

Lecture Notes on the Corner Growth Model

Timo Seppäläinen

February 8, 2009

MATHEMATICS DEPARTMENT, UNIVERSITY OF WISCONSIN-MADISON
E-mail address: `seppalai@math.wisc.edu`

2000 *Mathematics Subject Classification.* Primary 60K35
Key words and phrases. interacting particle systems

©Copyright Timo Seppäläinen.

ABSTRACT. These notes cover limit properties of the corner growth model. Version February 8, 2009.

Contents

Notation	vii
Chapter 1. The corner growth model and some of its relatives	1
1.1. The corner growth model	1
Comments	6
Chapter 2. Deterministic large scale limits	7
2.1. Law of large numbers	7
2.2. Explicit limit for geometric weights	10
2.3. M/M/1 queues in series	17
2.4. Hydrodynamic limit for M/M/1 queues in series	21
2.5. Queues and the last-passage model revisited	31
Comments	34
Exercises	35
Chapter 3. The last-passage Markov chain	37
3.1. Derivation of the determinantal transition probabilities	39
3.2. From determinantal formula to orthogonal polynomial ensemble	45
Comments	48
Chapter 4. Tracy-Widom distribution	49
4.1. Airy function and kernel	49
4.2. Tracy-Widom distribution	51
Comments	53
Chapter 5. Distributional limit for the last-passage time	55
5.1. From orthogonal polynomial ensemble to Meixner kernel	56
5.2. Meixner polynomials	57
5.3. Airy asymptotics for the Meixner kernel	63
Comments	81
Appendix A. Probability theory	83
Appendix B. Linear algebra and determinants	87
Appendix C. Analysis	93
C.1. Convex functions	93
C.2. Complex variables	95
Appendix D. Orthogonal Polynomials	99
Bibliography	105

Notation

Some general conventions. Constant C can sometimes change from line to line. Analogously, the meaning of ε as a quantity that can be chosen arbitrarily small may also change within the same proof.

$(a)_{[n]}$	descending factorial: $(a)_{[0]} = 1$, and $(a)_{[n]} = a(a-1)\cdots(a-n+1)$ for $n \in \mathbb{N}$.
\mathbb{N}	the set $\{1, 2, 3, \dots\}$ of natural numbers
$[n]$	the set $\{1, 2, 3, \dots, n\}$ for $n \in \mathbb{N}$
$x_{1,n}$	shorthand for the n -vector (x_1, \dots, x_n)
\mathbb{Z}_+	the set $\{0, 1, 2, 3, \dots\}$ of nonnegative integers
$\bar{1}$	the vector $(1, 1, \dots, 1)$ of all ones
\setminus	set difference: $A \setminus B = A \cap B^c$

The corner growth model and some of its relatives

In this introductory chapter we define the corner growth model and then discuss somewhat informally models related to it. The precise mathematical study of the corner growth model begins in Chapter 2 with laws of large numbers.

1.1. The corner growth model

The *corner growth model* is a simple mathematical model of a randomly growing cluster that over time invades the entire first quadrant of the plane. The first quadrant of the plane is represented here by \mathbb{N}^2 , that is, by the set of points (i, j) with positive integer coordinates. At the outset each point $(i, j) \in \mathbb{N}^2$ is given a weight, or waiting time $Y_{i,j}$. See Figure 1. The values $\{Y_{i,j} : (i, j) \in \mathbb{N}^2\}$ are nonnegative random variables.

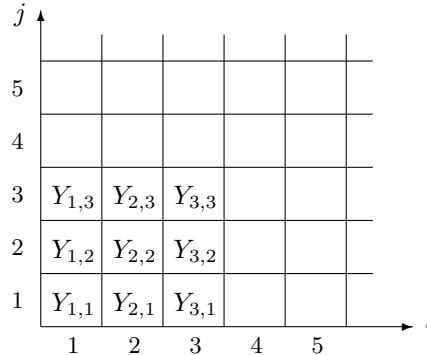


FIGURE 1. A portion of \mathbb{N}^2 . The squares represent points $(i, j) \in \mathbb{N}^2$. Each point (i, j) has a weight $Y_{i,j}$ attached to it.

The values $\{Y_{i,j}\}$ determine the evolution of a growing subset of \mathbb{N}^2 called the *cluster* whose value at time t is denoted by $\mathcal{B}(t)$. The general rule is that $Y_{i,j}$ is the time it takes to occupy point (i, j) but only *after* its two neighbors to the left and below are either occupied or lie outside \mathbb{N}^2 . So at the boundaries the rule is that point $(1, 1)$ needs no occupied neighbors to start, points $(1, j)$ on the left boundary wait only for the neighbor below to be occupied, and points $(i, 1)$ on the bottom boundary wait only for the left neighbor to be occupied. Alternatively, we can imagine that the outside boundary points $\{(i, j) : i = 0 \text{ or } j = 0\}$ are occupied at the outset.

To illustrate, if all $Y_{i,j} > 0$, then initially the cluster is empty: $\mathcal{B}(0) = \emptyset$. At time $Y_{1,1}$ the cluster occupies point $(1, 1)$:

$$\mathcal{B}(t) = \emptyset \text{ for } 0 \leq t < Y_{1,1}, \mathcal{B}(Y_{1,1}) = \{(1, 1)\}.$$

Next in line are points $(1, 2)$ and $(2, 1)$. Point $(1, 2)$ is occupied at time $Y_{1,1} + Y_{1,2}$ and point $(2, 1)$ is occupied at time $Y_{1,1} + Y_{2,1}$. And so on.

By contrast, if values $Y_{i,j} = 0$ are possible then the cluster may invade some points immediately: $\mathcal{B}(0) = \{(k, \ell) \in \mathbb{N}^2 : Y_{i,j} = 0 \forall (i, j) \in \{1, \dots, k\} \times \{1, \dots, \ell\}\}$.

Once occupied, a point remains occupied. Thus the growing cluster never loses points, only adds them. Such a model is called *totally asymmetric*.

It is convenient to describe the evolution in terms of the times when points join the cluster. Let $G(m, n)$ denote the time when point (m, n) becomes occupied. The above explanation is summarized by the equation

$$(1.1) \quad G(m, n) = G(m-1, n) \vee G(m, n-1) + Y_{m,n} \quad \text{for } (m, n) \in \mathbb{N}^2,$$

together with the boundary conditions

$$(1.2) \quad G(m, n) = 0 \quad \text{if } m = 0 \text{ or } n = 0.$$

Equation (1.1) can be iterated backwards until the corner $(1, 1)$ is reached, resulting in this *last-passage* formula for G :

$$(1.3) \quad G(m, n) = \max_{\pi \in \Pi(m, n)} \sum_{(i, j) \in \pi} Y_{i, j}, \quad (m, n) \in \mathbb{N}^2.$$

$\Pi(m, n)$ is the collection of nearest-neighbor up-right paths π from $(1, 1)$ to (m, n) . Figure 2 represents one such path for $(m, n) = (5, 4)$. Precisely speaking, an element π of $\Pi(m, n)$ is a sequence

$$\pi = \{(1, 1) = (i_1, j_1), (i_2, j_2), \dots, (i_{m+n-1}, j_{m+n-1}) = (m, n)\}$$

such that $(i_s, j_s) - (i_{s-1}, j_{s-1}) = (1, 0)$ or $(0, 1)$ for $s = 1, 2, \dots, m+n-1$. In terms of the last-passage times, the growing cluster at time t is given by

$$(1.4) \quad \mathcal{B}(t) = \{(m, n) \in \mathbb{N}^2 : G(m, n) \leq t\}.$$

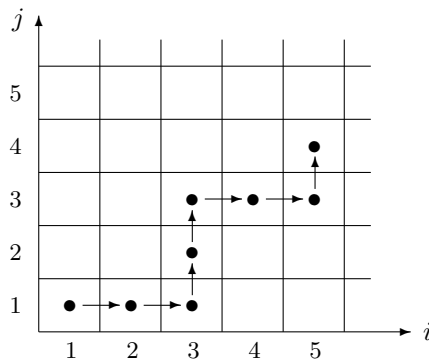


FIGURE 2. An admissible path from $(1, 1)$ to $(5, 4)$

More generally, we will consider last-passage times between two points (k, ℓ) and (m, n) in \mathbb{N}^2 such that $k \leq m$ and $\ell \leq n$:

$$(1.5) \quad G((k, \ell), (m, n)) = \max_{\pi \in \Pi((k, \ell), (m, n))} \sum_{(i, j) \in \pi} Y_{i, j}.$$

$\Pi((k, \ell), (m, n))$ is the collection of paths

$$\pi = \{(k, \ell) = (i_1, j_1), (i_2, j_2), \dots, (i_{m+n-k-\ell+1}, j_{m+n-k-\ell+1}) = (m, n)\}$$

such that $(i_s, j_s) - (i_{s-1}, j_{s-1}) = (1, 0)$ or $(0, 1)$ for $s = 1, 2, \dots, m + n - k - \ell + 1$. The earlier $G(m, n)$ is now $G((1, 1), (m, n))$ but we continue to use the notation $G(m, n)$ for this special case.

This model and others of its kind are called *directed last-passage percolation models*. “Directed” refers to the restrictions on admissible paths, and “last-passage” to the feature that the occupation time $G(m, n)$ is determined by the slowest path to (m, n) . (By contrast, in first-passage percolation occupation times are determined by quickest paths.)

When the weights $Y_{i,j}$ have exponential or geometric distribution, the growing cluster $\mathcal{B}(t)$ becomes a Markov chain in the state space of possible finite clusters in \mathbb{N}^2 . For later use in Chapter 2 we prove this claim here for the case of geometric weights, together with a description of the transition probability.

Fix a parameter $0 < p < 1$ with $q = 1 - p$ and let the independent random variables $\{Y_{i,j}\}$ have common probability distribution given by

$$(1.6) \quad \mathbf{P}\{Y_{i,j} = k\} = pq^{k-1}, \quad k \in \mathbb{N}.$$

As above, define the last-passage times $G(m, n)$ by (1.3) and the cluster process $\mathcal{B}(t)$ by (1.4). Since the last-passage times are integers, we can think of $\mathcal{B}(t)$ as a discrete-time process indexed by $t \in \mathbb{Z}_+$. Since each $Y_{i,j} \geq 1$, $\mathcal{B}(0) = \emptyset$ and for $t \geq 1$, $\mathcal{B}(t)$ is a subset of the square $[0, t] \times [0, t]$ and in particular a finite set. Thus $\mathcal{B}(t)$ is a discrete-time process in the countable state space

$$(1.7) \quad \Gamma = \{U \subseteq \mathbb{N}^2 : U \text{ is finite, and } (i, j) \in U \text{ implies that } U \text{ contains} \\ \text{the entire discrete rectangle } \{1, \dots, i\} \times \{1, \dots, j\}\}$$

We take the description above to include the empty set, so $\emptyset \in \Gamma$.

Let $U \in \Gamma$. A point $(m, n) \notin U$ is a *growth corner* or *growth site* for U if $\{1, \dots, m\} \times \{1, \dots, n\} \subseteq U \cup \{(m, n)\}$. In other words, both the left and lower neighbors of (m, n) lie either outside \mathbb{N}^2 or in U , and (m, n) can be added to U to create a new element $U \cup \{(m, n)\} \in \Gamma$.

PROPOSITION 1.1. *The process $\mathcal{B}(t)$ is a Markov chain on the state space Γ with initial state $\mathcal{B}(0) = \emptyset$ and with transition probability given by the following description: given $\mathcal{B}(t)$, $\mathcal{B}(t+1)$ is obtained by adding to $\mathcal{B}(t)$ each of its growth sites independently with probability p .*

PROOF. Let us use boldface symbols \mathbf{x}, \mathbf{y} to denote points of \mathbb{N}^2 . Fix $U \in \Gamma$ and let U' be the set of growth sites of U . Let $U' = L \cup M$ be an arbitrary partition of U' into two disjoint sets, one of which can be empty. We need to show that, for arbitrary sets $U_1, \dots, U_{t-1} \in \Gamma$ such that

$$\mathbf{P}\{\mathcal{B}(1) = U_1, \dots, \mathcal{B}(t-1) = U_{t-1}, \mathcal{B}(t) = U\} > 0$$

we have the Markov property:

$$(1.8) \quad \mathbf{P}[\mathcal{B}(t+1) = U \cup L \mid \mathcal{B}(1) = U_1, \dots, \mathcal{B}(t-1) = U_{t-1}, \mathcal{B}(t) = U] = p^{|L|} q^{|M|}.$$

There exist integers $s(\mathbf{y}) \leq t$ for $\mathbf{y} \in U$ such that

$$\begin{aligned} & \{\mathcal{B}(1) = U_1, \dots, \mathcal{B}(t-1) = U_{t-1}, \mathcal{B}(t) = U\} \\ &= \{G(\mathbf{y}) = s(\mathbf{y}) \text{ for } \mathbf{y} \in U \text{ and } G(\mathbf{x}) > t \text{ for } \mathbf{x} \in U'\}. \end{aligned}$$

Let $e_1 = (1, 0)$ and $e_2 = (0, 1)$ denote the standard basis vectors. For $\mathbf{x} \in U'$ let $S(\mathbf{x}) = G(\mathbf{x} - e_1) \vee G(\mathbf{x} - e_2)$ so that $G(\mathbf{x}) = S(\mathbf{x}) + Y_{\mathbf{x}}$. For the definition of $S(\mathbf{x})$ remember again the boundary conditions (1.2) if $\mathbf{x} - e_1$ or $\mathbf{x} - e_2$ lies outside \mathbb{N}^2 . For the calculation below note that the variables $\{S(\mathbf{x}) : \mathbf{x} \in U'\}$ are functions of $\{G(\mathbf{y}) : \mathbf{y} \in U\}$.

Write the conditional probability in (1.8) as a ratio of two probabilities. The numerator is

$$\begin{aligned} & \mathbf{P}\{\mathcal{B}(1) = U_1, \dots, \mathcal{B}(t-1) = U_{t-1}, \mathcal{B}(t) = U, \mathcal{B}(t+1) = U \cup L\} \\ &= \mathbf{P}\{G(\mathbf{y}) = s(\mathbf{y}) \text{ for } \mathbf{y} \in U, G(\mathbf{x}) = t+1 \text{ for } \mathbf{x} \in L \text{ and } G(\mathbf{x}) > t+1 \text{ for } \mathbf{x} \in M\} \\ &= \mathbf{P}\{G(\mathbf{y}) = s(\mathbf{y}) \text{ for } \mathbf{y} \in U, Y_{\mathbf{x}} = t+1 - S(\mathbf{x}) \text{ for } \mathbf{x} \in L \\ &\quad \text{and } Y_{\mathbf{x}} > t+1 - S(\mathbf{x}) \text{ for } \mathbf{x} \in M\} \end{aligned}$$

by the independence of $\{Y_{\mathbf{x}} : \mathbf{x} \in U'\}$ and $\{G(\mathbf{y}) : \mathbf{y} \in U\}$

$$\begin{aligned} &= \mathbf{E}\left[\prod_{\mathbf{y} \in U} \mathbf{1}\{G(\mathbf{y}) = s(\mathbf{y})\} \cdot \prod_{\mathbf{x} \in L} pq^{t-S(\mathbf{x})} \cdot \prod_{\mathbf{x} \in M} q^{t+1-S(\mathbf{x})}\right] \\ &= p^{|L|}q^{|M|} \mathbf{E}\left[\prod_{\mathbf{y} \in U} \mathbf{1}\{G(\mathbf{y}) = s(\mathbf{y})\} \cdot \prod_{\mathbf{x} \in U'} q^{t-S(\mathbf{x})}\right] \\ &= p^{|L|}q^{|M|} \mathbf{P}\{G(\mathbf{y}) = s(\mathbf{y}) \text{ for } \mathbf{y} \in U \text{ and } Y_{\mathbf{x}} > t - S(\mathbf{x}) \text{ for } \mathbf{x} \in U'\} \\ &= p^{|L|}q^{|M|} \mathbf{P}\{\mathcal{B}(1) = U_1, \dots, \mathcal{B}(t-1) = U_{t-1}, \mathcal{B}(t) = U\}. \end{aligned}$$

Equation (1.8) has been verified. \square

The reader should notice how crucially the last calculation depended on the special structure of the geometric distribution.

To complement the proof by calculation, here is the seasoned probabilist's hand-waiving proof of Proposition 1.1, by a clever construction. Give each point $\mathbf{x} \in \mathbb{N}^2$ a $\{0, 1\}$ -valued sequence $(\zeta_k^{\mathbf{x}})_{k \in \mathbb{N}}$ whose entries come from independent coin flips with success probability $\mathbf{P}\{\zeta_k^{\mathbf{x}} = 1\} = p$. Let the initial cluster be empty. At each time $t = 1, 2, 3, \dots$ inspect the first coin of each growth site. If this coin is 1, include the site in $\mathcal{B}(t)$. If this coin is 0, leave the site out of $\mathcal{B}(t)$ but discard the used coin and move the next coin to the front. This procedure realizes the transition probability described in Proposition 1.1. The waiting time for site \mathbf{x} is $Y_{\mathbf{x}} = \inf\{k : \zeta_k^{\mathbf{x}} = 1\}$ which has distribution (1.6).

Totally asymmetric simple exclusion process. Next we relate the last-passage model to a discrete-time particle process which is one variant of the *totally asymmetric simple exclusion process* (TASEP). TASEP is a Markov process that describes the motion of particles on the integer lattice \mathbb{Z} . Particles are constrained by the exclusion rule which entails that two particles cannot occupy the same integer site at the same time. Label the particles with integers $i \in I$ where the index set I is a subinterval of \mathbb{Z} , possibly all of \mathbb{Z} . Let $X_i(t) \in \mathbb{Z}$ be the position of particle i at time $t \in \mathbb{Z}_+$. Each particle retains its label throughout the evolution.

One can imagine several ways of updating the particle locations. Let us consider the following discrete-time evolution where particles make random jumps to the right but subject to the exclusion rule. During time period $(t-1, t)$, each particle flips a p -coin to decide whether it attempts to jump one step to the right. The coin flips are all independent. Suppose the coin flip for particle i indicates a jump attempt. Then inspect the next site $X_i(t-1) + 1$. If this site was vacant at time $t-1$, particle i moves one step to the right. But if at time $t-1$ the next site to the right was occupied, particle i remains in its current location. Also if the coin indicated no jump attempt particle i remains in its current location.

Here are the rules summarized. Assume given an initial particle configuration $\{X_i(0)\}$. Then for each $t \in \mathbb{N}$:

(1.9) If site $X_i(t-1) + 1$ is occupied at time $t-1$ then $X_i(t) = X_i(t-1)$.

(1.10) If site $X_i(t-1) + 1$ is vacant at time $t-1$ then

$$X_i(t) = \begin{cases} X_i(t-1) + 1 & \text{with probability } p \\ X_i(t-1) & \text{with probability } q. \end{cases}$$

This defines a Markov process $\{X_i(t)\}_{i \in I}$ in discrete time $t \in \mathbb{Z}_+$ because the rules for the step from time $t-1$ to t do not look into the past before time $t-1$ and because the coin flips used in that step are independent of the past evolution.

To connect TASEP with the corner growth model, let the particles evolve from the following special initial configuration: all sites to the left of the origin are occupied, and all sites to the right of the origin, including the origin itself, are vacant. Thinking of a long line of vehicles stopped behind a red light we call this the “jam initial condition.” At time $t = 0$ the light turns green, and the vehicles start moving randomly to the right. Let us label the particle-vehicles from right to left with positive integers, so that initially $X_i(0) = -i$.

As a corollary of Proposition (1.1) we show that TASEP with jam initial condition is actually the last-passage model in disguise. Let the particles determine a growing cluster $\mathcal{A}(t)$ on \mathbb{N}^2 by the formula

$$(1.11) \quad \mathcal{A}(t) = \{(i, j) \in \mathbb{N}^2 : 1 \leq j \leq X_i(t) + i\}.$$

In other words, the height of the column above i in the cluster $\mathcal{A}(t)$ is $X_i(t) - X_i(0)$, the number of steps taken by particle i up to time t .

PROPOSITION 1.2. *Let $\mathcal{B}(t)$ be the cluster process of the last-passage model with geometric weights (1.6). Then the processes $\{\mathcal{A}(t) : t \in \mathbb{Z}_+\}$ and $\{\mathcal{B}(t) : t \in \mathbb{Z}_+\}$ are equal in distribution.*

PROOF. Initially $\mathcal{A}(0) = \emptyset = \mathcal{B}(0)$. In light of Proposition 1.1, we only need to observe that $\mathcal{A}(t)$ is a Markov chain with the transition probability described in Proposition 1.1. Point (i, j) is a growth site of $\mathcal{A}(t)$ iff $X_i(t) + i = j - 1$ and either $i = 1$ or $X_{i-1}(t) + i - 1 \geq j$. This is equivalent to saying that the site $X_i(t) + 1 = j - i$ is vacant at time t . Thus by rule (1.10) an independent p -coin flip determines whether X_i jumps to $j - i$ and thereby (i, j) joins $\mathcal{A}(t + 1)$.

Since the distribution of a discrete-time Markov chain on a countable state space is uniquely determined by the initial state and the transition probability, the claim of the Proposition follows. \square

Most often TASEP evolution is studied in continuous time. To run the process in continuous time give each particle a rate 1 Poisson process on the time line $(0, \infty)$ (a “Poisson clock”) and stipulate that whenever the Poisson clock of a particle jumps, this particle attempts to jump one step to the right. The jump is executed if the position on the right is vacant, otherwise not, so that the exclusion rule is not violated.

The waiting times between successive jump attempts of a particle have rate 1 exponential distribution. This is a continuous distribution with density e^{-t} on $(0, \infty)$. Consequently simultaneous jump attempts by different particles do not happen, or more precisely, happen with probability 0, and we do not need a rule for resolving conflicts that might arise from simultaneous jump attempts.

Proposition 1.2 can be proved again, this time for the last-passage model whose weights $\{Y_{i,j}\}$ are i.i.d. rate 1 exponential random variables.

A queueing model. The last-passage model arises also in situations that do not immediately seem to describe a growing cluster. Here is a queueing situation. Imagine that there are n customers that go through a system with m service stations. The customers are labeled $1, \dots, n$ and the servers are labeled $1, \dots, m$. Each customer visits the servers $1, \dots, m$ in order and then leaves the system. Each server serves one customer at a time, takes the customers in the order in which they arrive (first-in-first-out or FIFO discipline), and rests if there are no customers waiting to be served. Let $Y_{i,j}$ be the amount of service time that customer j requires with server i . Initially all n customers are queued up at server 1. At time $t = 0$ customer 1 begins service at server 1. At time $Y_{1,1}$ server 1 is finished with customer 1. At this point server 1 begins serving customer 2, while customer 1 moves to server 2. And so on. Let $\tau_{k,\ell}$ be the time when customer ℓ leaves server k . How is $\tau_{k,\ell}$ expressed in terms of $\{Y_{i,j}\}$?

A moment's thought reveals that $\tau_{k,\ell} = G(k, \ell)$. To see this, note that the rules of the queueing process imply that the service of customer ℓ with server k starts at time $\tau_{k,\ell-1} \vee \tau_{k-1,\ell}$, for this is the earliest time at which server k is done with customer $\ell-1$ and customer ℓ is through with server $k-1$. (To make this correct for $k=1$ or $\ell=1$ we add the boundary conditions $\tau_{i,0} = \tau_{0,j} = 0$ for all $i, j \geq 1$.) After the service customer ℓ departs server k and so $\tau_{k,\ell} = \tau_{k,\ell-1} \vee \tau_{k-1,\ell} + Y_{k,\ell}$. We have exactly the same equations that determine $G(m, n)$.

Comments

Among the seminal works to connect the last-passage model with queueing systems is Glynn and Whitt [GW91]. Rost [Ros81] connected the growth model with TASEP in one of the early papers on hydrodynamic limits of asymmetric particle systems, but he did not utilize the last-passage formulation.

Deterministic large scale limits

In this chapter we begin the study of the corner growth model with results that describe deterministic limits on large scales. We work with the last-passage times defined in the previous chapter:

$$(2.1) \quad G((k, \ell), (m, n)) = \max_{\pi \in \Pi((k, \ell), (m, n))} \sum_{(i, j) \in \pi} Y_{i, j},$$

for $(k, \ell), (m, n) \in \mathbb{N}^2$. $\Pi((k, \ell), (m, n))$ is the collection of nearest-neighbor up-right paths π from (k, ℓ) to (m, n) . For the special case where the paths start at $(1, 1)$ we write $G(m, n) = G((1, 1), (m, n))$. The first result in Section 2.1 is the existence of the deterministic limit

$$\lim_{N \rightarrow \infty} N^{-1}G(\lfloor Nx \rfloor, \lfloor Ny \rfloor) = \Psi(x, y) \quad \text{almost surely.}$$

In Section 2.2 we add boundary conditions to the last passage model to help us find an explicit formula for $\Psi(x, y)$ in the case of geometric weights. Then we turn to study a queueing model related to the last passage model with geometric weights, namely discrete-time M/M/1 queues in series. Section 2.3 establishes the invariant distributions for the queues. Section 2.4 proves a hydrodynamic limit for a bi-infinite system of such queues, a model that can also be called a discrete-time zero range particle system. The hydrodynamic limit gives another proof of the explicit limit earlier done in Section 2.2. (The second proof is not fundamentally different.) In the final Section 2.5 we develop a connection between the last-passage model with boundaries from Section 2.2 and the system of queues.

