Finite section method. Under the assumptions above, the finite sections are stable:

$$T_n(a) \text{ is invertible for all sufficiently large } n, \qquad \sup_{n\geq n_0} \|T_n(a)^{-1}\| < \infty.$$

Moreover, in a strong sense,

$$T_n(a)^{-1} \;\approx\; P_n\, T(a)^{-1}\, P_n \qquad (n\to\infty),$$

with an error concentrated near the boundary (hence effectively low-rank after suitable truncations).

(2) Gohberg–Semencul-type structure. There are explicit representations of $T_n(a)^{-1}$ in terms of Toeplitz and Hankel matrices built from the Fourier coefficients of $a_\pm^{\pm 1}$, showing that

$$T_n(a)^{-1} = T_n(a_+^{-1})\, T_n(a_-^{-1}) \;+\; (\text{boundary correction}),$$

where the boundary correction has small rank compared to n (often rank at most $2r$ for symbols with finite bandwidth, and more generally a correction whose operator norm decays as $n\to\infty$ under additional regularity/analyticity).

(3) Exponentially small errors for analytic symbols. If a extends analytically and nonvanishingly to an annulus $\{r < |z| < R\}$ with $r < 1 < R$, then the Fourier coefficients of $a_\pm^{\pm 1}$ decay exponentially, and the boundary correction above becomes exponentially small in n (in operator norm or entrywise, depending on the formulation). This is the regime in which approximate inverses are particularly effective.

In applications, one typically combines the operator factorization (which yields a manageable expression for $T(a)^{-1}$) with one of the finite-n approximation principles above. This allows one to control kernels or partition functions in the limit $n\to\infty$ without requiring an explicit closed form for $T_n(a)^{-1}$.

0.5. The special case of rational symbols. An important and particularly tractable situation arises when the symbol $a(z)$ is a rational function of z. In this case many of the abstract constructions discussed above admit explicit realizations, and several difficulties present for general symbols disappear.

Assume that $a(z)$ is rational and has no zeros or poles on the unit circle. Then a automatically belongs to the Wiener algebra $\mathcal{W}$, and its Wiener–Hopf factorization can be carried out explicitly by separating the zeros and poles inside and outside the unit disk. More precisely, one can write

$$a(z) = c\, z^{\kappa} \prod_{j=1}^{J} \frac{1-\overline{\alpha_j} z}{1-\alpha_j z} \prod_{k=1}^{K} \frac{1-\beta_k z^{-1}}{1-\overline{\beta_k} z^{-1}},$$

where $|\alpha_j| < 1$, $|\beta_k| < 1$, $c \neq 0$, and $\kappa \in \mathbb{Z}$ is the winding number of a. From this representation the Wiener–Hopf factors a_+ and a_- are obtained by grouping together the factors that are analytic inside and outside the unit disk, respectively.

In the rational case the associated Toeplitz and Hankel operators have additional structure. In particular, the Hankel operators $H(a)$ and $H(\widetilde{a})$ have finite rank, with rank equal to the total number of poles of a inside or outside the unit disk. As a consequence, the correction terms appearing in identities such as

$$T(a)^{-1} = T(a_+^{-1})\, T(a_-^{-1}) + (\text{Hankel correction})$$

are finite-rank operators.

This finite-rank property has two important implications. First, the inverse Toeplitz operator $T(a)^{-1}$ differs from the simple product $T(a_+^{-1})T(a_-^{-1})$ only by an explicitly computable, finite-dimensional correction. Second, for finite Toeplitz matrices $T_n(a)$, the boundary corrections in the approximate inverse are uniformly finite rank, independent of n. In particular, away from the boundary the entries of $T_n(a)^{-1}$ stabilize exactly once n is sufficiently large.

In asymptotic problems, and especially in determinantal point process limits, rational symbols therefore lead to kernels that are finite-rank perturbations of translation-invariant kernels. This feature is often crucial for explicit computations and for identifying integrable structures in the limiting process.

Bibliography

[1] I. Gohberg, S. Goldberg and N. Krupnik, Trace and Determinants of Linear operators.
[2] D.J. Daley and D. Verre-Jones, An Introduction to the Theory of Point Processes Volume II