2.1. Law of large numbers

The general objective is to understand the behavior of the random variables $G(m, n)$ for large values of m and n . The most basic result is a law of large numbers that says that, along any direction $(x, y) \in \mathbb{R}_+^2$, the value $G(\lfloor Nx \rfloor, \lfloor Ny \rfloor)$ grows asymptotically at a precise, deterministic rate.

THEOREM 2.1. *Assume that the random variables $\{Y_{i, j} : (i, j) \in \mathbb{N}^2\}$ are independent and identically distributed and satisfy $0 \leq Y_{i, j} < \infty$. Then there exists a deterministic function $\Psi : (0, \infty)^2 \rightarrow [0, \infty]$ such that for all $(x, y) \in (0, \infty)^2$*

$$(2.2) \quad \Psi(x, y) = \lim_{N \rightarrow \infty} N^{-1}G(\lfloor Nx \rfloor, \lfloor Ny \rfloor) \quad \text{almost surely.}$$

Either $\Psi = \infty$ or $\Psi < \infty$ on all of $(0, \infty)^2$. In the latter case Ψ is superadditive, concave, continuous, homogeneous, and symmetric on $(0, \infty)^2$. Ψ is nondecreasing in both arguments, and more quantitatively $\Psi(x + h, y) \geq \Psi(x, y) + h\mathbf{E}(Y_{1, 1})$ for $h > 0$.

Here are precise expressions for several properties listed above, for $(x_1, y_1), (x_2, y_2) \in (0, \infty)^2$, $0 < s < 1$ and $c > 0$: superadditivity is

$$(2.3) \quad \Psi(x_1, y_1) + \Psi(x_2, y_2) \leq \Psi(x_1 + x_2, y_1 + y_2),$$

concavity is

$$(2.4) \quad s\Psi(x_1, y_1) + (1-s)\Psi(x_2, y_2) \leq \Psi(s(x_1, y_1) + (1-s)(x_2, y_2)),$$

homogeneity is

$$(2.5) \quad \Psi(cx_1, cy_1) = c\Psi(x_1, y_1)$$

and symmetry $\Psi(x_1, y_1) = \Psi(y_1, x_1)$.

PROOF OF THEOREM 2.1. The idea is to exploit the superadditivity

$$(2.6) \quad G(k, \ell) + G((k+1, \ell+1), (k+m, \ell+n)) \leq G(k+m, \ell+n)$$

that is a direct consequence of (2.1).

We start the proof of (2.2) with an integer point $(x, y) \in \mathbb{N}^2$. For $0 \leq m < n$ let $Z_{m,n} = G((mx+1, my+1), (nx, ny))$. $Z_{m,n}$ is a superadditive process that satisfies the assumptions of Corollary A.3 in Appendix A. Thus by that theorem there exists a function $\Psi : \mathbb{N}^2 \rightarrow [0, \infty]$ such that

$$(2.7) \quad \Psi(x, y) = \lim_{N \rightarrow \infty} N^{-1}G(Nx, Ny) \quad \text{almost surely, for all } (x, y) \in \mathbb{N}^2.$$

From the limit itself we get homogeneity (2.5) for $(x, y) \in \mathbb{N}^2$ and $c \in \mathbb{N}$. For superadditivity rewrite (2.6) as

$$(2.8) \quad \begin{aligned} N^{-1}G(Nx_1, Ny_1) + N^{-1}G((Nx_1+1, Ny_1+1), (Nx_1+Nx_2, Ny_1+Ny_2)) \\ \leq N^{-1}G(Nx_1+Nx_2, Ny_1+Ny_2). \end{aligned}$$

The first and last terms converge by (2.7). For the middle term we can assert convergence to $\Psi(x_2, y_2)$ along a subsequence because $G((Nx_1+1, Ny_1+1), (Nx_1+Nx_2, Ny_1+Ny_2))$ has the same distribution as $G(Nx_2, Ny_2)$. (This point is in Lemma A.1 in Appendix A.) Taking the limit in (2.8) along this subsequence leads to superadditivity (2.3) for $(x_1, y_1), (x_2, y_2) \in \mathbb{N}^2$. Lastly, it is immediate that $\Psi(x, y)$ is nondecreasing in both arguments separately.

Suppose $\Psi(a, b) = \infty$ for some $(a, b) \in \mathbb{N}^2$. Pick an arbitrary point $(x, y) \in \mathbb{N}^2$. Take $k \in \mathbb{N}$ large enough so that $kx > a$ and $ky > b$. Then

$$\begin{aligned} \Psi(x, y) &= k^{-1}\Psi(kx, ky) = \lim_{N \rightarrow \infty} (kN)^{-1}G(Nkx, Nky) \geq \lim_{N \rightarrow \infty} (kN)^{-1}G(Na, Nb) \\ &= k^{-1}\Psi(a, b) = \infty. \end{aligned}$$

Thus one infinite value forces $\Psi = \infty$ on all of \mathbb{N}^2 .

Let us finish off the case $\Psi = \infty$ by showing that the limit (2.2) holds with $\Psi(x, y) = \infty$ for any fixed $(x, y) \in (0, \infty)^2$. Pick $k \in \mathbb{N}$ so that $x, y > 1/k$. Write each integer N as $N = Mk + r$ for $M \in \mathbb{Z}_+$ and a remainder $r \in \{0, \dots, k-1\}$. Then $\lfloor Nx \rfloor \geq \lfloor N/k \rfloor = M$ and similarly for y , from which

$$\liminf_{N \rightarrow \infty} N^{-1}G(\lfloor Nx \rfloor, \lfloor Ny \rfloor) \geq \liminf_{N \rightarrow \infty} N^{-1}G(M, M) = k^{-1}\Psi(1, 1) = \infty.$$

This completes the proof for the case $\Psi = \infty$.

For the remainder of the proof we can assume $\Psi < \infty$ on \mathbb{N}^2 . The next step is to extend the limit (2.7) and the properties of Ψ to rational points. For any rational $(x, y) \in (0, \infty)^2$ define

$$(2.9) \quad \Psi(x, y) = k^{-1}\Psi(kx, ky)$$

for any positive integer k such that $(kx, ky) \in \mathbb{N}^2$. The definition is independent of the choice of k by the homogeneity (2.5) already established for integers. Homogeneity (2.5), superadditivity (2.3)

and the monotonicity extend to all rational arguments, including rational $c > 0$ in (2.5). Keep the choice of k and write again $N = Mk + r$ with $r \in \{0, \dots, k-1\}$. Then

$$Mkx \leq \lfloor Mkx + rx \rfloor = \lfloor Nx \rfloor \leq Mkx + rx < (M+1)kx,$$

the same inequalities hold for y , and the monotonicity of last-passage times gives

$$G(Mkx, Mky) \leq G(\lfloor Nx \rfloor, \lfloor Ny \rfloor) \leq G((M+1)kx, (M+1)ky).$$

Divide by N and let $N \rightarrow \infty$:

$$k^{-1}\Psi(kx, ky) \leq \varliminf_{N \rightarrow \infty} \frac{1}{N}G(\lfloor Nx \rfloor, \lfloor Ny \rfloor) \leq \overline{\varliminf}_{N \rightarrow \infty} \frac{1}{N}G(\lfloor Nx \rfloor, \lfloor Ny \rfloor) \leq k^{-1}\Psi(kx, ky).$$

We now have the limit (2.2) for rational $(x, y) \in (0, \infty)^2$.

The final extension of the limiting function is done as follows. For $(x, y) \in (0, \infty)^2$ set

$$(2.10) \quad \Psi(x, y) = \sup\{\Psi(u, v) : 0 < u \leq x, 0 < v \leq y, \text{ and } u, v \in \mathbb{Q}\}.$$

For rational (x, y) this is an identity that follows from the monotonicity of $\Psi(x, y)$. Once more extend the properties.

For homogeneity, take $(x, y) \in (0, \infty)^2$ but first consider rational $c > 0$. Below u, v range over rationals and then $u_1 = cu, v_1 = cv$.

$$c\Psi(x, y) = c \sup_{\substack{u \leq x \\ v \leq y}} \Psi(u, v) = \sup_{\substack{u \leq x \\ v \leq y}} c\Psi(u, v) = \sup_{\substack{u \leq x \\ v \leq y}} \Psi(cu, cv) = \sup_{\substack{u_1 \leq cx \\ v_1 \leq cy}} \Psi(u_1, v_1) = \Psi(cx, cy).$$

For general $c > 0$ find rational c_1, c_2 so that $c_1 < c < c_2$. By monotonicity

$$c_1\Psi(x, y) = \Psi(c_1x, c_1y) \leq \Psi(cx, cy) \leq \Psi(c_2x, c_2y) = c_2\Psi(x, y).$$

Letting $c_1 \nearrow c$ and $c_2 \searrow c$ completes the proof of homogeneity.

Superadditivity (2.3) and monotonicity are fairly immediate from the definition (2.10).

Superadditivity and homogeneity together imply concavity: for $0 < s < 1$ and $(x_1, y_1), (x_2, y_2) \in (0, \infty)^2$,

$$(2.11) \quad \begin{aligned} s\Psi(x_1, y_1) + (1-s)\Psi(x_2, y_2) &= \Psi(sx_1, sy_1) + \Psi((1-s)x_2, (1-s)y_2) \\ &\leq \Psi(sx_1 + (1-s)x_2, sy_1 + (1-s)y_2). \end{aligned}$$

A finite concave function on an open set is continuous [Roc70, Theorem 10.1]. Hence we get the continuity of Ψ .

For the last and most general form of the limit (2.2) for $(x, y) \in (0, \infty)^2$ pick rational points $(x_1, y_1), (x_2, y_2)$ such that $0 < x_1 < x < x_2$ and $0 < y_1 < y < y_2$. Then by monotonicity

$$G(\lfloor Nx_1 \rfloor, \lfloor Ny_1 \rfloor) \leq G(\lfloor Nx \rfloor, \lfloor Ny \rfloor) \leq G(\lfloor Nx_2 \rfloor, \lfloor Ny_2 \rfloor)$$

and by passing to the limit

$$\Psi(x_1, y_1) \leq \varliminf_{N \rightarrow \infty} \frac{1}{N}G(\lfloor Nx \rfloor, \lfloor Ny \rfloor) \leq \overline{\varliminf}_{N \rightarrow \infty} \frac{1}{N}G(\lfloor Nx \rfloor, \lfloor Ny \rfloor) \leq \Psi(x_2, y_2).$$

Let (x_1, y_1) and (x_2, y_2) tend to (x, y) and use the continuity of the function Ψ . The limit (2.2) has now been justified for all points.

The remaining points are the symmetry and the inequality $\Psi(x+h, y) \geq \Psi(x, y) + h\mathbf{E}(Y_{1,1})$ for $h > 0$. Symmetry follows from the distributional symmetry $G(m, n) \stackrel{d}{=} G(n, m)$ and an argument like the one used for the middle term of (2.8). The bound on the slope comes from $G(\lfloor Nx + Nh \rfloor, \lfloor Ny \rfloor) \geq G(\lfloor Nx \rfloor, \lfloor Ny \rfloor) + \sum_{i=\lfloor Nx \rfloor+1}^{\lfloor Nx + Nh \rfloor} Y_{i, \lfloor Ny \rfloor}$. \square

2.2. Explicit limit for geometric weights

We turn to the problem of computing the limit

$$(2.12) \quad \Psi(x, y) = \lim_{N \rightarrow \infty} \frac{G(\lfloor Nx \rfloor, \lfloor Ny \rfloor)}{N}$$

explicitly, and for this we need very specialized assumptions. Essentially only one distribution can be currently handled: the exponential, and its discrete counterpart, the geometric. Fix a parameter $0 < p < 1$ with $q = 1 - p$ and take the random variables $\{Y_{i,j}\}$ to be i.i.d. geometric random variables with common probability distribution γ as given here:

$$(2.13) \quad \mathbf{P}\{Y_{i,j} = k\} = \gamma(k) = pq^k, \quad k \in \mathbb{Z}_+.$$

This is the geometric distribution with parameter p . Here is the result.

THEOREM 2.2. *Under assumption (2.13) the limit in (2.12) is given for $(x, y) \in (0, \infty)^2$ by*

$$(2.14) \quad \Psi(x, y) = p^{-1}(qx + qy + 2\sqrt{qxy}).$$

The boundary curve of the limit shape is an arc of the ellipse...

The rest of this section proves the theorem. The difficulty with finding the explicit limit has to do with the superadditivity. The subadditive ergodic theorem (Theorem A.2 in Appendix A gives only an asymptotic expression for the limit. Assuming the moment bounds needed for the subadditive ergodic theorem (Theorem A.2), in our case this limit expression would be

$$\Psi(x, y) = \lim_{N \rightarrow \infty} \frac{\mathbf{E}[G(Nx, Ny)]}{N} \quad \text{for } (x, y) \in \mathbb{N}^2.$$

We need a new ingredient to find an explicit formula for $\Psi(x, y)$. We shall augment the model with suitable boundary conditions so that it becomes meaningful to talk about a notion of *steady state* or *invariant distribution*. Through these invariant distributions we can do explicit calculations.

The last-passage model with boundaries lives on the quadrant \mathbb{Z}_+^2 and there are two parameters $0 < r < p < 1$. The weights $\{Y_{i,j} : (i, j) \in \mathbb{Z}_+^2\}$ are independent and distributed as follows: for $k \in \mathbb{Z}_+$,

$$(2.15) \quad \mathbf{P}\{Y_{0,0} = 0\} = 1,$$

$$(2.16) \quad \mathbf{P}\{Y_{i,0} = k\} = \frac{p-r}{1-r} \left(\frac{q}{1-r} \right)^k \quad \text{for } i \in \mathbb{N},$$

$$(2.17) \quad \mathbf{P}\{Y_{0,j} = k\} = r(1-r)^k \quad \text{for } j \in \mathbb{N},$$

$$(2.18) \quad \mathbf{P}\{Y_{i,j} = k\} = pq^k \quad \text{for } i, j \in \mathbb{N}.$$

To clarify, in the interior of the model (that is, on \mathbb{N}^2) we have the same old i.i.d. geometric weights given in (2.13). The boundary values $\{Y_{i,j} : (i, j) \in \mathbb{Z}_+^2 \setminus \mathbb{N}^2\}$ are an auxiliary construction that will assist us in the task of computing the limit $\Psi(x, y)$ explicitly as a function of p .

Let $G_0(m, n) = G((0, 0), (m, n))$ denote last-passage times over paths that emanate from the origin and are allowed to collect weights on the boundaries:

$$(2.19) \quad G_0(m, n) = \max_{\pi \in \Pi((0,0), (m,n))} \sum_{(i,j) \in \pi} Y_{i,j}, \quad (m, n) \in \mathbb{Z}_+^2.$$

$\Pi((0, 0), (m, n))$ is the collection of nearest-neighbor up-right paths π from $(0, 0)$ to (m, n) in \mathbb{Z}_+^2 . To be precise, paths in $\Pi((0, 0), (m, n))$ include the origin $(0, 0)$ even though the weight $Y_{0,0} = 0$. We also continue to use the abbreviation $G(m, n) = G((1, 1), (m, n))$.

In the last-passage model with boundaries we define some new random variables. Horizontal and vertical increments are given by

$$(2.20) \quad \begin{aligned} I_{i,j} &= G_0(i,j) - G_0(i-1,j) \quad \text{for } i \geq 1, j \geq 0 \\ \text{and } J_{i,j} &= G_0(i,j) - G_0(i,j-1) \quad \text{for } i \geq 0, j \geq 1. \end{aligned}$$

An alternative formula for $I_{i,j}$ develops as follows, if $i, j \geq 1$:

$$\begin{aligned} I_{i,j} &= G_0(i,j) - G_0(i-1,j) \\ &= G_0(i-1,j) \vee G_0(i,j-1) + Y_{i,j} - G_0(i-1,j-1) \\ &\quad - [G_0(i-1,j) - G_0(i-1,j-1)] \\ &= J_{i-1,j} \vee I_{i,j-1} + Y_{i,j} - J_{i-1,j} \\ &= (I_{i,j-1} - J_{i-1,j})^+ + Y_{i,j}. \end{aligned}$$

Similar formula works for $J_{i,j}$ by symmetry, so we have

$$(2.21) \quad \begin{aligned} I_{i,j} &= (I_{i,j-1} - J_{i-1,j})^+ + Y_{i,j} \\ J_{i,j} &= (J_{i-1,j} - I_{i,j-1})^+ + Y_{i,j} \end{aligned} \quad \text{for } (i,j) \in \mathbb{N}^2.$$

Define

$$(2.22) \quad X_{i,j} = I_{i+1,j} \wedge J_{i,j+1} \quad \text{for } (i,j) \in \mathbb{Z}_+^2.$$

We develop the invariant distributions. First a technical lemma.

LEMMA 2.3. *Let $0 < r < p < 1$. Let I, J and Y be independent geometric random variables with distributions*

$$P[I = k] = \frac{p-r}{1-r} \left(\frac{q}{1-r} \right)^k, \quad P[J = k] = r(1-r)^k, \quad P[Y = k] = pq^k$$

for $k \in \mathbb{Z}_+$. Let $I_1 = (I - J)^+ + Y$, $J_1 = (J - I)^+ + Y$ and $X = I \wedge J$. Then the triple (I_1, J_1, X) has the same distribution as (I, J, Y) .

PROOF. Compute the joint Laplace transform with $u, v, w > 0$:

$$\begin{aligned} E[e^{-uI_1 - vJ_1 - wX}] &= E[e^{-u(I-J)^+ - v(J-I)^+ - w(I \wedge J)}] E[e^{-(u+v)Y}] \\ &= \left\{ \sum_{0 \leq j \leq i} \frac{p-r}{1-r} \left(\frac{q}{1-r} \right)^i r(1-r)^j e^{-u(i-j) - wj} \right. \\ &\quad \left. + \sum_{0 \leq i < j} \frac{p-r}{1-r} \left(\frac{q}{1-r} \right)^i r(1-r)^j e^{-v(j-i) - wi} \right\} \sum_{k=0}^{\infty} pq^k e^{-(u+v)k} \\ &= \frac{p-r}{1-r - qe^{-u}} \cdot \frac{r}{1 - (1-r)e^{-v}} \cdot \frac{p}{1 - qe^{-w}} \\ &= E[e^{-uI}] \cdot E[e^{-vJ}] \cdot E[e^{-wY}]. \quad \square \end{aligned}$$

Let Σ be the set of doubly-infinite down-right paths $\sigma = \{\sigma(\ell) : \ell \in \mathbb{Z}\}$ in \mathbb{Z}_+^2 . A down-right path means that the increments between the points $\sigma(\ell) = (\sigma_1(\ell), \sigma_2(\ell)) \in \mathbb{Z}_+^2$ satisfy

$$(\sigma_1(\ell), \sigma_2(\ell)) - (\sigma_1(\ell-1), \sigma_2(\ell-1)) = (1, 0) \quad \text{or} \quad (0, -1),$$

that is, direction \rightarrow or \downarrow . The interior of the set enclosed by σ is defined by

$$\mathcal{U}(\sigma) = \{(i, j) \in \mathbb{Z}_+^2 : \exists \ell \in \mathbb{Z} \text{ such that } 0 \leq i < \sigma_1(\ell) \text{ and } 0 \leq j < \sigma_2(\ell)\}.$$

We admit the possibility that σ is the union of the i - and j -coordinate axes, in which case $\mathcal{U}(\sigma)$ is empty. For $l \in \mathbb{Z}$ the last-passage time increments along σ are the variables

$$Z_\ell(\sigma) = \begin{cases} G_0(\sigma(\ell+1)) - G_0(\sigma(\ell)) = I_{\sigma(\ell+1)} & \text{if } \sigma(\ell+1) - \sigma(\ell) = (1, 0), \\ G_0(\sigma(\ell)) - G_0(\sigma(\ell+1)) = J_{\sigma(\ell)} & \text{if } \sigma(\ell+1) - \sigma(\ell) = (0, -1). \end{cases}$$

THEOREM 2.4. *For any $\sigma \in \Sigma$, the random variables*

$$(2.23) \quad \{\{X_{i,j} : (i,j) \in \mathcal{U}(\sigma)\}, \{Z_\ell(\sigma) : \ell \in \mathbb{Z}\}\}$$

are independent and geometrically distributed. $X_{i,j}$ has the interior distribution (2.18) with parameter p . If $Z_\ell(\sigma) = I_{\sigma(\ell+1)}$ it has the horizontal increment distribution (2.16) with parameter $(p-r)/(1-r)$, while if $Z_\ell(\sigma) = J_{\sigma(\ell)}$ then it has the vertical increment distribution (2.17) with parameter r .

PROOF. We first consider the countable set of paths that connect the j -axis to the i -axis, in other words those for which there exist finite $n_0 < n_1$ such that $\sigma_1(\ell) = 0$ for $\ell \leq n_0$ and $\sigma_2(\ell) = 0$ for $\ell \geq n_1$. For these paths we argue by induction on $\mathcal{U}(\sigma)$. When $\mathcal{U}(\sigma)$ is the empty set, the statement reduces to the independence of the $Y_{i,j}$ -values on the i - and j -axes which is true by construction.

Let $\sigma \in \Sigma$ be a path that connects the j -axis to the i -axis and assume the statement holds for σ . Let us say (i, j) is a *growth corner* for $\mathcal{U}(\sigma)$ if, for some index $\ell_0 \in \mathbb{Z}$,

$$(\sigma(\ell_0 - 1), \sigma(\ell_0), \sigma(\ell_0 + 1)) = ((i, j + 1), (i, j), (i + 1, j)).$$

A new valid $\tilde{\sigma} \in \Sigma$ can be defined by replacing $\sigma(\ell_0)$ with $\tilde{\sigma}(\ell_0) = (i + 1, j + 1)$ but keeping all other points intact: $\tilde{\sigma}(\ell) = \sigma(\ell)$ for $\ell \neq \ell_0$. The interior gained the point (i, j) : $\mathcal{U}(\tilde{\sigma}) = \mathcal{U}(\sigma) \cup \{(i, j)\}$.

In going from σ to $\tilde{\sigma}$ the change brought about in the set of random variables (2.23) is that

$$(2.24) \quad \{I_{i+1,j}, J_{i,j+1}\}$$

have been replaced by

$$(2.25) \quad \{I_{i+1,j+1}, J_{i+1,j+1}, X_{i,j}\}.$$

By (2.21) variables (2.25) are determined by variables (2.24) together with $Y_{i+1,j+1}$. By the last-passage construction $Y_{i+1,j+1}$ is independent of (2.23) for the σ under consideration. By the induction assumption the variables $\{I_{i+1,j}, J_{i,j+1}, Y_{i+1,j+1}\}$ are independent and independent from all the other variables in (2.23). An application of Lemma 2.3 to this last triple then implies that the new variables $\{I_{i+1,j+1}, J_{i+1,j+1}, X_{i,j}\}$ are also independent, have the correct marginal distributions, and are independent of all the other random variables of $\tilde{\sigma}$. Consequently $\tilde{\sigma}$ satisfies the statement of the theorem.

We can build all the paths σ that connect the axes by starting with the empty \mathcal{U} and adding growth corners one at a time. Hence this inductive argument proves the theorem for this class of paths.

For an arbitrary σ the statement follows because the independence of the random variables in (2.23) follows from independence of finite subcollections. Consider any square $R = \{0 \leq i, j \leq M\}$ large enough so that the corner (M, M) lies outside $\sigma \cup \mathcal{U}(\sigma)$. Then the X and $Z(\sigma)$ variables associated to σ that lie in R are a subset of the variables of the path $\tilde{\sigma}$ that goes through the points $(0, M)$, (M, M) and $(M, 0)$. This path $\tilde{\sigma}$ connects the axes so the first part of the proof applies to it. Thus the variables in (2.23) that lie inside an arbitrarily large square are independent. \square

In particular this theorem tells us that along any horizontal line, for a fixed $n \in \mathbb{Z}_+$, the increment process $\{I_{i,n} = G_0(i, n) - G_0(i - 1, n) : i \in \mathbb{N}\}$ is i.i.d. geometric with parameter $(p-r)/(1-r)$, exactly as the boundary values (2.16). Similarly, on each vertical line at fixed $m \in \mathbb{Z}_+$ we

see an i.i.d. process $\{J_{m,j} = G_0(m, j) - G_0(m, j-1) : j \in \mathbb{N}\}$ of geometric variables with parameter r . From

$$(2.26) \quad G_0(m, n) = \sum_{j=1}^n J_{0,j} + \sum_{i=1}^m I_{i,n}.$$

we can compute the expectation

$$(2.27) \quad \mathbf{E}G_0(m, n) = n\mathbf{E}J_{0,1} + m\mathbf{E}I_{1,n} = m\frac{q}{p-r} + n\frac{1-r}{r}.$$

The ease of this computation is in marked contrast with $G(m, n)$ whose expectation seems very difficult to find explicitly.

Note that it is not claimed (and not true) that *all* the variables on the right-hand side of (2.26) are independent. The path taken there, from $(0, 0)$ vertically up to $(0, n)$ and then horizontally right to (m, n) , is not a down-right path for which Theorem 2.4 applies. In fact, the two sums on the right-hand side of (2.26) are so strongly correlated with each other that the variance of $G_0(N, N)$ is of order $N^{2/3}$, instead of order N which it would be if the variables were all independent.

For the purpose of deriving limits we extract these further consequences.

COROLLARY 2.5. *For $(x, y) \in \mathbb{R}_+^2$ we have the limit*

$$(2.28) \quad \lim_{N \rightarrow \infty} \frac{G_0(\lfloor Nx \rfloor, \lfloor Ny \rfloor)}{N} = \Psi_0(x, y) \equiv x\frac{q}{p-r} + y\frac{1-r}{r} \quad \text{almost surely.}$$

The limits of (2.12) and (2.28) satisfy the inequality

$$(2.29) \quad \Psi(x, y) \leq \Psi_0(x, y) \quad \text{for } (x, y) \in (0, \infty)^2.$$

PROOF. We can construct the random variables $G(m, n)$ and $G_0(m, n)$ together from the weights $\{Y_{i,j} : (i, j) \in \mathbb{Z}_+^2\}$. (A joint construction is called a *coupling* of $G(m, n)$ and $G_0(m, n)$.) In this construction it is clear that

$$(2.30) \quad G(m, n) \leq G_0(m, n)$$

since the steps from $(0, 0)$ to $(1, 0)$ to $(1, 1)$ followed by a path from $(1, 1)$ to (m, n) that is maximal for $G(m, n)$ is just one possible path for $G_0(m, n)$.

For the limit use increments:

$$(2.31) \quad N^{-1}G_0(\lfloor Nx \rfloor, \lfloor Ny \rfloor) = N^{-1} \sum_{j=1}^{\lfloor Ny \rfloor} J_{0,j} + N^{-1} \sum_{i=1}^{\lfloor Nx \rfloor} I_{i, \lfloor Ny \rfloor}.$$

The strong law of large numbers gives

$$\lim_{N \rightarrow \infty} N^{-1} \sum_{j=1}^{\lfloor Ny \rfloor} J_{0,j} = y\mathbf{E}J_{0,1} = y\frac{1-r}{r} \quad \text{a.s.}$$

For the sum $N^{-1} \sum_{i=1}^{\lfloor Nx \rfloor} I_{i, \lfloor Ny \rfloor}$ we cannot use the strong law of large numbers because the summands also depend on N . Nevertheless, for each N the summands are i.i.d. geometric variables with a constant parameter. By standard large deviation estimates (Lemma A.4), for each $\varepsilon > 0$ there is a constant $c(\varepsilon) > 0$ such that

$$\mathbf{P}\left\{\left|N^{-1} \sum_{i=1}^{\lfloor Nx \rfloor} I_{i, \lfloor Ny \rfloor} - x\mathbf{E}I_{1,0}\right| \geq \varepsilon\right\} \leq e^{-c(\varepsilon)N}.$$

Thus by the Borel-Cantelli lemma also

$$\lim_{N \rightarrow \infty} N^{-1} \sum_{i=1}^{\lfloor Nx \rfloor} I_{i, \lfloor Ny \rfloor} = x \mathbf{E} I_{1,0} = x \frac{q}{p-r} \quad \text{a.s.}$$

Inequality (2.29) follows from the existence of the limits and (2.30). \square

We develop a precise connection between the last-passage times G and G_0 . Recall the definition (2.1) of the last-passage time between any two points on the integer square lattice. A maximal path for $G_0(m, n)$ first collects some weights on one of the two axes, and then follows a path in the interior \mathbb{N}^2 . We separate the two contributions. For $(m, n) \in \mathbb{N}^2$,

$$(2.32) \quad G_0(m, n) = \max_{k \in [m]} \{G_0(k, 0) + G((k, 1), (m, n))\} \\ \vee \max_{k \in [n]} \{G_0(0, k) + G((1, k), (m, n))\}.$$

To clarify, in the first maximum on the right-hand side $(k, 0)$ is the last point of the path on the boundary and $(k, 1)$ the point where the interior path begins. The notation $[m]$ is for the integer interval $\{1, \dots, m\}$.

The key is to take formula (2.32) to the limit and use the known values for $\Psi_0(x, y)$ to solve for $\Psi(x, y)$. By the homogeneity of $\Psi(x, y)$ it is sufficient to identify the one-variable function

$$(2.33) \quad \psi(x) = \Psi(x, 1), \quad x > 0.$$

LEMMA 2.6. *ψ is a strictly increasing, concave, continuous function on $(0, \infty)$. Setting $\psi(0) = q/p$ extends ψ to a strictly increasing, concave, continuous function on \mathbb{R}_+ .*

PROOF. The properties of ψ on $(0, \infty)$ follow from the definition (2.33) and the corresponding properties of Ψ given in Theorem 2.1. We need to extend $\psi(x)$ continuously to $x = 0$. From (2.28) and (2.29) comes the inequality

$$\Psi(x, y) \leq x \frac{q}{p-r} + y \frac{1-r}{r}.$$

To minimize the right-hand side set $r = p\sqrt{y}/(\sqrt{qx} + \sqrt{y})$ and get

$$(2.34) \quad \Psi(x, y) \leq p^{-1} [qx + qy + 2\sqrt{qxy}].$$

By taking any particular path π from $(1, 1)$ to $(\lfloor Nx \rfloor, \lfloor Ny \rfloor)$ we get the lower bound

$$G(\lfloor Nx \rfloor, \lfloor Ny \rfloor) \geq \sum_{(i,j) \in \pi} Y_{i,j}$$

from which by the strong law of large numbers

$$(2.35) \quad \Psi(x, y) \geq p^{-1} q(x + y).$$

In particular we have the bounds

$$(2.36) \quad p^{-1} q(x + 1) \leq \psi(x) \leq p^{-1} [qx + q + 2\sqrt{qx}].$$

Letting $x \searrow 0$ brings the bounds together and verifies the claim about extending ψ to \mathbb{R}_+ . \square

PROPOSITION 2.7. *We have the identity*

$$(2.37) \quad \Psi_0(1, 1) = \sup_{0 \leq z \leq 1} \{\Psi_0(z, 0) + \psi(1 - z)\} \vee \sup_{0 \leq z \leq 1} \{\Psi_0(0, z) + \psi(1 - z)\}.$$

PROOF. By (2.32) for any $1 \leq k \leq N$

$$(2.38) \quad G_0(N, N) \geq \{G_0(k, 0) + G((k, 1), (N, N))\} \\ \bigvee \{G_0(0, k) + G((1, k), (N, N))\}.$$

We wish to use the fact that the random variable $G((k, 1), (N, N))$ has the same probability distribution as $G((1, 1), (N - k + 1, N))$. The elegant way to do this is to imagine the so-called *canonical construction* for the random variables we are working with. The probability space $(\Omega, \mathcal{F}, \mathbf{P})$ has $\Omega = (\mathbb{Z}_+)^{\mathbb{Z}_+^2}$ with generic element $\omega = (\omega_{i,j})_{(i,j) \in \mathbb{Z}_+^2}$ and the product σ -algebra \mathcal{F} . \mathbf{P} is the product probability measure that makes the coordinate random variables $Y_{i,j}(\omega) = \omega_{i,j}$ independent with marginal distributions (2.15)–(2.18). On Ω we have the shift maps $(T_{k,\ell}(\omega))_{i,j} = \omega_{i+k,j+\ell}$ for $(k, \ell) \in \mathbb{Z}_+^2$ that leave the probability distribution of the interior variables invariant: if f is any Borel function on \mathbb{R}^M and $(i_1, j_1), \dots, (i_M, j_M) \in \mathbb{N}^2$, then for any $k, \ell \geq 0$,

$$\int f(\omega_{i_1, j_1}, \dots, \omega_{i_M, j_M}) \mathbf{P}(d\omega) = \int f(\omega_{i_1+k, j_1+\ell}, \dots, \omega_{i_M+k, j_M+\ell}) \mathbf{P}(d\omega)$$

provided the integrals are well-defined. By considering the indices of the Y -variables that go into formula (2.1) it should be evident that

$$(2.39) \quad G((k, 1), (N, N)) = G((1, 1), (N - k + 1, N)) \circ T_{k-1, 0}.$$

Then from the invariance comes the distributional equality:

$$(2.40) \quad \mathbf{P}\{G((k, 1), (N, N)) \in B\} = \mathbf{P}\{G((1, 1), (N - k + 1, N)) \in B\}$$

for all Borel sets $B \subseteq \mathbb{R}$.

Take $z \in [0, 1)$. Then for $N > 1/(1 - z)$ (to guarantee $\lceil Nz \rceil < N$) take $k = 1 + \lceil Nz \rceil$ so that $N - k + 1 = N - \lceil Nz \rceil = \lfloor N(1 - z) \rfloor$. From (2.38)

$$(2.41) \quad N^{-1}G_0(N, N) \geq \{N^{-1}G_0(1 + \lceil Nz \rceil, 0) + N^{-1}G(\lfloor N(1 - z) \rfloor, N) \circ T_{\lceil Nz \rceil, 0}\} \\ \bigvee \{N^{-1}G_0(0, 1 + \lceil Nz \rceil) + N^{-1}G(N, \lfloor N(1 - z) \rfloor) \circ T_{0, \lceil Nz \rceil}\}.$$

We let $N \rightarrow \infty$. We know from Theorem 2.1 that $N^{-1}G(\lfloor N(1 - z) \rfloor, N) \rightarrow \psi(1 - z)$ a.s. This does not imply the a.s. limit

$$(2.42) \quad N^{-1}G(\lfloor N(1 - z) \rfloor, N) \circ T_{\lceil Nz \rceil, 0} \rightarrow \psi(1 - z)$$

for the shifted variables. The reason is that the distributions of the processes

$$\{N^{-1}G(\lfloor N(1 - z) \rfloor, N) : N \geq 1\} \quad \text{and} \quad \{N^{-1}G(\lfloor N(1 - z) \rfloor, N) \circ T_{\lceil Nz \rceil, 0} : N \geq 1\}$$

are not identical. But the equal distributions (2.40) give convergence in probability in (2.42), and this in turn implies a.s. convergence along a subsequence. The same argument works for the transposed shifted term $N^{-1}G(N, \lfloor N(1 - z) \rfloor) \circ T_{0, \lceil Nz \rceil}$ on the second line of (2.41), and the limit is the same $\psi(1 - z)$ by the symmetry $\Psi(x, y) = \Psi(y, x)$.

The G_0 -terms in (2.41) converge a.s. by Corollary 2.5. Thus we take $N \rightarrow \infty$ in (2.41) along a suitable subsequence to conclude that for $z \in [0, 1)$,

$$\Psi_0(1, 1) \geq \{\Psi_0(z, 0) + \psi(1 - z)\} \bigvee \{\Psi_0(0, z) + \psi(1 - z)\}.$$

Continuity extends the conclusion to $z = 1$ and thereby we have verified that \geq holds in (2.37).

We turn to prove \leq in (2.37). Introduce another integer parameter L . Consider $N \geq L$ so that $\lceil NL^{-1}m \rceil < \lceil NL^{-1}(m+1) \rceil$ which justifies the first inequality below. From (2.32) develop:

$$\begin{aligned} G_0(N, N) &\leq \max_{0 \leq m \leq L-1} \{G_0(\lceil NL^{-1}(m+1) \rceil, 0) + G(\lceil NL^{-1}m \rceil + 1, 1), (N, N)\} \\ &\quad \bigvee \max_{0 \leq m \leq L-1} \{G_0(0, \lceil NL^{-1}(m+1) \rceil) + G(\lceil NL^{-1}m \rceil + 1, 1), (N, N)\} \\ &= \max_{0 \leq m \leq L-1} \{G_0(\lceil NL^{-1}(m+1) \rceil, 0) + G(\lfloor N(1 - L^{-1}m) \rfloor, N) \circ T_{\lceil NL^{-1}m \rceil, 0}\} \\ &\quad \bigvee \max_{0 \leq m \leq L-1} \{G_0(0, \lceil NL^{-1}(m+1) \rceil) + G(N, \lfloor N(1 - L^{-1}m) \rfloor) \circ T_{0, \lceil NL^{-1}m \rceil}\} \end{aligned}$$

Letting $N \rightarrow \infty$ along a suitable subsequence gives

$$\begin{aligned} \Psi_0(1, 1) &\leq \max_{0 \leq m \leq L-1} \{\Psi_0(L^{-1}(m+1), 0) + \psi(1 - L^{-1}m)\} \\ &\quad \bigvee \max_{0 \leq m \leq L-1} \{\Psi_0(0, L^{-1}(m+1)) + \psi(1 - L^{-1}m)\} \end{aligned}$$

by the bilinearity of $\Psi_0(x, y)$ in x and y

$$\leq \sup_{0 \leq z \leq 1} \{\Psi_0(z, 0) + \psi(1 - z)\} \bigvee \sup_{0 \leq z \leq 1} \{\Psi_0(0, z) + \psi(1 - z)\} + CL^{-1}.$$

Letting $L \rightarrow \infty$ completes the proof. \square

The probabilistic part of the proof is over. The rest is analysis to extract the unknown ψ from (2.37).

COMPLETION OF THE PROOF OF THEOREM 2.14. We arrange things so that concave duality from Appendix C.1 applies. The first task is to simplify (2.37). Rewrite it with explicit formulas from (2.28):

$$(2.43) \quad \frac{q}{p-r} + \frac{1-r}{r} = \sup_{0 \leq z \leq 1} \left\{ z \frac{q}{p-r} + \psi(1-z) \right\} \bigvee \sup_{0 \leq z \leq 1} \left\{ z \frac{1-r}{r} + \psi(1-z) \right\}.$$

Observe that $(1-r)/r \geq q/(p-r)$ iff $r \leq 1 - \sqrt{q}$. We restrict the parameter r to the subset $(0, 1 - \sqrt{q}]$ of its original range $(0, p)$. Then we can drop the first expression in braces from the right-hand side of (2.43) because at each z -value the second expression in braces dominates. Replace the variable z with $x = 1 - z$, multiply through by -1 , and we have turned (2.43) into

$$-\frac{q}{p-r} = \inf_{0 \leq x \leq 1} \left\{ x \frac{1-r}{r} - \psi(x) \right\}, \quad r \in (0, 1 - \sqrt{q}].$$

Once more change variables to $y = (1-r)/r$ to turn the above equation into

$$(2.44) \quad -\frac{q(1+y)}{py-q} = \inf_{0 \leq x \leq 1} \{xy - \psi(x)\}, \quad y \in [p^{-1}(\sqrt{q} + q), \infty).$$

Define the function

$$\tilde{\psi}(x) = \begin{cases} -\infty, & x < 0 \\ \psi(x), & 0 \leq x \leq 1 \\ \psi(1), & x > 1. \end{cases}$$

By the monotonicity, continuity and concavity of ψ on $[0, 1]$, $\tilde{\psi} : \mathbb{R} \rightarrow [-\infty, \infty)$ is upper semicontinuous and concave. Equation (2.44) implies that

$$-\frac{q(1+y)}{py-q} = \inf_{x \in \mathbb{R}} \{xy - \tilde{\psi}(x)\}, \quad y \in [p^{-1}(\sqrt{q} + q), \infty).$$

In particular, the values $x > 1$ cannot yield the infimum above since $y > 0$ and $\tilde{\psi}$ is constant on $[1, \infty)$. In terms of concave duality, the above equation says that

$$(2.45) \quad \tilde{\psi}^*(y) = -\frac{q(1+y)}{py-q} \quad \text{for } y \in [p^{-1}(\sqrt{q}+q), \infty).$$

By Corollary C.2

$$(2.46) \quad \tilde{\psi}(x) = \tilde{\psi}^{**}(x) = \inf_{y \in \mathbb{R}} \{xy - \tilde{\psi}^*(y)\}.$$

Restrict the above to $x \in [0, 1]$ so that $\psi(x) = \tilde{\psi}(x)$. From (2.45) we can get a right derivative at $y = p^{-1}(\sqrt{q}+q)$:

$$(\tilde{\psi}^*)' \left(\left(\frac{\sqrt{q}+q}{p} \right)^+ \right) = \frac{q}{(py-q)^2} \Big|_{y=\frac{\sqrt{q}+q}{p}} = 1.$$

Hence by concavity, for $y < p^{-1}(\sqrt{q}+q)$ and $x \in [0, 1]$,

$$\tilde{\psi}^*(y) \leq \tilde{\psi}^* \left(\frac{\sqrt{q}+q}{p} \right) + y - \frac{\sqrt{q}+q}{p} \leq \tilde{\psi}^* \left(\frac{\sqrt{q}+q}{p} \right) + xy - x \frac{\sqrt{q}+q}{p}.$$

The conclusion is that for $x \in [0, 1]$ the infimum in (2.46) is not affected by restricting y to $[p^{-1}(\sqrt{q}+q), \infty)$. This is the range where we know $\tilde{\psi}^*(y)$ from (2.45). Consequently for $x \in [0, 1]$

$$(2.47) \quad \psi(x) = \inf_{y \geq p^{-1}(\sqrt{q}+q)} \left\{ xy + \frac{q(1+y)}{py-q} \right\} = p^{-1}(qx + q + 2\sqrt{qx}).$$

The second step above is calculus.

Going back to the definition (2.33) of ψ and by the homogeneity of Ψ , for $0 < x \leq y$,

$$\Psi(x, y) = y\Psi(x/y, 1) = y\psi(x/y) = p^{-1}(qx + qy + 2\sqrt{qxy}).$$

By symmetry of Ψ the same formula works for all $(x, y) \in (0, \infty)^2$. This completes the proof of formula (2.14). \square

2.3. M/M/1 queues in series

We turn to study discrete-time M/M/1 queues in series, a model which can also be called a discrete-time zero range process. This particle system is closely related to the last-passage model with geometric weights studied in Section 2.2. However, results from Section 2.2 will not be used. Instead, with the help of the queues, we give an alternate proof of Theorem 2.2 at the end of Section 2.4. The underlying idea of the proof is the same as for the one we just covered: explicit limits come from knowing explicit invariant distributions.

In this section we study the stationary behavior first of a single queue and then queues in series. In Section 2.4 we prove a hydrodynamic limit for the infinite process. In that proof a particle version of the last-passage model appears in a variational property of the process (see equation (2.72)) that gives a counterpart to equation (2.32).

As in the previous section, we have a fixed parameter $p = 1 - q \in (0, 1)$, and an auxiliary parameter r that varies in the interval $r \in (0, p)$. A single-server discrete-time M/M/1 queue operates as follows. Customers arrive one at a time at a service station, are served in the order in which they arrive, and leave after the service is complete. This scheme is called FIFO, short for first-in-first-out queueing discipline. The rule for the stochastic evolution is that for every time point $t \in \mathbb{N}$, during time period $(t-1, t)$ two things can happen. (i) If the queue is not empty at time $t-1$, then with probability p one customer departs by time t . (ii) One new customer arrives with probability r . Our convention is that it is **not** possible for a customer to arrive and depart during the same interval $(t-1, t)$. The name M/M/1 for this model comes from memoryless arrivals and services (that's two M's) and from having 1 server.

For $t \in \mathbb{N}$ let

$$a(t) = \mathbf{1}\{\text{a customer arrives during } (t-1, t)\}$$

and

$$d(t) = \mathbf{1}\{\text{a customer departs during } (t-1, t)\},$$

and for $t \in \mathbb{Z}_+$ let $Q(t)$ denote the number of customers in the system (queue length) at time t . The arrival process is $\{a(t) : t \in \mathbb{N}\}$, departure process is $\{d(t) : t \in \mathbb{N}\}$, and the queue length process is $\{Q(t) : t \in \mathbb{Z}_+\}$. We call the arrival process a mean r Bernoulli process because the variables $a(t)$ are independent mean r Bernoulli variables.

For a rigorous construction of the processes (a, Q, d) we need one more ingredient, namely the probability distribution β for the initial queue length $Q(0)$. Once β is picked, let $(\Omega, \mathcal{F}, \mathbf{P}_\beta)$ be a probability space on which are defined these three independent random objects: (i) β -distributed $Q(0)$, (ii) mean r Bernoulli process $\{a(t) : t \in \mathbb{N}\}$, and (iii) mean p Bernoulli process $\{\kappa(t) : t \in \mathbb{N}\}$. The variables $\kappa(t)$ signal possible service completions. On this probability space we define the queue length and departure processes for $t \in \mathbb{N}$ by

$$(2.48) \quad \begin{aligned} Q(t) &= (Q(t-1) - \kappa(t))_+ + a(t) \\ \text{and } d(t) &= \kappa(t) \cdot \mathbf{1}\{Q(t-1) \geq 1\}. \end{aligned}$$

Since $a(t)$ and $\kappa(t)$ are inputs that are independent of $Q(0), \dots, Q(t-1)$, the formula above defines $Q(t)$ as a discrete-time Markov chain with state space \mathbb{Z}_+ and transition matrix $\{P(x, y) : x, y \in \mathbb{Z}_+\}$ whose nonzero values are given by

$$(2.49) \quad \begin{aligned} P(x, x+1) &= \begin{cases} r, & x = 0 \\ qr, & x \geq 1 \end{cases} \\ P(x, x) &= \begin{cases} 1-r, & x = 0 \\ q(1-r) + pr, & x \geq 1 \end{cases} \\ P(x, x-1) &= p(1-r), \quad x \geq 1. \end{aligned}$$

We observe the state $Q(t)$ of the queue at integer times $t = 0, 1, 2, \dots$. Between successive integer times at most one arrival and at most one departure can happen. Consequently the transition probabilities satisfy $P(x, y) = 0$ for $y \notin \{x, x \pm 1\}$.

Set

$$(2.50) \quad u = \frac{qr}{p(1-r)} \in (0, 1)$$

and define this probability distribution on \mathbb{Z}_+ :

$$(2.51) \quad \alpha(0) = \frac{p-r}{p} \quad \text{and} \quad \alpha(x) = \frac{p-r}{pq} u^x \quad \text{for } x \geq 1.$$

LEMMA 2.8. *Let $0 < r < p$ and let the arrival process be a mean r Bernoulli process. The probability measure α is reversible for the transition P . If the distribution of the initial queue length $Q(0)$ is α , then under \mathbf{P}_α the process $Q(\cdot)$ is a stationary Markov chain, the departure process is a mean r Bernoulli process, and the departure process up to time t is independent of the queue length at time t , for all $t \in \mathbb{N}$. Precisely speaking, for all $t \in \mathbb{N}$, $k \in \mathbb{Z}_+$, and $(w_1, \dots, w_t) \in \{0, 1\}^t$:*

$$(2.52) \quad \mathbf{P}_\alpha\{Q(t) = k, (d(1), \dots, d(t)) = (w_1, \dots, w_t)\} = \alpha(k) \cdot r^{\sum_{i=1}^t w_i} (1-r)^{t - \sum_{i=1}^t w_i}.$$

PROOF. Reversibility amounts to checking $\alpha(x)P(x, y) = \alpha(y)P(y, x)$ for all pairs $(x, y) \in \mathbb{Z}_+$. Reversibility implies invariance (that is, $\alpha P = \alpha$) and this in turn that under \mathbf{P}_α the Markov chain $Q(\cdot)$ is stationary.

Equation (2.52) is checked by induction on t . We give parts of the calculation explicitly.

Case $t = 1$. Assume $k \geq 1$, $w_1 = 1$.

$$\begin{aligned} \mathbf{P}_\alpha\{Q(1) = k, d(1) = 1\} &= \mathbf{P}_\alpha\{Q(0) = k, \kappa(1) = 1, a(1) = 1\} \\ &\quad + \mathbf{P}_\alpha\{Q(0) = k + 1, \kappa(1) = 1, a(1) = 0\} \\ &= \alpha(k)pr + \alpha(k + 1)p(1 - r) = \alpha(k)r. \end{aligned}$$

Then $k = 0$, $w_1 = 1$.

$$\begin{aligned} \mathbf{P}_\alpha\{Q(1) = 0, d(1) = 1\} &= \mathbf{P}_\alpha\{Q(0) = 1, \kappa(1) = 1, a(1) = 0\} \\ &= \alpha(1)p(1 - r) = \alpha(0)r. \end{aligned}$$

The cases with $w_1 = 0$ come by complementation.

Induction step. Assume (2.52) is valid up to time $t - 1$. Abbreviate

$$D_{w_1, t} = \{(d(1), \dots, d(t)) = (w_1, \dots, w_t)\}.$$

Assume $k \geq 1$, $w_t = 1$.

$$\begin{aligned} \mathbf{P}_\alpha\{Q(t) = k, D_{w_1, t}\} &= \mathbf{P}_\alpha\{Q(t-1) = k, D_{w_1, t-1}, \kappa(t) = 1, a(t) = 1\} \\ &\quad + \mathbf{P}_\alpha\{Q(t-1) = k + 1, D_{w_1, t-1}, \kappa(t) = 1, a(t) = 0\} \end{aligned}$$

utilizing the independence of $(\kappa(t), a(t))$ from everything that has happened up to time $t - 1$

$$= \mathbf{P}_\alpha\{Q(t-1) = k, D_{w_1, t-1}\}pr + \mathbf{P}_\alpha\{Q(t-1) = k + 1, D_{w_1, t-1}\}p(1 - r)$$

by induction

$$\begin{aligned} &= \alpha(k)r^{\sum_{i=1}^{t-1} w_i}(1 - r)^{t-1-\sum_{i=1}^{t-1} w_i}pr + \alpha(k + 1)r^{\sum_{i=1}^{t-1} w_i}(1 - r)^{t-1-\sum_{i=1}^{t-1} w_i}p(1 - r) \\ &= \alpha(k)r^{\sum_{i=1}^t w_i}(1 - r)^{t-\sum_{i=1}^t w_i}. \end{aligned}$$

Similarly for the case $k = 0$, $w_t = 1$ and then the $w_t = 0$ cases by complementation. \square

We build on this lemma to describe invariant distributions for a process that consists of a bi-infinite sequence of queues. Customers departing queue i immediately join the arrival process of queue $i + 1$. The state of the process is $\eta(t) = (\eta_i(t) : i \in \mathbb{Z})$, where $\eta_i(\cdot)$ is the queue length process of queue i . The state space of the process $\eta(\cdot)$ is $\mathbf{X} = (\mathbb{Z}_+)^{\mathbb{Z}}$. Generic elements of \mathbf{X} are denoted by $\eta = (\eta_i)_{i \in \mathbb{Z}}$.

To define the evolution, assume given an initial state $\eta(0)$ and let $\{\kappa_i(t) : i \in \mathbb{Z}, t \in \mathbb{N}\}$ be a collection of i.i.d. $\{0, 1\}$ -valued Bernoulli random variables with mean p . Variable $\kappa_i(t)$ signals a possible service completion at server i during time period $(t - 1, t)$. The initial state $\eta(0)$ is independent of the process $\{\kappa_i(t)\}$. For $t \in \mathbb{N}$, the move from $\eta(t - 1)$ to $\eta(t)$ follows the equation

$$(2.53) \quad \eta_i(t) = (\eta_i(t - 1) - \kappa_i(t))_+ + \kappa_{i-1}(t) \cdot \mathbf{1}\{\eta_{i-1}(t - 1) \geq 1\}, \quad i \in \mathbb{Z}.$$

The equation expresses the two events that can happen at queue i . The first term on the right represents a possible departure at queue i , and the second term a new arrival to queue i from queue $i - 1$. The evolution $\eta(t)$ is well-defined for any initial state $\eta(0) \in \mathbf{X}$ and it is a Markov process. Let $a_i(t)$ be the arrival process to queue i and $d_i(t)$ the departure process from queue i . Since the departures from queue $i - 1$ feed directly into queue i ,

$$a_i(t) = d_{i-1}(t) = \kappa_{i-1}(t) \cdot \mathbf{1}\{\eta_{i-1}(t - 1) \geq 1\}.$$

Let us use the notation $a_i[s, t] = \{a_i(m) : s \leq m \leq t\}$ for segments of these processes.

In the particle system literature, process $\eta(\cdot)$ would be called a discrete-time *zero range process*. The name comes from the property that a particle interacts only with particles on its site, and not with particles at other sites (zero range of interaction). Our process is a simple case of the zero range process. More generally, the dynamics is determined by a function $g(\eta_i)$ that gives, as a function of the number of particles at site i , the probability or the rate of moving one particle out of site i .

Let $\nu = \alpha^{\otimes \mathbb{Z}}$ be the probability measure on \mathbf{X} under which the coordinates η_i are i.i.d. α -distributed. Explicitly, for any n distinct indices $i_1, \dots, i_n \in \mathbb{Z}$ and any $x_1, \dots, x_n \in \mathbb{Z}_+$,

$$\nu\{\eta \in \mathbf{X} : \eta_{i_1} = x_1, \dots, \eta_{i_n} = x_n\} = \prod_{j=1}^n \alpha(x_j).$$

Note that there is actually a family of measures indexed by $r \in (0, p)$ (considering p fixed).

LEMMA 2.9. *Consider $p \in (0, 1)$ fixed. Then for each $r \in (0, p)$ the probability measure ν is invariant for the process $\eta(\cdot)$. Let the initial state have distribution ν . Then at each time t and for each i the departure process $d_i[1, t]$ is a mean r Bernoulli process and independent of the queue lengths $\{\eta_j(t) : j \leq i\}$.*

PROOF. Fix $t \in \mathbb{N}$. Assuming that $\eta(0)$ has distribution ν we show that $\eta(t)$ also has distribution ν .

Let $Y_i = (\eta_i(t), (d_i[1, t]))$. Y_i is a Markov chain indexed by i and with countable state space $\mathbb{Z}_+ \times \{0, 1\}^t$. The Markov property is true because the only information from $\{Y_j : j < i\}$ used to compute Y_i is the arrival process $a_i[1, t] = d_{i-1}[1, t]$. The rest come from independent inputs. The transition probability from Y_{i-1} to Y_i can be described informally by saying: take the arrival process $a_i[1, t] = d_{i-1}[1, t]$, pick independently an α -distributed initial queue length $\eta_i(0)$ and the Bernoulli variables $\kappa_i[1, t]$, and compute the variables $Y_i = (\eta_i(t), (d_i[1, t]))$ according to equations (2.48).

Lemma 2.8 implies that the probability measure described by (2.52) is an invariant distribution for the Markov chain Y_i . Markov chain Y_i is irreducible, for the state $(t, (0, 0, \dots, 0))$ communicates with every other state. To move from state $(x, w_{1,t})$ to $(t, (0, \dots, 0))$, pick the next initial queue length $Q(0) = t - \sum_{i=1}^t w_i$ and $\kappa[1, t] = (0, \dots, 0)$. To move from $(t, (0, \dots, 0))$ to $(x, w_{1,t})$, pick $Q(0) = x + \sum_{i=1}^t w_i$ and $\kappa[1, t] = w_{1,t}$. Finally, Y_i is aperiodic because state $(t, (0, \dots, 0))$ has period 1.

Thus the convergence theorem for Markov chains applies [Dur04, Ch. 5, Theorem (5.5)] and implies that $\{Y_i\}$ is the stationary Markov chain with each Y_i distributed as in (2.52). To make this assertion rigorously for a fixed segment $(Y_{i_0}, Y_{i_0+1}, \dots, Y_{i_1})$, imagine the Markov chain Y_i starting at index value $i = N$ and take N to $-\infty$.

Fix $i_0 \in \mathbb{Z}$. We appeal once more to Lemma 2.8 to prove the following claim: for $i \geq i_0$, the variables $\{\eta_{i_0}(t), \eta_{i_0+1}(t), \dots, \eta_i(t), d_i[1, t]\}$ are mutually independent, with each $\eta_j(t)$ α -distributed and $d_i[1, t]$ a mean r Bernoulli process. This claim finishes the lemma.

The case $i = i_0$ is the above conclusion that Y_{i_0} has the distribution in (2.52).

Assume the claim is true for some $i \geq i_0$. Then by the independence built into the inputs, also the larger collection of random variables $\{\eta_{i_0}(t), \eta_{i_0+1}(t), \dots, \eta_i(t), d_i[1, t], \eta_{i+1}(0), \kappa_{i+1}[1, t]\}$ is mutually independent. Let I and J be the functions defined by (2.48) that give the queue length and the departure process as functions of the inputs:

$$\eta_{i+1}(t) = I(a_{i+1}[1, t], \eta_{i+1}(0), \kappa_{i+1}[1, t]) \quad \text{and} \quad d_{i+1}[1, t] = J(a_{i+1}[1, t], \eta_{i+1}(0), \kappa_{i+1}[1, t]).$$

It seems clearest to make the argument through an expectation. Let f_1, f_2, f_3 be bounded functions defined on the appropriate spaces so that the expectations below make sense.

$$\begin{aligned} & \mathbf{E}_\nu[f_1(\eta_{i_0}(t), \dots, \eta_i(t))f_2(\eta_{i+1}(t))f_3(d_{i+1}[1, t])] \\ &= \mathbf{E}_\nu[f_1(\eta_{i_0}(t), \dots, \eta_i(t))f_2(I(d_i[1, t], \eta_{i+1}(0), \kappa_{i+1}[1, t])) \\ & \quad \times f_3(J(d_i[1, t], \eta_{i+1}(0), \kappa_{i+1}[1, t]))] \end{aligned}$$

by the induction hypothesis

$$\begin{aligned} &= \mathbf{E}_\nu[f_1(\eta_{i_0}(t), \dots, \eta_i(t))] \\ & \quad \times \mathbf{E}_\nu[f_2(I(d_i[1, t], \eta_{i+1}(0), \kappa_{i+1}[1, t]))f_3(J(d_i[1, t], \eta_{i+1}(0), \kappa_{i+1}[1, t]))] \\ &= \mathbf{E}_\nu[f_1(\eta_{i_0}(t), \dots, \eta_i(t))]\mathbf{E}_\nu[f_2(\eta_{i+1}(t))f_3(d_{i+1}[1, t])] \end{aligned}$$

by another application of Lemma 2.8 and the part of the induction assumption that says the arrival process $a_{i+1}[1, t]$ is mean r Bernoulli

$$= \mathbf{E}_\nu[f_1(\eta_{i_0}(t), \dots, \eta_i(t))]\mathbf{E}_\nu[f_2(\eta_{i+1}(t))]\mathbf{E}_\nu[f_3(d_{i+1}[1, t])].$$

Since we already know that $\eta_{i+1}(t)$ and $d_{i+1}[1, t]$ have the desired marginal distributions, this extends the claim to $i + 1$ and thereby completes the proof of the lemma. \square

2.4. Hydrodynamic limit for M/M/1 queues in series

A *hydrodynamic limit* is a law of large numbers that describes deterministic behavior of an interacting particle system on large space and time scales. We consider a sequence of processes $\eta^N(\cdot)$ indexed by $N \in \mathbb{N}$. Each process is of the type studied in the previous section, with the state $\eta^N(t) = (\eta_i^N(t) : i \in \mathbb{Z}) \in \mathbf{X}$ that consists of a sequence of queue lengths and evolution described by equation (2.53). Let $d_i^N(\cdot)$ be the departure process from queue η_i^N . We write \mathbf{P}^N for the probability measure on the probability space on which processes $(\eta^N(\cdot), d^N(\cdot))$ are defined.

The only hypothesis needed is that the initial states $\eta^N(0)$ approximate a deterministic profile.

There exists a nondecreasing function $U_0 : \mathbb{R} \rightarrow \mathbb{R}$ such that $U_0(0) = 0$ and, for all $a < b$ in \mathbb{R} and $\varepsilon > 0$,

$$(2.54) \quad \lim_{N \rightarrow \infty} \mathbf{P}^N \left\{ \left| N^{-1} \sum_{i=\lfloor Na \rfloor + 1}^{\lfloor Nb \rfloor} \eta_i^N(0) - (U_0(b) - U_0(a)) \right| \geq \varepsilon \right\} = 0.$$

EXAMPLE 2.10. Here is the simplest example of initial variables that satisfy assumption (2.54). Suppose U'_0 is continuous. For each N let the variables $\{\eta_i^N(0) : i \in \mathbb{Z}\}$ be independent with means $\mathbf{E}^N[\eta_i^N(0)] = U'_0(i/N)$. Assume there is a finite constant C such that $\mathbf{E}^N[\eta_i^N(0)^2] \leq C$ for all N and i .

Define a function g on \mathbb{R}_+ by

$$(2.55) \quad g(x) = \begin{cases} -(\sqrt{p(1-x)} - \sqrt{qx})^2, & 0 \leq x \leq p \\ 0, & x > p. \end{cases}$$

Then define a function U on $\mathbb{R}_+ \times \mathbb{R}$ by setting $U(0, x) = U_0(x)$ and

$$(2.56) \quad U(t, x) = \sup_{y: y \leq x} \left\{ U_0(y) + tg\left(\frac{x-y}{t}\right) \right\}, \quad t > 0, x \in \mathbb{R}.$$

The resulting function $U(t, x)$ continues to be nondecreasing in x . This operation defines a semi-group, in the sense that once U is defined as above, then for all $0 \leq s < t$,

$$(2.57) \quad U(t, x) = \sup_{y: y \leq x} \left\{ U(s, y) + (t-s)g\left(\frac{x-y}{t-s}\right) \right\}.$$

Since $g \leq 0$ and $U(s, y)$ is nondecreasing in y , $U(t, x)$ is nonincreasing in t .

Function U describes the process on space and time scales of order N as made precise in the next theorem.

THEOREM 2.11. *Assume (2.54). Then for all $a < b$ in \mathbb{R} , $t > 0$ and $\varepsilon > 0$,*

$$(2.58) \quad \lim_{N \rightarrow \infty} \mathbf{P}^N \left\{ \left| N^{-1} \sum_{i=\lfloor Na \rfloor + 1}^{\lfloor Nb \rfloor} \eta_i^N(\lfloor Nt \rfloor) - (U(t, b) - U(t, a)) \right| \geq \varepsilon \right\} = 0$$

and

$$(2.59) \quad \lim_{N \rightarrow \infty} \mathbf{P}^N \left\{ \left| N^{-1} \sum_{s=1}^{\lfloor Nt \rfloor} d_{\lfloor Na \rfloor}^N(s) - (U(0, a) - U(t, a)) \right| \geq \varepsilon \right\} = 0.$$

Along the way to the proof we will also give another proof of Theorem 2.2. This appears as Corollary 2.16 in the final stage of the proof.

The theorem says that the function g of (2.55) summarizes all the information needed about the process for describing its large scale evolution. In the proof we define this function in terms of the limit of a system that starts with infinitely many customers in a single queue and the rest of the system empty. It turns out that g is the concave dual of the *particle flux* f defined as follows:

$$f(\rho) = r \text{ with } r \text{ uniquely determined by } \sum_{x=0}^{\infty} x \alpha(x) = \rho.$$

Here α is the stationary distribution of a single queue defined in (2.51). The flux $f(\rho)$ gives the average departure rate as a function of the mean queue length, or particle density, ρ . (This meaning of the parameter r was clear from Lemma 2.8.) The explicit formula is

$$(2.60) \quad f(\rho) = \frac{1}{2} (1 + \rho - \sqrt{(1 + \rho)^2 - 4p\rho}), \quad \rho \in [0, \infty).$$

The connection between g and f is, for $\rho, x \in \mathbb{R}_+$,

$$(2.61) \quad f(\rho) = \inf_{x \in \mathbb{R}_+} \{\rho x - g(x)\} \quad \text{and} \quad g(x) = \inf_{\rho \in \mathbb{R}_+} \{\rho x - f(\rho)\}.$$

This will arise in the final stage of the proof.

What formula (2.56) tells us about the particle system becomes clearer through its connection with partial differential equations. For this discussion let us assume that

$$U_0 \text{ is Lipschitz continuous.}$$

Then definition (2.56) can be used to show that U is a Lipschitz function on $\mathbb{R}_+ \times \mathbb{R}$.

The relevant p.d.e. for U is the following *Hamilton-Jacobi equation*

$$(2.62) \quad U_t(t, x) + f(U_x(t, x)) = 0, \quad U(0, x) = U_0(x).$$

As a Lipschitz function U is differentiable almost everywhere. If (t, x) is a point of differentiability for U , then the partial derivatives $U_t(t, x)$ and $U_x(t, x)$ at this point satisfy equation (2.62).

For each $t \geq 0$ we can also define the partial derivative $\rho(t, x) = U_x(t, x)$ at Lebesgue almost every x . This function represents the limiting density of customer particles, for we can paraphrase limit (2.58) by saying that the random Radon measure $N^{-1} \sum_i \eta_i^N(\lfloor Nt \rfloor) \delta_{i/N}$ converges vaguely,

in probability, to the measure $\rho(t, x)dx$. Formally differentiating through equation (2.62) suggests that the correct p.d.e. for $\rho(t, x)$ is the *scalar conservation law*

$$(2.63) \quad \rho_t(t, x) + f(\rho(t, x))_x = 0, \quad \rho(0, x) = (U_0)_x(x).$$

A measurable function $\rho(t, x)$ is a *weak solution* of the initial value problem (2.63) if it satisfies this integral criterion for all compactly supported, continuously differentiable test functions ϕ on $\mathbb{R}_+ \times \mathbb{R}$:

$$(2.64) \quad \int_0^\infty \int_{\mathbb{R}} [\rho(t, x)\phi_t(t, x) + f(\rho(t, x))\phi_x(t, x)] dx dt + \int_{\mathbb{R}} \rho(0, x)\phi(0, x) dx = 0.$$

That $\rho(t, x) = U_x(t, x)$ satisfies (2.64) can be checked by multiplying equation (2.62) with ϕ_x and integrating by parts. Thus one message of the hydrodynamic limit is that the large scale motion of customer particles is governed by the conservation law (2.63).

We leave the discussion of the p.d.e. side at this informal level because a complete treatment would take up a significant amount of space. Here are some closing comments. The claims made above can be proved with small adjustments to the arguments used in Sections 3.3 and 3.4 of [Eva98]. Equation (2.56) is an example of a *Hopf-Lax formula* and in principle it identifies U as the unique *viscosity solution* of the Hamilton-Jacobi equation (2.62). The viscosity solution is the weak solution appropriate for these equations, originally developed in [CEL84] and [CL83]. [Eva98, Chapter 10] contains a general treatment of viscosity solutions for Hamilton-Jacobi equations, but the hypotheses used do not cover our case. To show that (2.56) defines the unique viscosity solution to (2.62), one can adapt the proof from [Eva98, Chapter 10] to show that U is a viscosity solution and then appeal to a uniqueness theorem for unbounded viscosity solutions from [Ish84].

Like the Hamilton-Jacobi equation, the conservation law (2.63) can possess multiple weak solutions. There are auxiliary conditions that select the physically relevant *entropy solution*. The solution $\rho(t, x) = U_x(t, x)$ defined above is the entropy solution, and the hydrodynamic limit shows that the particle system converges to the entropy solution.

The rest of this section proves Theorem 2.11. The only item we utilize from the past development is the existence of the general limit of Theorem 2.1. The proof separates naturally into four stages:

Stage 1. Derivation of a particle-level variational property that mimics formula (2.56).

Stage 2. Initial definition of the function g by the limit of the process that starts with the special initial conditions we already met in the discussion about exclusion and queues in Section 1.1.

Stage 3. Limits (2.58) and (2.59) under assumption (2.54) but with only partial knowledge about g .

Stage 4. Explicit evaluation of the function g . This is equivalent to proving Theorem 2.2.

Stage 1: The server process and the envelope property. For the proof we switch point of view: instead of having the customer particles jump from queue to queue, we let the server particles jump from customer to customer. Let random variable $z_i(t)$ represent the position of server i at time t . The connection between the queue lengths $\eta_i(t)$ and the server positions $z_i(t)$ is

$$(2.65) \quad \eta_i(t) = z_i(t) - z_{i-1}(t).$$

Completion of a service at queue i is now signified by a *leftward* jump of z_i . The entire process of server particles is denoted $z(t) = (z_i(t))_{i \in \mathbb{Z}}$.

As in the earlier description of the customer motion, we take i.i.d. $\{0, 1\}$ -valued mean p random variables $\{\kappa_i(t) : i \in \mathbb{Z}, t \in \mathbb{N}\}$ that are independent of the initial state $z(0)$. Variable $\kappa_i(t)$ signals a possible service completion at server i during time period $(t - 1, t)$. Reflecting the earlier rule

(2.53), the evolution from state $z(t-1)$ to state $z(t)$ obeys the equation

$$(2.66) \quad z_i(t) = z_i(t-1) - \kappa_i(t) \cdot \mathbf{1}\{z_{i-1}(t-1) < z_i(t-1)\} \quad i \in \mathbb{Z}, t \in \mathbb{N}.$$

Each server completes a service each time $t \in \mathbb{N}$ with probability p , provided this server had at least one customer in its queue at time $t-1$. The departure process from queue i is given by $d_i(t) = z_i(t-1) - z_i(t)$. Also, comparison with rule (1.9)–(1.10) shows that $z_i(t) + i$ is an exclusion process where particles march to the left. We make no use of this connection.

The state space of the server process $z(t)$ is

$$\mathbf{Z} = \{(z_i)_{i \in \mathbb{Z}} : z_i \in \mathbb{Z} \cup \{-\infty\}, z_{i-1} \leq z_i \forall i\}.$$

The value $-\infty$ is included to admit the possibility that there is a first server j in which case $z_i = -\infty$ for $i < j$. Servers at $-\infty$ never move and do not communicate in any way with the rest of the process. A similar convention for a last server is not really needed. Any server k can be regarded as the “last server” because as far as the evolution of servers $(z_i : i \leq k)$ goes, the subsequent servers $(z_i : i > k)$ are irrelevant.

To construct a queue length process $\eta(\cdot)$ from a given initial configuration $\eta(0)$, define an initial server configuration by

$$(2.67) \quad z_i(0) = \begin{cases} \sum_{j=1}^i \eta_j(0), & i > 0 \\ 0, & i = 0 \\ -\sum_{j=i+1}^0 \eta_j(0), & i < 0, \end{cases}$$

run the server process $z(\cdot)$ according to (2.66) and define $\eta(t)$ by (2.65).

The next *envelope property* for coupled server processes is our tool for proving the hydrodynamic limit. Let \mathcal{M} be a countable index set, and let $z(\cdot)$ and $\{x^m(\cdot) : m \in \mathcal{M}\}$ be server processes defined on the same probability space so that they obey the same coin flips. In other words each process $z(\cdot)$ and $x^m(\cdot)$ obeys equation (2.66) by itself, but the values $\{\kappa_i(t)\}$ are shared by all. We say that the server processes are coupled through the shared coin flips $\{\kappa_i(t)\}$.

LEMMA 2.12. (Envelope property) *In the situation described above, assume the initial configurations satisfy*

$$(2.68) \quad z_i(0) = \sup_{m \in \mathcal{M}} x_i^m(0) \quad \text{for all } i \in \mathbb{Z}, \text{ a.s.}$$

Then

$$(2.69) \quad z_i(t) = \sup_{m \in \mathcal{M}} x_i^m(t) \quad \text{for all } i \in \mathbb{Z} \text{ and } t \in \mathbb{Z}_+, \text{ a.s.}$$

PROOF. If $z_i(0) = -\infty$ then the hypothesis implies that $w_i^m(0) = -\infty$ for all $m \in \mathcal{M}$. These servers remain at $-\infty$ for the duration of the process, and so the conclusion holds for these indices i . We can restrict consideration to those particles that satisfy $z_i(t) \in \mathbb{Z}$ for all time.

Step 1. We claim that $x_i^m(t) \leq z_i(t)$ for all $i, t \in \mathbb{Z}_+$, and $m \in \mathcal{M}$ a.s. Proof is by induction on time. Suppose the property holds up to time $t-1$.

If $x_i^m(t-1) < z_i(t-1)$ then one jump cannot reverse the ordering.

If $x_i^m(t-1) = z_i(t-1)$ then since $x_{i-1}^m(t-1) \leq z_{i-1}(t-1)$, variable z_i cannot jump without variable x_i^m also jumping. Thus the ordering $x_i^m(t) \leq z_i(t)$ continues to hold at time t .

Step 2. We need to show that, given i and t , some $m \in \mathcal{M}$ satisfies $x_i^m(t) = z_i(t)$. Assume (2.69) holds up to time $t - 1$. Pick m_1 and m_2 such that

$$x_{i-1}^{m_1}(t-1) = z_{i-1}(t-1) \quad \text{and} \quad x_i^{m_2}(t-1) = z_i(t-1).$$

If $x_i^{m_2}$ does not jump during $(t-1, t)$ then neither can z_i by **Step 1** and we conclude that $x_i^{m_2}(t) = z_i(t)$.

Similarly, if both $x_i^{m_2}$ and z_i jump during $(t-1, t)$ again we have $x_i^{m_2}(t) = z_i(t)$.

Suppose $x_i^{m_2}$ jumps during $(t-1, t)$ but z_i does not. Then z_i must have been blocked by z_{i-1} , and we have

$$x_{i-1}^{m_1}(t-1) = z_{i-1}(t-1) = z_i(t-1).$$

By **Step 1** and the ordering in the process x^{m_1}

$$x_{i-1}^{m_1}(t-1) = z_{i-1}(t-1) = x_i^{m_1}(t-1) = z_i(t-1).$$

This implies that $x_i^{m_1}$ cannot jump during $(t-1, t)$ and thereby $x_i^{m_1}(t) = z_i(t)$. This completes the proof. \square

A consequence of the lemma is a form of monotonicity (proved in **Step 1** of the proof) that we also state for future reference.

LEMMA 2.13. *If z and x are two server processes coupled through shared coin flips and $z_i(0) \geq x_i(0)$ for all i , then $z_i(t) \geq x_i(t)$ continues to hold for all i and t .*

The auxiliary processes $x^m(\cdot)$ we use in Lemma 2.12 will be of the following special type. For $m \in \mathbb{Z}$ let the process $w^m(\cdot)$ evolve from the initial condition

$$(2.70) \quad w_i^m(0) = \begin{cases} 0, & i \geq m \\ -\infty, & i < m. \end{cases}$$

Then given an arbitrary initial configuration $z(0)$, for $m \in \mathbb{Z}$ define the processes

$$(2.71) \quad x_i^m(t) = z_m(0) + w_i^m(t).$$

Lemma 2.12 applies and yields

$$(2.72) \quad z_i(t) = \sup_{m \in \mathbb{Z}: m \leq i} \{z_m(0) + w_i^m(t)\}, \quad i \in \mathbb{Z}, t \in \mathbb{Z}_+.$$

The condition $m \leq i$ can be added to the supremum because $w_i^m(t) = -\infty$ for $m > i$. The equation is true also for the case $z_i(t) = -\infty$ because then $z_m(0) = -\infty$ for all $m \leq i$. The virtue of equation (2.72) is that inside the braces on the right the initial condition has been separated from the coin flips and for distinct m the processes $(w_{i-m}^m(\cdot) : i \in \mathbb{Z})$ are identical in distribution.

Stage 2: Limit for process $w(\cdot) = w^0(\cdot)$. At this stage we define the function g in terms of a limit of a special case of the process itself. We could appeal to Theorem 2.2 to give an explicit formula for g . However, we prefer to derive the formula for g from the particle system in Stage 4 of the proof.

THEOREM 2.14. *Let $w(\cdot) = w^0(\cdot)$ be the process with initial condition given in (2.70) with $m = 0$. Then there exists a continuous, concave, nondecreasing function $g : [0, \infty) \rightarrow [-p, 0]$ such that $g(0) = -p$, $g(x) < 0$ for $0 \leq x < p$, $g(x) = 0$ for $p \leq x < \infty$, and for all $x \in \mathbb{R}_+$ and $t \in (0, \infty)$*

$$(2.73) \quad \lim_{N \rightarrow \infty} N^{-1} w_{\lfloor Nx \rfloor}(\lfloor Nt \rfloor) = tg(x/t) \quad \text{almost surely.}$$

PROOF. To prove the existence of the limit in (2.73) we could resort to another application of the subadditive ergodic theorem together with the approximations necessary, as was done in the proof of Theorem 2.1. Let us bypass this work and take the limit from Theorem 2.1 by recasting the problem as one involving a last passage model.

Define a growing cluster on \mathbb{N}^2 by

$$\tilde{\mathcal{A}}(t) = \{(i, j) \in \mathbb{N}^2 : 1 \leq j \leq -w_{i-1}(t)\}.$$

Rule (2.66) implies that $\tilde{\mathcal{A}}$ fills in growth sites with independent p -coin flips. Therefore by Proposition 1.1 the process $\{\tau(i, j) : (i, j) \in \mathbb{N}^2\}$ of hitting times

$$\tau(i, j) = \inf\{t \in \mathbb{Z}_+ : (i, j) \in \tilde{\mathcal{A}}(t)\} = \inf\{t \in \mathbb{Z}_+ : w_{i-1}(t) \leq -j\}$$

has the same distribution as the process $\{G(i, j) : (i, j) \in \mathbb{N}^2\}$ of last-passage times defined by (1.3) with weights $Y_{i,j}$ geometrically distributed as in (1.6). By Theorem 2.1 the event

$$\left\{ \lim_{N \rightarrow \infty} N^{-1}G(\lfloor Nx \rfloor, \lfloor Ny \rfloor) = \Psi(x, y) \right\}$$

has probability 1 for any fixed $(x, y) \in (0, \infty)^2$, where Ψ is the limit function of that theorem. By the distributional equality $\{\tau(i, j)\} \stackrel{d}{=} \{G(i, j)\}$ of the processes, also

$$(2.74) \quad \mathbf{P}\left\{ \lim_{N \rightarrow \infty} N^{-1}\tau(\lfloor Nx \rfloor, \lfloor Ny \rfloor) = \Psi(x, y) \right\} = 1.$$

We use a simple large deviation bound to show that $\Psi < \infty$. There are $\binom{2N-2}{N-1} \leq 4^N$ paths from $(1, 1)$ to (N, N) . Fix any particular path π from $(1, 1)$ to (N, N) . The geometric distribution has an exponential moment. By Lemma A.4 we can pick $t > 0$ large enough so that, for some $C > 0$,

$$\mathbf{P}\{G(N, N) \geq Nt\} \leq 4^N \mathbf{P}\left\{ \sum_{\mathbf{x} \in \pi} Y_{\mathbf{x}} \geq Nt \right\} \leq e^{-CN} \rightarrow 0 \quad \text{as } N \rightarrow \infty.$$

The first inequality above is true because the sum of weights along any path has the same distribution as the sum along π . Then $\Psi(1, 1) \leq t$ and by Theorem 2.1 $\Psi < \infty$ on all of $(0, \infty)^2$.

We approach the function g through its negative. Define the function

$$(2.75) \quad \tilde{g}(x) = \inf\{y > 0 : \Psi(x, y) \geq 1\}, \quad x > 0.$$

This function is finite because $\Psi(x, y) \geq p^{-1}(x+y)$ and nonincreasing by the coordinatewise monotonicity of Ψ . By concavity of Ψ , $\Psi(x_1, y_1) \geq 1$ and $\Psi(x_2, y_2) \geq 1$ imply $\Psi(sx_1 + (1-s)x_2, sy_1 + (1-s)y_2) \geq 1$, and thereby

$$\tilde{g}(sx_1 + (1-s)x_2) \leq sy_1 + (1-s)y_2.$$

Letting $y_i \searrow \tilde{g}(x_i)$ shows \tilde{g} convex on $(0, \infty)$. Thereby \tilde{g} is also continuous on $(0, \infty)$. If $\tilde{g}(x) > 0$ then $\Psi(x, y) = 1$ iff $y = \tilde{g}(x)$ by the continuity and strict coordinatewise monotonicity of Ψ .

By the strong law of large numbers

$$N^{-1}G(\lfloor Nx \rfloor, 1) = N^{-1} \sum_{i=1}^{\lfloor Nx \rfloor} Y_{i,1} \xrightarrow[N \rightarrow \infty]{} x \mathbf{E}Y_{1,1} = xp^{-1}$$

and thereby $\Psi(x, y) \geq 1$ for all $x \geq p$ and $y > 0$. This implies $\tilde{g}(x) = 0$ for $x \geq p$. By symmetry $\Psi(x, y) \geq 1$ for all $x > 0$ and $y \geq p$ which implies $\tilde{g}(x) \leq p$.

To show $\tilde{g}(0+) \geq p$ we use another large deviation estimate. Fix a small $\varepsilon \in (0, 1)$ and let $x \in (0, p\varepsilon/4)$ and $y = p(1-\varepsilon)$. Consider N large enough so that $(\lfloor Nx \rfloor, \lfloor Ny \rfloor) \in \mathbb{N}^2$. For an

individual path π from $(1, 1)$ to $(\lfloor Nx \rfloor, \lfloor Ny \rfloor)$, by Lemma A.4,

$$\begin{aligned} \mathbf{P}\left\{\sum_{\mathbf{x} \in \pi} Y_{\mathbf{x}} \geq N(1 - \varepsilon/4)\right\} &\leq \mathbf{P}\left\{\sum_{\mathbf{x} \in \pi} (Y_{\mathbf{x}} - \mathbf{E}Y_{\mathbf{x}}) \geq N\varepsilon/2\right\} \leq e^{-(\lfloor Nx \rfloor + \lfloor Ny \rfloor - 1)A(\varepsilon/2)} \\ &\leq e^{-C(\varepsilon)N} \end{aligned}$$

for a constant $C(\varepsilon) > 0$ and large enough N . The rearranged right-hand side inside the probability came from

$$\sum_{\mathbf{x} \in \pi} \mathbf{E}Y_{\mathbf{x}} = (\lfloor Nx \rfloor + \lfloor Ny \rfloor - 1)p^{-1} \leq N(x + y)p^{-1} \leq N(1 - 3\varepsilon/4).$$

Next estimate the number of paths as a function of x and large N . Put temporarily $k = \lfloor Nx \rfloor - 1$ and $\ell = \lfloor Ny \rfloor - 1$. By Stirling's formula, for large N ,

$$\begin{aligned} \text{the number of paths from } (1, 1) \text{ to } (\lfloor Nx \rfloor, \lfloor Ny \rfloor) &= \binom{\lfloor Nx \rfloor + \lfloor Ny \rfloor - 2}{\lfloor Nx \rfloor - 1} \\ &= \binom{k + \ell}{k} \leq C \left(\frac{k + \ell}{k\ell}\right)^{1/2} \frac{(k + \ell)^{k + \ell}}{k^k \ell^\ell} \leq \frac{C}{\sqrt{Nx}} \exp[k \log(1 + \ell/k) + \ell \log(1 + k/\ell)] \\ &\leq C(Nx)^{-1/2} e^{Nx \log(1 + C_1 y/x) + Nx} \leq C e^{N\delta(x) - \frac{1}{2} \log Nx} \end{aligned}$$

where $\delta(x) \rightarrow 0$ as $x \rightarrow 0$. In the second last inequality we used $\ell/k \leq C_1 y/x$ for large N and $\log(1 + k/\ell) \leq k/\ell$. The two estimates together give

$$\mathbf{P}\{G(\lfloor Nx \rfloor, \lfloor Ny \rfloor) \geq N(1 - \varepsilon/4)\} \leq \exp[-N(C(\varepsilon) - \delta(x) + \frac{1}{2N} \log Nx)].$$

If x is fixed small enough relative to ε , the last bound above tends to 0 as $N \rightarrow \infty$. This implies $\Psi(x, y) \leq 1 - \varepsilon/4$ which forces $\tilde{g}(x) \geq y = p(1 - \varepsilon)$. Since this is true for small enough x we have $\tilde{g}(0+) \geq p(1 - \varepsilon)$, and since $\varepsilon > 0$ can be taken arbitrarily small we have $\tilde{g}(0+) \geq p$. Together with the earlier bound $\tilde{g}(x) \leq p$ this gives $\tilde{g}(0+) = p$.

Take the previous result and turn it around by symmetry to say $\Psi(x, y) \leq 1 - \varepsilon/4$ for $x = p(1 - \varepsilon)$ and $y \in (0, p\varepsilon/4)$. Then $\tilde{g}(x) \geq p\varepsilon/4 > 0$. Since $\varepsilon > 0$ can be taken arbitrarily small we conclude that $\tilde{g}(x) > 0$ for all $x \in (0, p)$.

To summarize, we have shown that the function

$$(2.76) \quad g(x) = \begin{cases} -\tilde{g}(x), & x > 0 \\ -p, & x = 0 \end{cases}$$

has the properties claimed in the statement of the theorem. It remains to prove the limit (2.73).

Fix $x, t > 0$. Let Ω_1 be the event on which

$$N^{-1}\tau(\lfloor Nx' \rfloor, \lfloor Ny \rfloor) \rightarrow \Psi(x', y)$$

for all (x', y) such that x' is either x or rational and $y = t\tilde{g}(x'/t) \pm \varepsilon$ for rational $\varepsilon > 0$. For $y \leq 0$ the condition is irrelevant and can be ignored. Restrict consideration to the event Ω_1 which has probability 1 by (2.74).

First the upper bound for the limit (2.73). Suppose $g(x/t) < 0$. Otherwise there is nothing to prove. Take $x' > x$ close enough to x and $\varepsilon > 0$ small enough so that $t\tilde{g}(x'/t) \geq t\tilde{g}(x/t) - \varepsilon$ and $y = t\tilde{g}(x'/t) - \varepsilon > 0$. Then

$$\Psi(x', y) = t\Psi(x'/t, \tilde{g}(x'/t) - \varepsilon/t) \leq t - \delta$$

for some $\delta > 0$. (More precisely, $\Psi(x', y+h) - \Psi(x', y) \geq h/p$ for $h > 0$ by the law of large numbers.) Consequently for all large enough N , $\tau(\lfloor Nx' \rfloor, \lfloor Ny \rfloor) \leq Nt - N\delta/2$, which in turn implies again for large enough N

$$\begin{aligned} w_{\lfloor Nx \rfloor}(\lfloor Nt \rfloor) &\leq w_{\lfloor Nx' \rfloor - 1}(\lfloor Nt \rfloor) \leq -\lfloor Ny \rfloor \leq -Nt\tilde{g}(x'/t) + N\varepsilon + 1 \\ &\leq -Nt\tilde{g}(x/t) + 2N\varepsilon + 1. \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary,

$$\overline{\lim}_{N \rightarrow \infty} N^{-1} w_{\lfloor Nx \rfloor}(\lfloor Nt \rfloor) \leq tg(x/t) \quad \text{on the event } \Omega_1.$$

The lower bound comes by an analogous argument. Suppose $y = t\tilde{g}(x/t) + \varepsilon$. Then

$$\Psi(x, y) = t\Psi(x/t, \tilde{g}(x/t) + \varepsilon/t) \geq t + \delta$$

for some $\delta > 0$. For large enough N , $\tau(\lfloor Nx \rfloor, \lfloor Ny \rfloor) \geq Nt + N\delta/2$, which implies

$$w_{\lfloor Nx \rfloor}(\lfloor Nt \rfloor) \geq w_{\lfloor Nx \rfloor - 1}(\lfloor Nt \rfloor) > -\lfloor Ny \rfloor \geq -Nt\tilde{g}(x/t) - N\varepsilon.$$

We have shown that the limit (2.73) holds on the event Ω_1 for $x > 0$.

It remains to argue the limit $N^{-1}w_0(\lfloor Nt \rfloor) \rightarrow tg(0) = -tp$ for $x = 0$. This is the classical strong law of large numbers because $w_0(t)$ advances by p -coin flips, without obstruction. This completes the proof of the theorem. \square

Stage 3: Limits for the particle system. Assume given the nondecreasing function U_0 . Let the function g be the one given in Theorem 2.14, defined by the limit (2.73). Once we have a more general limit we can prove that this function g actually agrees with (2.55). Let $U(t, x)$ be defined by (2.56) in terms of the given initial profile U_0 and the function g defined by limit (2.73).

On the particle side assume given the initial queue length configurations $\eta^N(0)$ that satisfy assumption (2.54). Define initial server configurations $z^N(0)$ by (2.67), run the server processes according to equation (2.66), and define the queue length processes $\eta^N(\cdot)$ by (2.65).

Assumption (2.54) gives

$$(2.77) \quad \lim_{N \rightarrow \infty} \mathbf{P} \{ |N^{-1}z_{\lfloor Nx \rfloor}^N(0) - U_0(x)| \geq \varepsilon \} = 0 \quad \text{for } x \in \mathbb{R} \text{ and } \varepsilon > 0.$$

The statement that needs to be proved is that, for $(t, x) \in (0, \infty) \times \mathbb{R}$ and $\varepsilon > 0$,

$$(2.78) \quad \lim_{N \rightarrow \infty} \mathbf{P} \{ |N^{-1}z_{\lfloor Nx \rfloor}^N(\lfloor Nt \rfloor) - U(t, x)| \geq \varepsilon \} = 0.$$

Since

$$\sum_{i=\lfloor Na \rfloor + 1}^{\lfloor Nb \rfloor} \eta_i^N(\lfloor Nt \rfloor) = z_{\lfloor Nb \rfloor}^N(\lfloor Nt \rfloor) - z_{\lfloor Na \rfloor}^N(\lfloor Nt \rfloor)$$

and

$$\sum_{s=1}^{\lfloor Nt \rfloor} d_{\lfloor Na \rfloor}^N(s) = z_{\lfloor Na \rfloor}^N(0) - z_{\lfloor Na \rfloor}^N(\lfloor Nt \rfloor)$$

both limits (2.58) and (2.59) of Theorem 2.11 follow from (2.78).

Rewrite the variational identity (2.72) for each process $z^N(\cdot)$ and replace the discrete variables with integer parts of scaled, continuous variables.

$$(2.79) \quad N^{-1}z_{\lfloor Nx \rfloor}^N(\lfloor Nt \rfloor) = \sup_{y \in \mathbb{R} : y \leq x} \{ N^{-1}z_{\lfloor Ny \rfloor}^N(0) + N^{-1}w_{\lfloor Nx \rfloor}^{N, \lfloor Ny \rfloor}(\lfloor Nt \rfloor) \}, \quad x \in \mathbb{R}, t > 0.$$

Each process $z^N(\cdot)$ has coin flips $\{\kappa_i^N(t) : i \in \mathbb{Z}, t \in \mathbb{N}\}$ for generating its dynamics, and the processes $\{w^{N, m}(\cdot) : m \in \mathbb{Z}\}$ are coupled with z^N through these coin flips.

Distributionally the process $\{w_{\lfloor Nx \rfloor}^{N, \lfloor Ny \rfloor}(\lfloor Nt \rfloor) : t \geq 0\}$ is equal to the process $\{w_{\lfloor Nx \rfloor - \lfloor Ny \rfloor}(\lfloor Nt \rfloor) : t \geq 0\}$ discussed in Theorem 2.14, and hence by that theorem,

$$(2.80) \quad \lim_{N \rightarrow \infty} N^{-1} w_{\lfloor Nx \rfloor}^{N, \lfloor Ny \rfloor}(\lfloor Nt \rfloor) = tg\left(\frac{x-y}{t}\right) \quad \text{in probability, for all } t > 0 \text{ and } y \leq x \text{ in } \mathbb{R}.$$

To appeal to the limit (2.73) precisely in its stated form, the position argument $\lfloor Nx \rfloor - \lfloor Ny \rfloor$ in $w_{\lfloor Nx \rfloor - \lfloor Ny \rfloor}(\lfloor Nt \rfloor)$ must be replaced by something of the form $\lfloor Nr \rfloor$. To this end take $x' > x$, use $\lfloor N(x' - y) \rfloor \geq \lfloor Nx \rfloor - \lfloor Ny \rfloor \geq \lfloor N(x - y) \rfloor$ and the monotonicity of $w_i(t)$ in i , and after passing to the limit use the continuity of g to take $x' \searrow x$. Presently we cannot assert almost sure convergence in (2.80) because the process itself changes with N .

Fix $x \in \mathbb{R}$ and $t > 0$. The path towards the limit (2.78) is pretty clear since naively taking $N \rightarrow \infty$ in (2.79) leads to a variational formula of the type (2.56). Due to the presence of the supremum we argue separately upper and lower bounds, beginning with the easier lower bound.

Given $\varepsilon > 0$, pick $y \in (-\infty, x]$ such that

$$U_0(y) + tg\left(\frac{x-y}{t}\right) \geq U(t, x) - \varepsilon.$$

By the convergence in probability in (2.77) and (2.80), we can find N_0 such that for $N \geq N_0$

$$\mathbf{P}^N \{N^{-1} z_{\lfloor Ny \rfloor}^N(0) \geq U_0(y) - \varepsilon\} \geq 1 - \varepsilon/2$$

and

$$\mathbf{P}^N \left\{ N^{-1} w_{\lfloor Nx \rfloor}^{N, \lfloor Ny \rfloor}(\lfloor Nt \rfloor) \geq tg\left(\frac{x-y}{t}\right) - \varepsilon \right\} \geq 1 - \varepsilon/2.$$

These bounds combined with (2.79) give

$$\mathbf{P}^N \{N^{-1} z_{\lfloor Nx \rfloor}^N(\lfloor Nt \rfloor) \geq U(t, x) - 3\varepsilon\} \geq 1 - \varepsilon \quad \text{for } N \geq N_0.$$

This is one half of the desired limit (2.78).

To get the other half we must bound the supremum in (2.79) from above and for that we need some truncation, discretization and estimation. The first step is to restrict the supremum in (2.79) to a bounded interval of y -values. With (t, x) fixed, define for $v < x$

$$\zeta^N(v) = \sup_{y \in [v, x]} \{z_{\lfloor Ny \rfloor}^N(0) + w_{\lfloor Nx \rfloor}^{N, \lfloor Ny \rfloor}(\lfloor Nt \rfloor)\}.$$

LEMMA 2.15. *If $v < x - tp$ then*

$$\lim_{N \rightarrow \infty} \mathbf{P}^N \{z_{\lfloor Nx \rfloor}^N(\lfloor Nt \rfloor) \neq \zeta^N(v)\} = 0.$$

PROOF. Forget the scaling for a moment and work with the basic variational equality (2.72) with fixed i . Suppose that in that formula $w_i^\ell(t) = 0$ for some index $\ell \leq i$. Then, by $w_i^m(t) \leq 0$ and by the monotonicity of the initial server locations $\{z_m(0)\}$, for $m < \ell$

$$z_m(0) + w_i^m(t) \leq z_m(0) \leq z_\ell(0) = z_\ell(0) + w_i^\ell(t).$$

This implies that indices $m < \ell$ cannot contribute to the supremum, and so

$$z_i(t) = \max_{\ell \leq m \leq i} \{z_m(0) + w_i^m(t)\}.$$

Pick $y \in (v, x - tp)$. For the situation at hand the above argument implies

$$\mathbf{P}^N \{z_{\lfloor Nx \rfloor}^N(\lfloor Nt \rfloor) \neq \zeta^N(v)\} \leq \mathbf{P}^N \{w_{\lfloor Nx \rfloor}^{N, \lfloor Ny \rfloor}(\lfloor Nt \rfloor) < 0\}.$$

Initially $w_i^{N, \lfloor Ny \rfloor}(0) = 0$ for $i = \lfloor Ny \rfloor, \lfloor Ny \rfloor + 1, \lfloor Ny \rfloor + 2, \dots$. Since the particles move with p -coin flips and a jump is possible only after the previous particle has jumped out of the way, the first jump time

$$\tau^N = \inf\{s \in \mathbb{N} : w_{\lfloor Nx \rfloor}^{N, \lfloor Ny \rfloor}(s) < 0\}$$

is a sum of $\lfloor Nx \rfloor - \lfloor Ny \rfloor + 1$ i.i.d. variables Y_i with common distribution $P[Y_i = k] = pq^{k-1}$ for $k \in \mathbb{N}$. These variables have mean p^{-1} . By the law of large numbers $N^{-1}\tau^N \rightarrow (x - y)p^{-1} > t$ in probability, and consequently

$$\mathbf{P}^N\{w_{\lfloor Nx \rfloor}^{N, \lfloor Ny \rfloor}(\lfloor Nt \rfloor) < 0\} \rightarrow 0. \quad \square$$

Fix $v < x - tp$. To complete the proof of (2.78) it is enough to show

$$(2.81) \quad \lim_{N \rightarrow \infty} \mathbf{P}\{N^{-1}\zeta^N(v) \geq U(t, x) + \varepsilon\} = 0.$$

Utilizing the continuity of the function g of Theorem 2.14, pick a partition $v = y_0 < y_1 < \dots < y_m = x$ such that

$$\left| tg\left(\frac{x - y_j}{t}\right) - tg\left(\frac{x - y_{j+1}}{t}\right) \right| \leq \frac{\varepsilon}{4} \quad \text{for } j = 0, \dots, m-1.$$

By the monotonicity Lemma 2.13, if $m_1 < m_2$ then $w_i^{m_2}(t) \leq w_i^{m_1}(t)$ because this inequality holds at time 0. We bound $\zeta^N(v)$ in terms of partition points as follows:

$$\begin{aligned} \zeta^N(v) &= \max_{0 \leq j < m} \max_{\lfloor Ny_j \rfloor \leq m \leq \lfloor Ny_{j+1} \rfloor} \{z_m^N(0) + w_{\lfloor Nx \rfloor}^{N, m}(\lfloor Nt \rfloor)\} \\ &\leq \max_{0 \leq j < m} \{z_{\lfloor Ny_{j+1} \rfloor}^N(0) + w_{\lfloor Nx \rfloor}^{N, \lfloor Ny_j \rfloor}(\lfloor Nt \rfloor)\}. \end{aligned}$$

The last line multiplied by N^{-1} converges in probability to

$$\begin{aligned} &\max_{0 \leq j < m} \left\{ U_0(y_{j+1}) + tg\left(\frac{x - y_j}{t}\right) \right\} \\ &\leq \max_{0 \leq j < m} \left\{ U_0(y_{j+1}) + tg\left(\frac{x - y_{j+1}}{t}\right) \right\} + \frac{\varepsilon}{4} \\ &\leq U(t, x) + \frac{\varepsilon}{4}. \end{aligned}$$

This implies limit (2.81) and finishes the proof of (2.78). We have now proved Theorem 2.11 except for the explicit characterization (2.55) of g .

Stage 4: Computation of the profile g . To complete the proof of the hydrodynamic limit we apply the limit (2.78) proved thus far to a stationary system to compute the function g explicitly. From Theorem 2.14 we already know the values $g(0) = p$ and $g(x) = 0$ for $x \geq p$.

Fix the parameter $r \in (0, p)$. Let the initial queue lengths $\eta(0) = (\eta_i(0))_{i \in \mathbb{Z}}$ have the product distribution $\nu = \alpha^{\otimes \mathbb{Z}}$ with marginal α defined in (2.51). According to Lemma 2.9 ν is invariant for the queue length process $\eta(\cdot)$ and in this situation the departure process $-z_0(t)$ is a Bernoulli r process. Thus

$$U_0(y) = \lim_{N \rightarrow \infty} N^{-1} z_{\lfloor Ny \rfloor}(0) = y \mathbf{E} \eta_0(0) = y \frac{r(1-r)}{p-r}$$

and

$$U(1, 0) = \lim_{N \rightarrow \infty} N^{-1} z_0(N) = -r.$$

Substituting these into (2.56) with $(t, x) = (1, 0)$ gives

$$(2.82) \quad -r = \sup_{y \leq 0} \left\{ y \frac{r(1-r)}{p-r} + g(-y) \right\}.$$

The above equation is valid for the range $0 < r < p$. By inspection, it also works for $r = 0$. Define the new variable

$$(2.83) \quad \rho = \frac{r(1-r)}{p-r} = \mathbf{E}\eta_0(0)$$

which is a strictly increasing convex function of $r \in [0, p)$ onto $[0, \infty)$. Then $r = f(\rho)$ is a strictly increasing concave function of $\rho \in [0, \infty)$. Solving explicitly from (2.83) gives formula (2.60) for f . Extend g to a concave upper semicontinuous function on all of \mathbb{R} by setting $g \equiv -\infty$ on $(-\infty, 0)$. Replace y by $-y$, and then (2.82) can be written as

$$(2.84) \quad f(\rho) = \inf_{y \in \mathbb{R}} \{y\rho - g(y)\}.$$

For $\rho < 0$ the infimum on the right gives $-\infty$ because $g(y)$ stays bounded as $y \nearrow \infty$. Thus we can extend $f(\rho)$ to a concave upper semicontinuous function by setting $f(\rho) = -\infty$ on $(-\infty, 0)$. By concave duality (Corollary C.2 in Appendix C.1)

$$(2.85) \quad g(y) = \inf_{\rho \geq 0} \{\rho y - f(\rho)\}.$$

With some calculus one can solve explicitly for

$$(2.86) \quad g(y) = -(\sqrt{p(1-y)} - \sqrt{qy})^2 \quad \text{for } 0 \leq y < p.$$

This verifies (2.55) and completes the proof of Theorem 2.11. Along the way we verified the concave duality (2.61) of the flux f and the special shape g .

Let us also point out that the last computation reproves the limit for the last-passage model with geometric weights.

COROLLARY 2.16. *Let the i.i.d. weights $\{Y_{i,j}\}_{(i,j) \in \mathbb{N}^2}$ have common geometric distribution $P[Y_{i,j} = k] = pq^{k-1}$ for $k \in \mathbb{N}$. Define the limit $\Psi(x, y)$ of the last-passage times as in (2.2). Then*

$$(2.87) \quad \Psi(x, y) = p^{-1}(x + y + 2\sqrt{qxy}) \quad \text{for } (x, y) \in (0, \infty)^2.$$

PROOF. Combining the definition (2.75) of \tilde{g} with $\tilde{g} = -g$ (2.76) and the explicit expression (2.86) for g above gives this statement: for $0 < x < p$, $\Psi(x, \tilde{g}(x)) = 1$ for $\tilde{g}(x) = (\sqrt{p(1-x)} - \sqrt{qx})^2$. Formula (2.87) can be derived from this with the help of the homogeneity of Ψ . Given $(x, y) \in (0, \infty)^2$, let $c^{-1} = \Psi(x, y)$. Then $\Psi(cx, cy) = 1$ from which $cy = \tilde{g}(cx)$, and c can be found from the last equation. \square

2.5. Queues and the last-passage model revisited

In this section we establish a connection between the last-passage model with boundaries studied in Section 2.2 and the M/M/1 queues. As before $0 < p < 1$ is fixed. To each $r \in (0, p)$ associate a “dual” parameter

$$(2.88) \quad r_* = \frac{p-r}{1-r} \in (0, p).$$

Let $\{G_0(m, n) : (m, n) \in \mathbb{Z}_+^2\}$ be the last-passage times in the last-passage model on \mathbb{Z}_+^2 as defined in (2.19), with these weight distributions for independent $\{Y_{i,j}\}$: for $k \in \mathbb{N}$,

$$(2.89) \quad \mathbf{P}\{Y_{0,0} = 0\} = 1,$$

$$(2.90) \quad \mathbf{P}\{Y_{i,0} = k\} = r_*(1 - r_*)^{k-1} \quad \text{for } i \in \mathbb{N},$$

$$(2.91) \quad \mathbf{P}\{Y_{0,j} = k\} = r(1 - r)^{k-1} \quad \text{for } j \in \mathbb{N},$$

$$(2.92) \quad \mathbf{P}\{Y_{i,j} = k\} = pq^{k-1} \quad \text{for } i, j \in \mathbb{N}.$$

These distributions are exactly those used in (2.15)–(2.18) in Section 2.2 except that the geometric variables have been shifted by 1.

To create a queueing situation that matches this last-passage picture, we construct a system in which the departure process from a fixed queue is a Bernoulli r process, and an individual customer jumps from server to server as a Bernoulli r_* process.

Consider an infinite sequence of M/M/1 servers labeled by $i \in \mathbb{N}$ as in Section 1.1. Customers enter the system at server 1 and move from server i to $i + 1$ in order, under FIFO discipline. But instead of having all customers queue up at server 1 at time $t = 0$, imagine that the system has been in operation for a long time, fed by a mean r Bernoulli arrival process at server $i = 1$. Consequently we can assume that the queue length process $\{\eta_i(t) : i \in \mathbb{N}\}$ for times $t \leq -1$ is stationary with product distribution $\alpha^{\otimes \mathbb{N}}$, as described in Lemma 2.9.

During period $(-1, 0)$ a special customer labeled 0 shows up, outside the arrival process, and enters the system at server 1. After customer 0 the Bernoulli r arrival process resumes as before, from period $(0, 1)$ onwards. (Perhaps customer 0 is you, cutting in line?) We need to be precise about what customer 0 sees upon arrival, so let us stipulate that customer 0 arrives during $(-1, 0)$ after the service events of period $(-1, 0)$ have taken place.

Let $\tau(0, j)$ denote the arrival time of customer $j \in \mathbb{Z}_+$ at server 1. Thus $\tau(0, 0) = 0$, and then, since the subsequent arrival process is a mean r Bernoulli process, processes $\{\tau(0, j) : j \in \mathbb{Z}_+\}$ and $\{G_0(0, j) : j \in \mathbb{Z}_+\}$ are equal in distribution. For $i \geq 1, j \geq 0$ let $\tau(i, j)$ be the time when customer j departs server i and joins the queue at server $i + 1$. More precisely, if $\tau(i, j) = t \in \mathbb{N}$ then the event in question happens during period $(t - 1, t)$. The result is that the last-passage process with boundaries captures this queueing process.

THEOREM 2.17. *The processes $\{G_0(m, n) : (m, n) \in \mathbb{Z}_+^2\}$ and $\{\tau(m, n) : (m, n) \in \mathbb{Z}_+^2\}$ are equal in distribution.*

Let us first be clear about what needs to be proved. We already know the equality $\{\tau(0, j) : j \in \mathbb{Z}_+\} \stackrel{d}{=} \{G_0(0, j) : j \in \mathbb{Z}_+\}$ on the j -axis. Suppose we know the similar equality in distribution on the i -axis. Then induction takes over and proves Theorem 2.17. Namely, after time $\tau(i - 1, j) \vee \tau(i, j - 1)$ the first successful p -coin flip at server i sends customer j away from server i and marks the occurrence of time $\tau(i, j)$. In other words, the system $\{\tau(i, j)\}$ follows the last-passage recipe with weight distributions (2.89)–(2.92). One can write a rigorous argument along the lines of Proposition 1.1 by showing that $\{\tau(i, j)\}$ and $\{G_0(i, j)\}$ determine the same cluster processes in the sense of distributions.

Thus it suffices to prove the equality in distribution on the i -axis:

$$\text{PROPOSITION 2.18. } \{\tau(i, 0) : i \in \mathbb{Z}_+\} \stackrel{d}{=} \{G_0(i, 0) : i \in \mathbb{Z}_+\}$$

The remainder of the section proves the proposition. As already mentioned above, we assume that during period $(-1, 0)$ all service completions are done “by time $-\frac{1}{2}$ ” before customer 0 arrives at server 1. Consequently the queues $\{\eta_i(-\frac{1}{2}) : i \in \mathbb{N}\}$ that customer 0 sees upon arrival are the

equilibrium queues $\{\eta_i(-1) : i \in \mathbb{N}\}$ from time $t = -1$ modified by the service completions of period $(-1, 0)$.

LEMMA 2.19. *The distribution seen by the arriving customer 0 is*

$$(2.93) \quad \mathbf{P}\{\eta_1(-\tfrac{1}{2}) = x_1, \eta_2(-\tfrac{1}{2}) = x_2, \dots, \eta_n(-\tfrac{1}{2}) = x_n\} = (1-u)u^{x_1} \cdot \prod_{i=2}^n \alpha(x_i)$$

for $x_1, \dots, x_n \in \mathbb{Z}_+$.

PROOF. As indicated $\eta(-1)$ has $\alpha^{\otimes \mathbb{N}}$ distribution. From time -1 to $-1/2$ there is no arrival from the outside, but a round of potential services happens, triggered by variables $\{\kappa_i(0)\}$. The evolution is given by the equations

$$(2.94) \quad \begin{aligned} \eta_1(-\tfrac{1}{2}) &= (\eta_1(-1) - \kappa_1(0))_+ \\ \eta_i(-\tfrac{1}{2}) &= (\eta_i(-1) - \kappa_i(0))_+ + d_{i-1}(0), \quad i \geq 2. \end{aligned}$$

Let us compute the distribution of $(\eta_1(-\tfrac{1}{2}), d_1(0))$:

$$\begin{aligned} \mathbf{P}\{\eta_1(-\tfrac{1}{2}) = x, d_1(0) = 1\} &= \mathbf{P}\{\eta_1(-1) = x+1, \kappa_1(0) = 1\} = \frac{p-r}{pq} u^{x+1} p \\ &= (1-u)u^x r. \end{aligned}$$

Thus the queue $\eta_1(-\tfrac{1}{2})$ has distribution $(1-u)u^x$ and is independent of the Bernoulli r arrival $a_2(0) = d_1(0)$ sent to server 2. The queues $\{\eta_i(-1) : i \geq 2\}$ are initially in the $\alpha^{\otimes \{2,3,4,\dots\}}$ equilibrium and independent of $(\eta_1(-\tfrac{1}{2}), d_1(0))$. It follows from Lemma 2.9 that after the service step the resulting queues $\{\eta_i(-\tfrac{1}{2}) : i \geq 2\}$ are again in the $\alpha^{\otimes \{2,3,4,\dots\}}$ equilibrium and independent of $\eta_1(-\tfrac{1}{2})$. \square

Starting from r_* defined in (2.88) we define dual counterparts of the constant u and the measure α from (2.50) and (2.51):

$$u_* = \frac{ur^*}{p(1-r_*)} = \frac{p-r}{p} = \alpha(0)$$

and

$$\alpha_*(0) = \frac{p-r_*}{p} = u \quad \text{and} \quad \alpha_*(x) = \frac{p-r_*}{pq} u_*^x \quad \text{for } x \geq 1.$$

Next, we record the situation seen by customer 0 upon arrival in a different way. The customers he sees ahead of himself in the queues are the customers $-1, -2, -3, \dots$ that arrived before him. They have stayed in order, with customer -1 ahead of customer 0, customer -2 ahead of customer -1 , and so on. Let S_ℓ be the label of the server at which customer $\ell \in -\mathbb{N}$ resides at time $t = -\frac{1}{2}$ when customer 0 arrives. Let

$$\xi_{-k} = S_{-k} - S_{-k+1}, \quad k \in \mathbb{N}.$$

Variable ξ_{-k} counts the number of servers between customer $-k+1$ and the next customer $-k$. $\xi_{-k} = 0$ if customers $-k+1$ and $-k$ are at the same server. $\xi_{-k} = y > 0$ if customer $-k+1$ is at the front of his queue at server S_{-k+1} , the next $y-1$ queues ahead are empty of customers, and the next customer $-k$ is at server $S_{-k+1} + y$. From (2.93) we derive the distribution of $\{\xi_{-k}\}$.

LEMMA 2.20. *Variables $\{\xi_{-k} : k \in \mathbb{N}\}$ have the product distribution $\alpha_*^{\otimes \mathbb{N}}$.*

PROOF. Let $N \in \mathbb{N}$, $k_1, \dots, k_N \in \mathbb{Z}_+$, $y_1, \dots, y_N \in \mathbb{N}$ and set $K_m = \sum_{i=1}^m (1+k_i)$. Define the event

$$\begin{aligned} A = \{ & \xi_{-1} = \dots = \xi_{-k_1} = 0, \xi_{-k_1-1} = y_1, \xi_{-k_1-2} = \dots = \xi_{-k_1-k_2-1} = 0, \xi_{-k_1-k_2-2} = y_2, \\ & \dots, \xi_{-K_{N-1}-1} = \dots = \xi_{-K_N+1} = 0, \xi_{-K_N} = y_N \}. \end{aligned}$$

To clarify, this event prescribes $k_1 + \dots + k_N$ ξ -values zero, and N ξ -values equal to $y_1, \dots, y_N > 0$. It suffices to show that

$$(2.95) \quad \mathbf{P}(A) = \alpha_*(0)^{k_1 + \dots + k_N} \cdot \prod_{i=1}^N \alpha_*(y_i).$$

This is sufficiently general for the conclusion because an event with adjacent positive ξ -values can be prescribed by setting some $k_i = 0$, and the last value can be zero by summing over $y_N > 0$ and taking the complement.

Let us abbreviate $\eta_i(-\frac{1}{2}) = \eta_i$ for this proof. Set $Y_n = y_1 + \dots + y_n$. Restate A in terms of $\{\eta_i\}$:

$$A = \left\{ \eta_1 = k_1, \eta_2 = \dots = \eta_{y_1} = 0, \eta_{y_1+1} = k_2 + 1, \eta_{y_1+2} = \dots = \eta_{y_1+y_2} = 0, \right. \\ \left. \dots, \eta_{Y_{N-1}+1} = k_N + 1, \eta_{Y_{N-1}+2} = \dots = \eta_{y_N} = 0, \eta_{Y_N+1} > 0 \right\}.$$

In words, this event specifies that $\eta_1 = k_1$ (the special first queue in (2.93)), and after that there are $Y_N - N$ empty queues, $N - 1$ queues with lengths $k_i + 1$ ($2 \leq i \leq N$), and finally a nonempty queue. By (2.93)

$$\mathbf{P}(A) = (1 - u)u^{k_1} \cdot \alpha(0)^{y_1 + \dots + y_N - N} \cdot \left\{ \prod_{i=2}^N \alpha(k_i + 1) \right\} \cdot (1 - \alpha(0))$$

which simplifies to the right-hand side of (2.95). \square

Now we wish to switch around the meaning of customers and servers. To make the text intelligible we continue to use the terms customer and server as used up to now, for the customers $j \in \mathbb{Z}_+$ and servers $i \in \mathbb{N}$ introduced in the beginning of this section. We call the new entities **customers* and **servers*. Think of the variables $\{\xi_{-k}\}$ as initial **queue* lengths. **Queue* $-k$ receives a **customer* when S_{-k} increases by 1 (customer $-k$ jumps from one queue to the next) and sends off a **customer* when S_{-k+1} increases by 1 (customer $-k + 1$ jumps from one queue to the next). Think of the variables S_{-k} as the positions of **servers* that jump when a **customer* completes service.

The variables $\{\dots, \xi_{-3}, \xi_{-2}, \xi_{-1}\}$ are initially in i.i.d. α_* equilibrium. Lemma 2.9 implies that the **departure* process from **queue* ξ_{-1} is a Bernoulli r_* process. However, the **departure* process from **queue* ξ_{-1} corresponds to the motion of S_0 , the position of customer 0 among the servers. Precisely speaking, if i **customers* have departed from **queue* ξ_{-1} during time $1, \dots, t$, then the position of customer 0 among the servers satisfies $S_0(t) = i + 1$. $S_0(t) = i + 1$ rather than i because $S_0(0) = 1$ and not 0.

Consequently, the marginal distribution of the position process $S_0(\cdot)$ of customer 0 is the same as obtained by flipping an r_* coin to determine when to jump to the next server. The time $\tau(i, 0)$ when customer 0 departs server i ($i \geq 1$) is then a sum of i i.i.d. geometric variables with common distribution $\mathbf{P}[Y = s] = r_*(1 - r_*)^{s-1}$, $s \in \mathbb{N}$. The distributional equality claimed in Proposition 2.18 holds. This completes the proof of Theorem 2.17.

Comments

Section 2.1. Moment assumptions on $Y_{i,j}$ under which the limit function Ψ of Theorem 2.1 is finite were investigated by [Mar04].

Section 2.2. The proof of the explicit limit in Section 2.2 adapts some calculations from [BCS06]. The limit for the geometric case satisfies

$$\Psi(x, y) = m(x + y) + 2\sqrt{\sigma^2 xy}$$

where $m = \mathbf{E}(Y_{1,1})$ and $\sigma^2 = \mathbf{Var}(Y_{1,1})$ are the mean and the variance of the weight distribution. Additional evidence on behalf of this formula comes from asymptotics of $\Psi(x, y)$ at the boundary as $y \searrow 0$ (see [Mar04]), but presently it is not known if this is the correct general limit formula.

Section 2.3. A standard references for the theory of M/M/1 queues is the monograph [Kel79].

Section 2.4. In general, hydrodynamic limits are laws of large numbers that describe the behavior of a particle system over large space and time scales. Introductions to the subject can be found in [DMP91], [KL99] and [Spo91]. The zero range process was introduced by Spitzer [Spi70]. A hydrodynamic limit for multidimensional zero range processes was proved by Rezakhanlou [Rez91]. The envelope approach to hydrodynamic limits used in Section 2.4 works for a special class of totally asymmetric systems in one dimension. It was introduced through several examples in the papers [Sep98], [Sep97] and [Sep00]. Assumption (2.54) was formulated in terms of convergence in probability, and consequently the results (2.58) and (2.59) are of the same type. If we assume almost sure convergence in (2.54) and work harder to derive summable deviation estimates, we can get the conclusions also with probability 1. See [Sep99].

Exercises

EXERCISE 2.1. Explain how the proof of Theorem 2.1 actually proves the stronger statement that there exists an event of probability 1 on which limit (2.2) holds simultaneously for all $(x, y) \in (0, \infty)^2$.

CHAPTER 3

The last-passage Markov chain

Let us recall the setting of the last-passage model with geometric weights. The parameter $0 < p < 1$ is fixed and $q = 1 - p$. The geometric distribution supported by nonnegative integers is

$$\gamma(x) = \begin{cases} pq^x, & x \in \mathbb{Z}_+ \\ 0, & x \in \mathbb{Z} \setminus \mathbb{Z}_+. \end{cases}$$

For $k \in \mathbb{Z}_+$ let $\gamma_k = \gamma^{*k}$ denote the k th convolution power of γ . Then $\gamma_0(x) = \mathbf{1}\{x = 0\}$, $\gamma_1 = \gamma$, and in general for $k \geq 1$ γ_k is the negative binomial distribution:

$$\gamma_k(x) = \gamma^{*k}(x) = \sum_{\substack{(x_1, \dots, x_k) \in \mathbb{Z}^k \\ x_1 + \dots + x_k = x}} \gamma(x_1) \cdots \gamma(x_k) = p^k \binom{x+k-1}{x} q^x, \quad x \in \mathbb{Z}_+.$$

The values at negative integers are all zero: $\gamma_k(x) = 0$ for all $k \in \mathbb{Z}_+$ and $x < 0$. Probabilistically speaking, for $k \geq 1$ γ_k is the probability distribution of a sum $S_k = Y_1 + \dots + Y_k$ of k i.i.d. γ -distributed terms Y_i . If an experiment with success probability p is repeated, then γ_k is the distribution of the number of failures that occur before the k th success. The identity $\gamma_k * \gamma_\ell = \gamma_{k+\ell}$ holds for all $k, \ell \in \mathbb{Z}_+$.

Let $\{Y_{i,j} : (i,j) \in \mathbb{N}^2\}$ be i.i.d. γ -distributed weights or waiting times associated to the points of the positive quadrant \mathbb{N}^2 of the planar integer lattice. For each point $(m,n) \in \mathbb{N}^2$ let $\Pi(m,n)$ be the set of up-right paths

$$(3.1) \quad \pi = \{(1,1) = (i_1, j_1), (i_2, j_2), \dots, (i_{m+n-1}, j_{m+n-1}) = (m,n)\}$$

that connect $(1,1)$ to (m,n) and whose steps are restricted to satisfy $(i_{s+1}, j_{s+1}) - (i_s, j_s) = (1,0)$ or $(0,1)$. Define the last-passage time $G(m,n)$ of point $(m,n) \in \mathbb{N}^2$ by

$$(3.2) \quad G(m,n) = \max_{\pi \in \Pi(m,n)} \sum_{(i,j) \in \pi} Y_{i,j}.$$

An alternative way to express this is

$$(3.3) \quad G(m,n) = \max\{G(m-1,n), G(m,n-1)\} + Y_{m,n}, \quad (m,n) \in \mathbb{N}^2,$$

with boundary conditions $G(m,0) = G(0,n) = 0$ for $m, n \in \mathbb{N}$.

In this section we look at the process $\{G(m,n)\}$ as a Markov chain indexed by m . Fix the vertical dimension n . Define the n -vector $\mathbf{G}(m) = (G(m,1), G(m,2), \dots, G(m,n))$ for $m \in \mathbb{Z}_+$. The initial value is $\mathbf{G}(0) = (0,0, \dots, 0)$. For the state space of the Markov chain $\mathbf{G}(m)$ we take

$$\mathbf{U}_n = \{z \in \mathbb{Z}^n : z_1 \leq z_2 \leq \dots \leq z_n\}.$$

With the initial value $\mathbf{G}(0) = (0,0, \dots, 0)$ the chain $\mathbf{G}(m)$ actually lives in the smaller space $\mathbf{U}_n^+ = \{z \in \mathbb{Z}_+^n : 0 \leq z_1 \leq z_2 \leq \dots \leq z_n\}$. For compact expression of some formulas it is at times convenient to add the coordinate $y_0 = -\infty$ to a vector $y \in \mathbf{U}_n$.

Define the discrete difference operator for functions on \mathbb{Z} by

$$(3.4) \quad Df(x) = f(x+1) - f(x).$$

On the subspace of functions for which there exists $a > -\infty$ such that $f(y) = 0$ for $y \leq a$, define also the operator

$$(3.5) \quad D^{-1}f(x) = \sum_{y:y < x} f(y).$$

On this subspace of functions $D(D^{-1}f) = D^{-1}(Df) = f$. Powers of the operators are defined in the usual way in terms of composition. By definition $D^0f = f$, $D^1f = Df$, and for $j > 0$ $D^{j+1}f = D(D^j f)$ while $D^{-j-1}f = D^{-1}(D^{-j}f)$. Then for all integers $i, j \in \mathbb{Z}$, $D^{i+j}f = D^i(D^j f)$.

Note that in general the inverse of D is not unique, by analogy with indefinite integration. The operator

$$Af(x) = - \sum_{y:y \geq x} f(y)$$

would also serve for functions that vanish at large integers. Definition (3.5) is the one that is useful for this section because the applications are to probability distributions supported on \mathbb{Z}_+ .

The goal of this section is to derive representations for probabilities of the geometric last-passage model. The first theorem gives a determinantal formula for the transition probability.

THEOREM 3.1. *Fix $n \in \mathbb{N}$. The transition probabilities of the Markov chain $\mathbf{G}(m)$ are given by*

$$(3.6) \quad \mathbf{P}\{\mathbf{G}(m) = y \mid \mathbf{G}(\ell) = x\} = \det_{i,j \in [n]} [D^{j-i} \gamma_{m-\ell}(y_j - x_i)],$$

for $x, y \in \mathbf{U}_n$ and $0 \leq \ell \leq m$.

As a corollary we derive a determinantal expression for the distribution function of $G(m, n)$.

THEOREM 3.2. *For $m, n \in \mathbb{N}$ and $t \in \mathbb{Z}$*

$$\mathbf{P}\{G(m, n) \leq t\} = \det_{i,j \in [n]} [D^{j-i-1} \gamma_m(t+1)].$$

If $t < 0$ then the first column ($j = 1$) of the determinant above vanishes and the formula returns 0 as it should.

For the purpose of extracting asymptotics we turn this formula into one of the orthogonal polynomial ensemble type. Our notation for the Vandermonde determinant is $\Delta_n(x) = \prod_{1 \leq i < j \leq n} (x_j - x_i)$. Note that below the summations are not restricted to vectors $x \in \mathbf{U}_n$, although they could be because the functions inside the sums are symmetric in the coordinates (x_1, \dots, x_n) and the Vandermonde vanishes unless the coordinates are distinct.

THEOREM 3.3. *For $m \geq n$ in \mathbb{N} and $t \in \mathbb{Z}_+$*

$$(3.7) \quad \mathbf{P}\{G(m, n) \leq t\} = \frac{1}{Z_{m,n,p}} \sum_{\substack{x \in \mathbb{Z}_+^n \\ \forall i \in [n]: x_i \leq t+n-1}} \Delta_n(x)^2 \prod_{i=1}^n \left\{ \binom{x_i + m - n}{x_i} q^{x_i} \right\}$$

where the normalization constant is given by

$$(3.8) \quad Z_{m,n,p} = \sum_{x \in \mathbb{Z}_+^n} \Delta_n(x)^2 \prod_{i=1}^n \left\{ \binom{x_i + m - n}{x_i} q^{x_i} \right\}.$$

3.1. Derivation of the determinantal transition probabilities

In this section we prove Theorems 3.1 and 3.2. First the transition probability of a single time step. Recall the convention of the extra coordinate $y_0 = -\infty$ for vectors $y = (y_1, \dots, y_n) \in \mathbf{U}_n$.

LEMMA 3.4. *The Markov chain $\mathbf{G}(m)$ has this time-homogeneous one-step transition probability for $x, y \in \mathbf{U}_n$:*

$$(3.9) \quad \mathbf{P}\{\mathbf{G}(m+1) = y \mid \mathbf{G}(m) = x\} = \prod_{k=1}^n \gamma(y_k - \max(x_k, y_{k-1})) = \det_{i,j \in [n]} [D^{j-i} \gamma(y_j - x_i)].$$

The first equality in (3.9) follows from (3.3) but the second needs a proof and we state it as a separate lemma.

LEMMA 3.5. *For $x, y \in \mathbf{U}_n$, with the additional convention $y_0 = -\infty$,*

$$(3.10) \quad \det_{i,j \in [n]} [D^{j-i} \gamma(y_j - x_i)] = \prod_{k=1}^n \gamma(y_k - \max(x_k, y_{k-1})).$$

PROOF. Proof is by induction on n . The case $n = 1$ is clear. Assume (3.10) is true for $n - 1$. Expand the determinant on the left-hand side of line (3.10) along the last row $i = n$:

$$(3.11) \quad \det_{i,j \in [n]} [D^{j-i} \gamma(y_j - x_i)] = \sum_{\ell=1}^{n-2} (-1)^{\ell+n} D^{\ell-n} \gamma(y_\ell - x_n) \det_{i \in [n-1], j \in [n] \setminus \{\ell\}} [D^{j-i} \gamma(y_j - x_i)]$$

$$(3.12) \quad - D^{-1} \gamma(y_{n-1} - x_n) \det_{i \in [n-1], j \in [n] \setminus \{n-1\}} [D^{j-i} \gamma(y_j - x_i)]$$

$$(3.13) \quad + \gamma(y_n - x_n) \det_{i,j \in [n-1]} [D^{j-i} \gamma(y_j - x_i)].$$

We show that each term in the sum on the right-hand side of line (3.11) vanishes. Fix an index $1 \leq \ell \leq n - 2$ for the moment.

Case 1: $y_\ell \leq x_n$. Check by induction for $j < 0$ and $j > 0$ that

$$(3.14) \quad \forall j \in \mathbb{Z}: D^j f(x) \text{ is a linear combination of } \{f(y) : y \leq x + j\}.$$

Hence since $\gamma(y) = 0$ for $y < 0$, $D^j \gamma(x) = 0$ for $x < -j$. It follows that $D^{\ell-n} \gamma(y_\ell - x_n) = 0$.

Case 2: $y_\ell > x_n$. Let us write D_y when the difference operator acts on the variable y , so that for $k \in \mathbb{Z}$

$$(3.15) \quad D^{k+1} f(x + y) = D^k f(x + y + 1) - D^k f(x + y) = D_y D^k f(x + y).$$

This is a convenient notational trick. Set temporarily

$$\tilde{y}_j = \begin{cases} y_j, & 1 \leq j \leq \ell - 1 \\ y_{j+1}, & \ell \leq j \leq n - 1. \end{cases}$$

In the first step below apply (3.15) in the columns $j = \ell + 1, \dots, n$. Use the linearity of the determinant in the columns to bring the operators D_{y_j} outside. The effect is to reduce the powers of D by 1 and thereby bring the index j to the interval $[n - 1]$:

$$\det_{i \in [n-1], j \in [n] \setminus \{\ell\}} [D^{j-i} \gamma(y_j - x_i)] = D_{y_{\ell+1}} \cdots D_{y_n} \det_{i,j \in [n-1]} [D^{j-i} \gamma(\tilde{y}_j - x_i)]$$

by the induction assumption that (3.10) holds for $n - 1$

$$\begin{aligned}
&= D_{y_{\ell+1}} \cdots D_{y_n} \left\{ \prod_{k=1}^{n-1} \gamma(\tilde{y}_k - \max(x_k, \tilde{y}_{k-1})) \right\} \\
&= D_{y_{\ell+1}} \cdots D_{y_n} \left\{ \prod_{k=1}^{\ell-1} \gamma(y_k - \max(x_k, y_{k-1})) \cdot \gamma(y_{\ell+1} - \max(x_\ell, y_{\ell-1})) \right. \\
&\quad \left. \times \prod_{k=\ell+1}^{n-1} \gamma(y_{k+1} - \max(x_k, y_k)) \right\}
\end{aligned}$$

utilizing the assumption $y_n \geq \cdots \geq y_\ell > x_n$

$$\begin{aligned}
&= D_{y_{\ell+1}} \cdots D_{y_n} \left\{ \prod_{k=1}^{\ell-1} \gamma(y_k - \max(x_k, y_{k-1})) \cdot pq^{y_{\ell+1} - \max(x_\ell, y_{\ell-1})} \right. \\
&\quad \left. \times \prod_{k=\ell+1}^{n-1} pq^{y_{k+1} - y_k} \right\} \\
&= \prod_{k=1}^{\ell-1} \gamma(y_k - \max(x_k, y_{k-1})) pq^{-\max(x_\ell, y_{\ell-1})} \\
&\quad \times D_{y_{\ell+1}} \cdots D_{y_n} \left\{ q^{y_{\ell+1}} \cdot \prod_{k=\ell+1}^{n-1} pq^{y_{k+1} - y_k} \right\}.
\end{aligned}$$

The last factor equals

$$p^{n-\ell-1} D_{y_{\ell+1}} \cdots D_{y_n} q^{y_n} = 0$$

because the function q^{y_n} is constant in each y_j such that $\ell + 1 \leq j \leq n - 1$, and this range of indices is nonempty because we are presently in the case $1 \leq \ell \leq n - 2$.

We have now verified that each term $1 \leq \ell \leq n - 2$ in the sum on the right-hand side of line (3.11) vanishes.

Next we show that lines (3.12) and (3.13) together make up the right-hand side of (3.10). This will complete the proof of the lemma.

Case 1: $y_{n-1} \leq x_n$. By (3.14) $D^{-1}\gamma(y_{n-1} - x_n) = 0$. Only line (3.13) remains, which by $y_{n-1} \leq x_n$ and induction equals the left-hand side of (3.10).

Case 2: $y_{n-1} > x_n$. Note first that for $x \geq 0$,

$$D\gamma(x) = pq^{x+1} - pq^x = -p\gamma(x),$$

and by induction for all $j \geq 1$ and $x \geq 0$

$$D^j\gamma(x) = -p^j\gamma(x) = -pD^{j-1}\gamma(x).$$

Consider the determinant $\det_{i \in [n-1], j \in [n] \setminus \{n-1\}} [D^{j-i}\gamma(y_j - x_i)]$ on line (3.12). Since $y_n \geq y_{n-1} > x_n \geq x_i$ is assumed, we can write the last column with index $j = n$ in this determinant as

$$\{D^{n-i}\gamma(y_n - x_i)\}_{i \in [n-1]} = -p\{D^{n-1-i}\gamma(y_n - x_i)\}_{i \in [n-1]}.$$

Take the factor $-p$ outside the determinant by linearity, and apply the induction assumption to write the determinant as

$$(3.16) \quad \begin{aligned} & -p \prod_{k=1}^{n-2} \gamma(y_k - \max(x_k, y_{k-1})) \cdot \gamma(y_n - \max(x_{n-1}, y_{n-2})) \\ & = \prod_{k=1}^{n-1} \gamma(y_k - \max(x_k, y_{k-1})) \cdot (-p)q^{y_n - y_{n-1}}. \end{aligned}$$

Noting that $D^{-1}\gamma(x) = 1 - q^x$ for $x > 0$, we write line (3.12) together with its minus sign as

$$\prod_{k=1}^{n-1} \gamma(y_k - \max(x_k, y_{k-1})) \cdot (1 - q^{y_{n-1} - x_n})pq^{y_n - y_{n-1}}.$$

Apply induction on line (3.13) and add it together with the line above to get this expression for the sum of lines (3.12) and (3.13):

$$\begin{aligned} & \prod_{k=1}^{n-1} \gamma(y_k - \max(x_k, y_{k-1})) \cdot \left((1 - q^{y_{n-1} - x_n})pq^{y_n - y_{n-1}} + pq^{y_n - x_n} \right) \\ & = \prod_{k=1}^{n-1} \gamma(y_k - \max(x_k, y_{k-1})) \cdot pq^{y_n - y_{n-1}} = \prod_{k=1}^n \gamma(y_k - \max(x_k, y_{k-1})). \end{aligned}$$

The last equality used $y_{n-1} > x_n$. To summarize, in both cases lines (3.12) and (3.13) together make up the right-hand side of (3.10). This completes the proof of the lemma. \square

This finishes the proof of formula (3.9) for the single step transition. Next a convolution identity.

PROPOSITION 3.6. *Let $f, g : \mathbb{Z} \rightarrow \mathbb{C}$ and $a \in \mathbb{Z}$ be such that $f(x) = g(x) = 0$ for $x < a$. Then the following identity holds for $n \times n$ determinants with $i, j \in [n]$ and for all integer vectors $(x_1, \dots, x_n), (z_1, \dots, z_n) \in \mathbb{Z}^n$:*

$$(3.17) \quad \sum_{y \in \mathbf{U}_n} \det[D^{j-i}f(y_j - x_i)] \det[D^{j-i}g(z_j - y_i)] = \det[D^{j-i}(f * g)(z_j - x_i)].$$

The sum on the left in (3.17) is actually finite because large negative y_j -values eliminate the f -values while large positive y_i -values eliminate the g -values. The proof of Proposition 3.6 depends on the next lemma which sets the stage for an application of the Cauchy-Binet identity (B.7).

LEMMA 3.7. *With assumptions as in Proposition 3.6,*

$$(3.18) \quad \begin{aligned} & \sum_{y \in \mathbf{U}_n} \det[D^{j-i}f(y_j - x_i)] \det[D^{j-i}g(z_j - y_i)] \\ & = \sum_{y \in \mathbf{U}_n} \det[D^{1-i}f(y_j - x_i)] \det[D^{i-1}g(z_i - y_j)]. \end{aligned}$$

PROOF. Note that the second g -determinant in (3.18) was transposed. Identity (3.18) is proved by repeatedly moving operators D from columns of the f -determinant to corresponding columns of the g -determinant via (3.19).

First check this summation by parts formula by induction on $\ell - k$: for any \mathbb{C} -valued functions ϕ and ψ on \mathbb{Z} , any integers x, z and $k < \ell$,

$$(3.19) \quad \sum_{y=k}^{\ell} D\phi(y-x)\psi(z-y) = \sum_{y=k}^{\ell} \phi(y-x)D\psi(z-y) + \phi(\ell+1-x)\psi(z-\ell) - \phi(k-x)\psi(z+1-k).$$

You do not see a minus sign in front of the sum on the right because the summation variable y comes with a minus sign.

Beginning with the left-hand side of (3.18), transpose the g -determinant, express the determinants in terms of columns, and write the summation index as $y = (y', y_n)$ with $y' = (y_1, \dots, y_{n-1}) \in U_{n-1}$:

$$(3.20) \quad \sum_{y' \in U_{n-1}} \sum_{y_n=y_{n-1}}^M \det[D^{1-i}f(y_1-x_i), \dots, D^{n-1-i}f(y_{n-1}-x_i), D^{n-i}f(y_n-x_i)]$$

$$(3.21) \quad \times \det[D^{i-1}g(z_i-y_1), \dots, D^{i-n+1}g(z_i-y_{n-1}), D^{i-n}g(z_i-y_n)].$$

Upper summation limit M above was chosen large enough so that $z_i - M < a$ for each i to guarantee that each $g(z_i - y_n) = 0$ for $y_n > M$.

Apply summation by parts (3.19) to the inner sum over y_n . This takes one operator D from the last column of the determinant on line (3.20), leaving $D^{n-1-i}f(y_n-x_i)$, and puts this D in the last column of the determinant on line (3.21), turning this column into $D^{i-n+1}g(z_i-y_n)$. (More precisely, first the D comes out of the f -determinant by linearity of determinant in columns, then moves in front of the g -determinant by summation by parts, and then slips into the last column of the g -determinant.) The boundary terms are (for fixed y')

$$(3.22) \quad \det[D^{1-i}f(y_1-x_i), \dots, D^{n-1-i}f(y_{n-1}-x_i), D^{n-1-i}f(M+1-x_i)] \\ \times \det[D^{i-1}g(z_i-y_1), \dots, D^{i-n+1}g(z_i-y_{n-1}), D^{i-n}g(z_i-M)]$$

$$(3.23) \quad - \det[D^{1-i}f(y_1-x_i), \dots, D^{n-1-i}f(y_{n-1}-x_i), D^{n-1-i}f(y_{n-1}-x_i)] \\ \times \det[D^{i-1}g(z_i-y_1), \dots, D^{i-n+1}g(z_i-y_{n-1}), D^{i-n}g(z_i+1-y_{n-1})]$$

The boundary terms vanish, (3.22) because M was chosen large enough and (3.23) because the two last columns of the f -determinant are identical. The result is that the sum on lines (3.20)–(3.21) has become

$$(3.24) \quad \sum_{y \in U_n} \det[D^{1-i}f(y_1-x_i), \dots, D^{n-1-i}f(y_{n-1}-x_i), D^{n-1-i}f(y_n-x_i)]$$

$$(3.25) \quad \times \det[D^{i-1}g(z_i-y_1), \dots, D^{i-n+1}g(z_i-y_{n-1}), D^{i-n+1}g(z_i-y_n)].$$

Next we repeat the procedure for y_{n-1} . With $y' = (y_1, \dots, y_{n-2}, y_n)$ fixed the summation is $\sum_{y_{n-1}=y_{n-2}}^{y_n}$. One D operator moves from the next-to-last column of the f -determinant on line (3.24) to the next-to-last column of the g -determinant on line (3.25). The boundary terms coming from the summation by parts are

$$\det[D^{1-i}f(y_1-x_i), \dots, D^{n-2-i}f(y_{n-2}-x_i), D^{n-2-i}f(y_n+1-x_i), D^{n-1-i}f(y_n-x_i)] \\ \times \det[D^{i-1}g(z_i-y_1), \dots, D^{i-n+1}g(z_i-y_n), D^{i-n+1}g(z_i-y_n)] \\ - \det[D^{1-i}f(y_1-x_i), \dots, D^{n-2-i}f(y_{n-2}-x_i), D^{n-2-i}f(y_{n-2}-x_i), D^{n-1-i}f(y_n-x_i)] \\ \times \det[D^{i-1}g(z_i-y_1), \dots, D^{i-n+1}g(z_i+1-y_{n-2}), D^{i-n+1}g(z_i-y_n)]$$

These vanish because determinants with repeated columns appear.

When this step has been done once for each of y_n, y_{n-1}, \dots, y_2 sum on lines (3.24)–(3.25) has become

$$(3.26) \quad \sum_{y \in \mathbf{U}_n} \det[D^{1-i}f(y_1 - x_i), D^{1-i}f(y_2 - x_i), D^{2-i}f(y_3 - x_i), \dots, D^{n-1-i}f(y_n - x_i)]$$

$$(3.27) \quad \times \det[D^{i-1}g(z_i - y_1), D^{i-1}g(z_i - y_2), D^{i-2}g(z_i - y_3), \dots, D^{i-n+1}g(z_i - y_n)].$$

In both determinants the first two columns are in agreement with the goal which is the right-hand side of (3.18). This procedure is next repeated for y_n, y_{n-1}, \dots, y_3 after which the first three columns are done. And so on, until the sum has turned into the right-hand side of (3.18). \square

PROOF OF PROPOSITION 3.6. Start with the left-hand side of (3.17). Apply identity (3.18):

$$\begin{aligned} & \sum_{y \in \mathbf{U}_n} \det[D^{j-i}f(y_j - x_i)] \det[D^{j-i}g(z_j - y_i)] \\ &= \sum_{y \in \mathbf{U}_n} \det[D^{1-i}f(y_j - x_i)] \det[D^{i-1}g(z_i - y_j)] \end{aligned}$$

by symmetry and vanishing of determinants with identical columns

$$= \frac{1}{n!} \sum_{y \in \mathbb{Z}^n} \det[D^{1-i}f(y_j - x_i)] \det[D^{i-1}g(z_i - y_j)]$$

by the generalized Cauchy-Binet identity (B.7)

$$(3.28) \quad = \det \left[\sum_{y \in \mathbb{Z}} D^{1-i}f(y - x_i) D^{j-1}g(z_j - y) \right].$$

Thus it remains to check that for $i, j \geq 1$

$$(3.29) \quad \sum_{y \in \mathbb{Z}} D^{1-i}f(y - x) D^{j-1}g(z - y) = D^{j-i}(f * g)(z - x).$$

The sum actually ranges over a finite interval of integers. Let D_z act on the z -variable. Note that

$$(3.30) \quad (f * \mathbf{1}_{\mathbb{N}})(x) = \sum_{y \in \mathbb{Z}} f(x - y) \mathbf{1}_{\mathbb{N}}(y) = \sum_{z: z < x} f(z) = D^{-1}f(x).$$

Use linearity of operators and associativity of convolution:

$$\begin{aligned} \sum_{y \in \mathbb{Z}} D^{1-i}f(y - x) D^{j-1}g(z - y) &= \sum_{y \in \mathbb{Z}} (\mathbf{1}_{\mathbb{N}}^{*(i-1)} * f)(y - x) D_z^{j-1}g(z - y) \\ &= D_z^{j-1}((\mathbf{1}_{\mathbb{N}}^{*(i-1)} * f) * g)(z - x) = D_z^{j-1}(\mathbf{1}_{\mathbb{N}}^{*(i-1)} * (f * g))(z - x) \\ &= D_z^{j-1}D^{1-i}(f * g)(z - x) = D^{j-i}(f * g)(z - x). \end{aligned}$$

This completes the proof of Proposition 3.6 \square

PROOF OF THEOREM 3.1. The case $m = \ell + 1$ is in (3.9).

The general case follows from an inductive argument. Suppose formula (3.6) has been proved for time steps of size m_1 and m_2 . Use the Markov property, the convolution formula (3.17), and the

identity $\gamma_j * \gamma_k = \gamma_{j+k}$. The determinants are still $n \times n$. Let $x, y \in \mathbf{U}_n$.

$$\begin{aligned} & \mathbf{P}\{\mathbf{G}(m + m_1 + m_2) = y \mid \mathbf{G}(m) = x\} \\ &= \sum_{z \in U_n} \mathbf{P}\{\mathbf{G}(m + m_1 + m_2) = y \mid \mathbf{G}(m + m_1) = z\} \mathbf{P}\{\mathbf{G}(m + m_1) = z \mid \mathbf{G}(m) = x\} \\ &= \sum_{z \in U_n} \det[D^{j-i} \gamma_{m_2}(y_j - z_i)] \det[D^{j-i} \gamma_{m_1}(z_j - x_i)] = \det[D^{j-i} \gamma_{m_1+m_2}(y_j - x_i)]. \quad \square \end{aligned}$$

PROOF OF THEOREM 3.2. We need to verify

$$(3.31) \quad \mathbf{P}\{G(m, n) \leq t\} = \det_{i,j \in [n]} [D^{j-i-1} \gamma_m(t+1)].$$

Use formula (3.6) for the transition probabilities and then take sums inside to the relevant columns.

$$\begin{aligned} \mathbf{P}\{G(m, n) \leq t\} &= \sum_{x: x_1 \leq \dots \leq x_n \leq t} \mathbf{P}\{\mathbf{G}(m) = x\} = \sum_{x: x_1 \leq \dots \leq x_n \leq t} \det[D^{j-i} \gamma_m(x_j)] \\ &= \sum_{x_1 \leq \dots \leq x_{n-1} \leq t} \sum_{x_n = x_{n-1}}^t \det[D^{j-i} \gamma_m(x_j)] \\ (3.32) \quad &= \sum_{x_1 \leq \dots \leq x_{n-1} \leq t} \det \left[D^{1-i} \gamma_m(x_1), \dots, D^{n-1-i} \gamma_m(x_{n-1}), \sum_{x_n = x_{n-1}}^t D^{n-i} \gamma_m(x_n) \right]. \end{aligned}$$

In the last column use one D to create a telescoping sum:

$$(3.33) \quad \sum_{x_n = x_{n-1}}^t D^{n-i} \gamma_m(x_n) = \sum_{x_n = x_{n-1}}^t (D^{n-1-i} \gamma_m(x_n + 1) - D^{n-1-i} \gamma_m(x_n)) \\ = D^{n-1-i} \gamma_m(t+1) - D^{n-1-i} \gamma_m(x_{n-1}).$$

Substitute this back into (3.32) and notice that the last term above is identical with the second last column in (3.32). Since a determinant with two identical columns vanishes, line (3.32) becomes

$$\sum_{x_1 \leq \dots \leq x_{n-1} \leq t} \det \left[D^{1-i} \gamma_m(x_1), \dots, D^{n-1-i} \gamma_m(x_{n-1}), D^{n-1-i} \gamma_m(t+1) \right].$$

Repeat this step for x_{n-1}, \dots, x_2 , each time letting identical columns eliminate one of the resulting terms, to turn the above into

$$\sum_{-\infty < x_1 \leq t} \det \left[D^{1-i} \gamma_m(x_1), D^{1-i} \gamma_m(t+1), \dots, D^{n-2-i} \gamma_m(t+1), D^{n-1-i} \gamma_m(t+1) \right].$$

Lastly, take the remaining sum into the first column as in (3.33) and use the definition of D^{-1} :

$$\sum_{x_1 = -\infty}^t D^{1-i} \gamma_m(x_1) = D^{-1} (D^{1-i} \gamma_m)(t+1) = D^{-i} \gamma_m(t+1).$$

Substituting the above for the first column completes the transformation of line (3.32) into the right-hand side of (3.31). \square

3.2. From determinantal formula to orthogonal polynomial ensemble

To prove Theorem 3.3 we prove this proposition.

PROPOSITION 3.8. *There exists a finite constant $C_{m,n,p}$ such that for integers $m \geq n \geq 1$ and $t \in \mathbb{Z}$,*

$$(3.34) \quad \det_{i,j \in [n]} [D^{j-i-1} \gamma_m(t+1)] = C_{m,n,p} \sum_{\substack{x \in \mathbb{Z}_+^n \\ \forall i \in [n]: x_i \leq t+n-1}} \Delta_n(x)^2 \prod_{i=1}^n \binom{x_i + m - n}{x_i} q^{x_i}.$$

We can observe already that both sides of (3.34) vanish if $t < 0$. On the left, the values $D^{-i} \gamma_m(t+1)$ in column $j = 1$ vanish by (3.14). On the right $\Delta_n(x)$ vanishes if x has two identical coordinates, hence if there are fewer than n distinct values available for the x_i 's.

This section is devoted to the proof of the proposition and at the end we derive Theorem 3.3 from it. For $a \in \mathbb{R}$ denote the descending factorials by $(a)_{[0]} = 1$ and for $n \in \mathbb{N}$, $(a)_{[n]} = a(a-1)(a-2) \cdots (a-n+1)$. For $n \in \mathbb{Z}_+$, $(n)_{[n]} = n!$.

LEMMA 3.9. *For $x \in \mathbb{Z}$ and $n \in \mathbb{N}$ the convolution powers of $\mathbf{1}_{\mathbb{N}}$ satisfy*

$$(3.35) \quad \mathbf{1}_{\mathbb{N}}^{*n}(x) = \frac{(x-1)_{[n-1]}}{(n-1)!} \mathbf{1}\{x \geq n\}.$$

PROOF. First we fix $n \in \mathbb{N}$ and prove by induction on $x \geq n+1$ that

$$(3.36) \quad \sum_{y=1}^{x-n} (x-y-1)_{[n-1]} = \frac{(x-1)_{[n]}}{n}.$$

The case $x = n+1$ simplifies to the identity $(n-1)_{[n-1]} = n^{-1}(n)_{[n]}$. Assume (3.36) is true for x . Then, by a change of summation index $y = z+1$,

$$\sum_{y=1}^{x+1-n} (x-y)_{[n-1]} = \sum_{z=0}^{x-n} (x-z-1)_{[n-1]} = (x-1)_{[n-1]} + \frac{(x-1)_{[n]}}{n} = \frac{(x)_{[n]}}{n}.$$

Turn to (3.35) with induction on n . The case $n = 1$ is clear. Assuming (3.35) is true for n ,

$$\mathbf{1}_{\mathbb{N}}^{*(n+1)}(x) = \sum_{y \in \mathbb{Z}} \mathbf{1}_{\mathbb{N}}^{*n}(x-y) \mathbf{1}_{\mathbb{N}}(y) = \sum_{y=1}^{x-n} \frac{(x-y-1)_{[n-1]}}{(n-1)!} = \frac{(x-1)_{[n]}}{n!} \mathbf{1}\{x \geq n+1\}$$

with the last equality from (3.36). Note that the last sum above is zero if $x \leq n$ by the convention on empty sums. \square

Recall from (3.30) that $D^{-k}f = (\mathbf{1}_{\mathbb{N}}^{*k}) * f$ for $k \geq 0$. (The convention for convolution is that $g^{*0} = \mathbf{1}_{\{0\}}$ for any function g so that $g^{*0} * f = f$.)

$$\begin{aligned} D^{j-i-1} \gamma_m(t+1) &= D^{-i}(D^{j-1} \gamma_m)(t+1) = \sum_{y \in \mathbb{Z}} \mathbf{1}_{\mathbb{N}}^{*i}(t+1-y)(D^{j-1} \gamma_m)(y) \\ &= \sum_{y \in \mathbb{Z}} \frac{(t-y)_{[i-1]}}{(i-1)!} \mathbf{1}\{y \leq t+1-i\} (D^{j-1} \gamma_m)(y) \\ &= \sum_{y=-n+1}^t \frac{(t-y)_{[i-1]}}{(i-1)!} (D^{j-1} \gamma_m)(y) \end{aligned}$$

The summation limits on the last line come as follows. The condition $y \leq t + 1 - i$ can be replaced by $y \leq t$ due to the vanishing of the numerator $(t - y)_{[i-1]}$ for the added y -values. For $1 \leq j \leq n$, $D^{j-1}\gamma_m(y) = 0$ for $y \leq -n$.

Now consider the determinant

$$\det_{i,j \in [n]} [D^{j-i-1}\gamma_m(t+1)] = \det_{i,j \in [n]} \left[\sum_{y=-n+1}^t \frac{(t-y)_{[i-1]}}{(i-1)!} (D^{j-1}\gamma_m)(y) \right].$$

The expression $(t-y)_{[i-1]}/(i-1)!$ in row i is a polynomial in y of precise degree $i-1$ with leading term $(-1)^{i-1}y^{i-1}/(i-1)!$. By adding suitable multiples of rows $1, \dots, i-1$ to row i we can convert this polynomial in row i into $(-1)^{i-1}(y+n-1)^{i-1}/(i-1)!$ (same leading term). Row 1 is not affected. These row operations preserve the determinant, and so

$$\begin{aligned} \det_{i,j \in [n]} [D^{j-i-1}\gamma_m(t+1)] &= \det_{i,j \in [n]} \left[\sum_{y=-n+1}^t \frac{(y+n-1)^{i-1}}{(i-1)!} (-1)^{i-1} (D^{j-1}\gamma_m)(y) \right] \\ &= \det_{i,j \in [n]} \left[\sum_{y=0}^{t+n-1} \frac{y^{i-1}}{(i-1)!} (-1)^{i-1} (D^{j-1}\gamma_m)(y-n+1) \right] \end{aligned}$$

by the simple observation $\det[(-1)^i a_{i,j}] = \det[(-1)^j a_{i,j}]$

$$= \det_{i,j \in [n]} \left[\sum_{y=0}^{t+n-1} \frac{y^{i-1}}{(i-1)!} (-1)^{j-1} (D^{j-1}\gamma_m)(y-n+1) \right]$$

by the generalized Cauchy-Binet identity

$$= \frac{1}{n!} \sum_{y_1, \dots, y_n=0}^{t+n-1} \det_{i,j \in [n]} \left[\frac{y_j^{i-1}}{(i-1)!} \right] \det_{i,j \in [n]} \left[(-1)^{i-1} (D^{i-1}\gamma_m)(y_j-n+1) \right]$$

and by identifying the $n \times n$ Vandermonde

$$(3.37) \quad = \left(\prod_{k=1}^n \frac{1}{k!} \right) \sum_{y_1, \dots, y_n=0}^{t+n-1} \Delta_n(y) \det_{i,j \in [n]} \left[(-1)^{i-1} (D^{i-1}\gamma_m)(y_j-n+1) \right].$$

A couple more transformations are needed on this last determinant.

LEMMA 3.10. *Fix now $m \geq n \geq 1$. There exist polynomials g_k of degree k such that for integers $x \geq -m+1$ and $0 \leq k < n$*

$$(3.38) \quad (-1)^k D^k \gamma_m(x) = q^x g_k(x) \prod_{\ell=k+1}^{m-1} (x+\ell).$$

The fixed parameters p and m appear in g_k . If $k = m-1$ the empty product equals 1.

PROOF. Induction on k . First $k = 0$. For $x \in \mathbb{Z}_+$

$$\gamma_m(x) = p^m \binom{x+m-1}{m-1} q^x = q^x \frac{p^m}{(m-1)!} \prod_{\ell=1}^{m-1} (x+\ell).$$

For integers $x \in \{-m+1, \dots, -1\}$ both sides vanish.

Assume (3.38) is valid for $k - 1 \in \{0, \dots, n - 2\}$.

$$\begin{aligned}
(-1)^k D^k \gamma_m(x) &= -(-1)^{k-1} D^{k-1} \gamma_m(x+1) + (-1)^{k-1} D^{k-1} \gamma_m(x) \\
&= -q^{x+1} g_{k-1}(x+1) \prod_{\ell=k}^{m-1} (x+1+\ell) + q^x g_{k-1}(x) \prod_{\ell=k}^{m-1} (x+\ell) \\
&= q^x \left\{ -q g_{k-1}(x+1) \cdot (x+m) + g_{k-1}(x) \cdot (x+k) \right\} \prod_{\ell=k+1}^{m-1} (x+\ell) \\
&\equiv q^x g_k(x) \prod_{\ell=k+1}^{m-1} (x+\ell)
\end{aligned}$$

where the last line defines the polynomial g_k of degree k . \square

Continuing from (3.37) we can write

$$\begin{aligned}
&\det_{i,j \in [n]} [D^{j-i-1} \gamma_m(t+1)] \\
&= \left(\prod_{k=1}^n \frac{1}{k!} \right) \sum_{y_1, \dots, y_n=0}^{t+n-1} \Delta_n(y) \det_{i,j \in [n]} \left[q^{y_j - n + 1} g_{i-1}(y_j - n + 1) \prod_{\ell=i}^{m-1} (y_j - n + 1 + \ell) \right] \\
&= C_{n,p} \sum_{y_1, \dots, y_n=0}^{t+n-1} \Delta_n(y) \prod_{j=1}^n \left\{ q^{y_j} \prod_{\ell=n+1}^m (y_j - n + \ell) \right\} \\
(3.39) \quad &\times \det_{i,j \in [n]} \left[g_{i-1}(y_j - n + 1) \prod_{\ell=i+1}^n (y_j - n + \ell) \right].
\end{aligned}$$

In the last equality we removed the q -factors and part of the product from each column in the determinant. All factors independent of y are collected into the constant $C_{n,p}$, which, as the notation suggests, is some function of n and q . We left some factors $y_j - n + \ell$ inside the determinant to take advantage of the next lemma.

LEMMA 3.11. *Let p_k be a polynomial of degree k for $k = 0, \dots, n - 1$ and let c_2, \dots, c_n be complex constants. Then there exists a constant A such that this identity holds for all complex vectors $x = (x_1, \dots, x_n)$:*

$$(3.40) \quad \det_{i,j \in [n]} \left[p_{i-1}(x_j) \prod_{\ell=i+1}^n (x_j + c_\ell) \right] = A \Delta_n(x).$$

For $i = n$ the value of the empty product $\prod_{\ell=i+1}^n$ is 1. Constant A depends on n , the coefficients in the p_k 's and the constants c_ℓ .

PROOF. The i, j entry in the determinant on the left-hand side of (3.40) is a polynomial of degree $n - 1$ in the variable x_j . If we write this entry in terms of coefficients as

$$p_{i-1}(x_j) \prod_{\ell=i+1}^n (x_j + c_\ell) = \sum_{k=0}^{n-1} a_{i,k} x_j^k$$

then we have the matrix product

$$\left(p_{i-1}(x_j) \prod_{\ell=i+1}^n (x_j + c_\ell) \right)_{i,j \in [n]} = (a_{i,j-1})_{i,j \in [n]} (x_j^{i-1})_{i,j \in [n]}.$$

Since determinants multiply, we have (3.40) with $A = \det(a_{i,j-1})$. \square

Apply (3.40) with $x_j = y_j$, $p_k(x) = g_k(x - n + 1)$ and $c_\ell = -n + \ell$. Then continuing from (3.39) we have, with a new constant $C_{n,p}$,

$$\begin{aligned}
& \det_{i,j \in [n]} [D^{j-i-1} \gamma_m(t+1)] \\
&= C_{n,p} \sum_{y_1, \dots, y_n=0}^{t+n-1} \Delta_n(y)^2 \prod_{j=1}^n \left\{ q^{y_j} \prod_{\ell=n+1}^m (y_j - n + \ell) \right\} \\
&= C_{n,p} \sum_{y_1, \dots, y_n=0}^{t+n-1} \Delta_n(y)^2 \prod_{j=1}^n q^{y_j} \frac{(y_j - n + m)!}{y_j!} \\
(3.41) \quad &= C_{m,n,p} \sum_{y_1, \dots, y_n=0}^{t+n-1} \Delta_n(y)^2 \prod_{j=1}^n \binom{y_j + m - n}{y_j} q^{y_j}.
\end{aligned}$$

In the last equality the expression was multiplied by $(m-n)!/(m-n)!$ and thereby the constant acquired a dependence on m also. This completes the proof of Proposition 3.8.

To prove Theorem 3.3 it remains to show that

$$(3.42) \quad 1 = C_{m,n,p} \sum_{y_1, \dots, y_n=0}^{\infty} \Delta_n(y)^2 \prod_{j=1}^n \binom{y_j + m - n}{y_j} q^{y_j}.$$

By the display that ended in (3.41) this follows from showing that $\det_{i,j \in [n]} [D^{j-i-1} \gamma_m(t+1)] \rightarrow 1$ as $t \rightarrow \infty$. Two observations imply this. (i) The diagonal elements in the determinant tend to 1: for $i = j$,

$$D^{j-i-1} \gamma_m(t+1) = D^{-1} \gamma_m(t+1) = \sum_{y: y < t+1} \gamma_m(y) \rightarrow 1 \quad \text{as } t \rightarrow \infty$$

since probabilities sum to 1. (ii) Elements above the diagonal ($j > i$) tend to 0 because $D^k \gamma_m(x) \rightarrow 0$ as $x \rightarrow \infty$ for $k \geq 0$. For $k = 0$ it is simply $D^0 \gamma_m(x) = \gamma_m(x) \rightarrow 0$. Assuming it is true for k , then for $k+1$

$$D^{k+1} \gamma_m(x) = D^k \gamma_m(x+1) - D^k \gamma_m(x) \rightarrow 0 \quad \text{as } x \rightarrow \infty.$$

Alternately, we can use the probabilistic connection from Theorem 3.2:

$$\lim_{t \rightarrow \infty} \det_{i,j \in [n]} [D^{j-i-1} \gamma_m(t+1)] = \lim_{t \rightarrow \infty} \mathbf{P}\{G(m, n) \leq t\} = 1.$$

In either case, we have verified formula (3.8) and completed the proof of Theorem 3.3.

Comments

This section follows Johansson's paper [Joh07].

Tracy-Widom distribution

4.1. Airy function and kernel

The Airy function of a complex argument z is defined by

$$(4.1) \quad \text{Ai}(z) = \frac{1}{2\pi i} \int_{\mathcal{C}} e^{\frac{1}{3}\zeta^3 - z\zeta} d\zeta$$

where the contour \mathcal{C} begins at a point “at infinity” in the sector $-\pi/2 + \delta \leq \arg \zeta \leq -\pi/6 - \delta$ and ends at a point at infinity in the conjugate sector $\pi/6 + \delta \leq \arg \zeta \leq \pi/2 - \delta$. For example, one could take the ray from $\infty e^{-\pi/3}$ to the origin together with the ray from the origin to $\infty e^{\pi/3}$.

We shall not verify that integral (4.1) is well-defined. This will be a consequence of the next argument that converts (4.1) to an integral over a path \mathcal{P}_s parallel to and to the right of the imaginary axis. Absolute convergence of this second integral will be easy to see. Path \mathcal{P}_s is given as $\zeta(t) = s + it$, $t \in (-\infty, \infty)$, where $s > 0$ is fixed. To show the equality

$$(4.2) \quad \int_{\mathcal{C}} e^{\frac{1}{3}\zeta^3 - z\zeta} d\zeta = \int_{\mathcal{P}_s} e^{\frac{1}{3}\zeta^3 - z\zeta} d\zeta$$

of the integrals, we show that integrals over horizontal “bridges” that connect the paths at imaginary levels it vanish as $|t| \rightarrow \infty$. For $c \geq s|t|^{-1}$ let \mathcal{L}_t be the horizontal line segment from $s + it$ to $c|t| + it$. We claim that

$$(4.3) \quad \sup_{s|t|^{-1} \leq c \leq \sqrt{3-|t|^{-1/2}}} \left| \int_{\mathcal{L}_t} e^{\frac{1}{3}\zeta^3 - z\zeta} d\zeta \right| \rightarrow 0 \quad \text{as } |t| \rightarrow \infty.$$

The upper bound of $\sqrt{3}$ minus a little for c comes from the angle $\pi/6$. Parametrize \mathcal{L}_t by $\zeta(u) = u + it$, write $z = x + iy$, and consider $|t|$ large enough so that $|t|^{3/2}/3 + x > 0$. Then, since $u^2 \leq c^2 t^2 \leq t^2(3 - |t|^{-1/2})$,

$$\begin{aligned} \left| \int_{\mathcal{L}_t} e^{\frac{1}{3}\zeta^3 - z\zeta} d\zeta \right| &\leq \int_s^{c|t|} \left| e^{\frac{1}{3}(u+it)^3 - z(u+it)} \right| du = e^{yt} \int_s^{c|t|} e^{u^3/3 - ut^2 - xu} du \\ &\leq e^{yt} \int_s^{c|t|} e^{-u(|t|^{3/2}/3 + x)} du \leq \frac{e^{yt - s(|t|^{3/2}/3 + x)}}{|t|^{3/2}/3 + x} \rightarrow 0 \quad \text{as } |t| \rightarrow \infty. \end{aligned}$$

This justifies (4.2), and we have the following representation for the Airy function: for an arbitrary fixed $s > 0$

$$(4.4) \quad \begin{aligned} \text{Ai}(z) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{\frac{1}{3}(s+it)^3 - z(s+it)} dt \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{s^3/3 - st^2 - xs + yt + i(s^2 t - t^3/3 - xt - ys)} dt. \end{aligned}$$

The modulus of the integrand in (4.4) is $e^{s^3/3 - st^2 - xs + yt}$. This is clearly integrable over $t \in (-\infty, \infty)$ for any z and decays faster than any polynomial. Hence dominated convergence justifies

differentiation inside the integral:

$$(4.5) \quad \text{Ai}^{(n)}(z) = \frac{(-1)^n}{2\pi} \int_{-\infty}^{\infty} (s+it)^n e^{\frac{1}{3}(s+it)^3 - z(s+it)} dt \quad \text{for } n \in \mathbb{Z}_+$$

where $s > 0$ is still arbitrary but fixed.

Since $e^{\frac{1}{3}(s+it)^3 - z(s+it)}$ is the antiderivative of $i((s+it)^2 - z)e^{\frac{1}{3}(s+it)^3 - z(s+it)}$ and vanishes as $t \rightarrow \pm\infty$, we have

$$0 = \frac{i}{2\pi} \int_{-\infty}^{\infty} ((s+it)^2 - z)e^{\frac{1}{3}(s+it)^3 - z(s+it)} dt = i(\text{Ai}''(z) - z\text{Ai}(z))$$

from which follows the simple second-order differential equation for Ai:

$$(4.6) \quad \text{Ai}''(z) = z\text{Ai}(z).$$

Rewrite (4.4) once more with $z = x + iy$ as

$$(4.7) \quad \text{Ai}(z) = \frac{e^{s^3/3 - xs}}{2\pi} \int_{-\infty}^{\infty} e^{-st^2 + yt} (\cos(s^2t - t^3/3 - xt - ys) + i \sin(s^2t - t^3/3 - xt - ys)) dt.$$

If $z = x$ is real so that $y = 0$, the integrand of the imaginary part becomes an odd function of t and vanishes. We conclude that $\text{Ai}(x)$ is real for real x , and thereby also the derivatives $\text{Ai}^{(n)}(x)$ are real for real x .

We insert an easy tail bound.

LEMMA 4.1. *For each $n \in \mathbb{Z}_+$ there is a constant $C_n < \infty$ such that for all $x > 0$,*

$$(4.8) \quad |\text{Ai}^{(n)}(x)| \leq C_n (x^{\frac{n}{2} - \frac{1}{4}} + x^{-\frac{n+1}{4}}) e^{-2x^{3/2}/3}.$$

PROOF. Take $s = x^{1/2}$ in (4.5). Then the exponent is

$$\frac{1}{3}(s+it)^3 - x(s+it) = -\frac{2}{3}x^{3/2} - x^{1/2}t^2 - \frac{1}{3}it^3.$$

Straightforward computation, after taking the modulus inside the integral:

$$\begin{aligned} |\text{Ai}^{(n)}(x)| &\leq C_n e^{-2x^{3/2}/3} \int_{-\infty}^{\infty} (x^{n/2} + |t|^n) e^{-x^{1/2}t^2} dt \\ &\leq C_n e^{-2x^{3/2}/3} \left(x^{n/2} \int_{-\infty}^{\infty} e^{-x^{1/2}t^2} dt + \int_{-\infty}^{\infty} |t|^n e^{-x^{1/2}t^2} dt \right) \\ &\leq C_n e^{-2x^{3/2}/3} (x^{(n/2)-1/4} + x^{-(n/4)-1/4}). \end{aligned} \quad \square$$

The Airy kernel is defined on $\mathbb{R} \times \mathbb{R}$ by

$$(4.9) \quad A(x, y) = \begin{cases} \frac{\text{Ai}(x)\text{Ai}'(y) - \text{Ai}'(x)\text{Ai}(y)}{x - y}, & x \neq y \\ \text{Ai}'(x)\text{Ai}'(x) - \text{Ai}(x)\text{Ai}''(x), & x = y. \end{cases}$$

The definition shows $A(x, y)$ real-valued, continuous and symmetric.

LEMMA 4.2. *The Airy kernel has the representation*

$$(4.10) \quad A(x, y) = \int_0^{\infty} \text{Ai}(x+t)\text{Ai}(y+t) dt, \quad \text{for } x, y \in \mathbb{R}.$$

For any $s \in \mathbb{R}$ there is a finite constant $C(s)$ such that

$$(4.11) \quad |A(x, y)| \leq C(s)e^{-x-y} \quad \text{for } x, y \geq s.$$

PROOF. Bound (4.8) guarantees that the integral in (4.10) is well-defined. Let D_1 and D_2 denote partial differential operators with respect to the first and second argument. Utilizing (4.6) one sees that

$$((D_1 + D_2)A)(x, y) = -\text{Ai}(x)\text{Ai}(y).$$

Also by basic differentiation rules

$$\frac{\partial}{\partial t} A(x+t, y+t) = ((D_1 + D_2)A)(x+t, y+t) = -\text{Ai}(x+t)\text{Ai}(y+t).$$

By estimate (4.8) $A(x+t, y+t) \rightarrow 0$ as $t \rightarrow \infty$, hence

$$\int_0^\infty \text{Ai}(x+t)\text{Ai}(y+t) dt = - \int_0^\infty \frac{\partial}{\partial t} A(x+t, y+t) dt = A(x, y).$$

Equation (4.10) is proved.

Next this upper bound: for each $t \in \mathbb{R}$ there is a finite constant $C_0(t)$ such that

$$(4.12) \quad |\text{Ai}(x)| \leq C_0(t)e^{-x} \quad \text{for } x \in [t, \infty).$$

For $x \geq 1$ this bound follows from (4.8). If $t < 1$ then for $x \in [t, 1]$ note that $\text{Ai}(x)$ is bounded while $e^{-x} \geq e^{-1}$.

From (4.10) and (4.12) for $x, y \geq s$,

$$|A(x, y)| \leq C_0(s)^2 \int_0^\infty e^{-x-y-2t} dt = C(s)e^{-x-y}. \quad \square$$

A consequence of (4.10) is that for any choice of real x_1, \dots, x_n the matrix $\{A(x_i, x_j)\}_{i, j \in [n]}$ is nonnegative definite:

$$(4.13) \quad \begin{aligned} \sum_{i, j \in [n]} A(x_i, x_j) u_i \bar{u}_j &= \int_0^\infty \sum_{i, j \in [n]} \text{Ai}(x_i+t)\text{Ai}(x_j+t) u_i \bar{u}_j dt \\ &= \int_0^\infty \left| \sum_{i \in [n]} \text{Ai}(x_i+t) u_i \right|^2 dt \geq 0 \end{aligned}$$

for any complex vector $(u_1, \dots, u_n) \in \mathbb{C}^n$.

4.2. Tracy-Widom distribution

We need to introduce the notion of a *Fredholm determinant* of an operator on a (possibly) infinite-dimensional space. More specifically, we will consider Fredholm determinants of integral operators.

To pave the way we restate equation (B.5) from a different point of view. Think of an $N \times N$ matrix $A = \{a(x, y)\}_{x, y \in [N]}$ as an integral operator on the L^2 space of the counting measure λ on the space $[N] = \{1, 2, \dots, N\}$. Thinking of a function f on $[N]$ as a column vector, we can write

$$Af(x) = \sum_{y \in [N]} a(x, y)f(y) = \int a(x, y)f(y) \lambda(dy).$$

Equation (B.5) can be written as

$$(4.14) \quad \det(I + A) = 1 + \sum_{n=1}^{\infty} \frac{1}{n!} \int_{[N]^n} \det \{a(x_i, x_j)\}_{i, j \in [n]} \lambda^{\otimes n}(dx_{1, n}).$$

If $n > N$ then $x_{i_1} = x_{i_2}$ for some distinct $i_1, i_2 \in [n]$ and so the rows $\{a(x_{i_1}, x_j)\}$ and $\{a(x_{i_2}, x_j)\}$ agree. Thus all terms for $n > N$ vanish in the sum.

Now generally, suppose we have a measure space $(X, \mathcal{B}, \lambda)$ and a measurable function $k(x, y)$ on $X \times X$ such that for $f \in L^2(\lambda)$

$$Kf(x) = \int_X k(x, y)f(y) \lambda(dy)$$

defines a new function $Kf \in L^2(\lambda)$. Then K is a linear operator on $L^2(\lambda)$ and $k(x, y)$ is the kernel of the operator. The *Fredholm determinant* is defined by

$$(4.15) \quad \det(I + K) = 1 + \sum_{n=1}^{\infty} \frac{1}{n!} \int_{X^n} \det \{k(x_i, x_j)\}_{i,j \in [n]} \lambda^{\otimes n}(dx_{1,n}),$$

provided the series is well-defined. The theory of Fredholm determinants can be developed very generally starting from determinants of finite-rank operators (defined as in (4.14)) and approximating more general operators in suitable norms. We refer to [GGK00].

The Tracy-Widom F_2 distribution, or Tracy-Widom GUE distribution, is defined in terms of the Fredholm determinant of the Airy kernel:

$$(4.16) \quad \begin{aligned} F_2(t) &= \det(I - A|_{L^2[t, \infty)}) \\ &= 1 + \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \int_{[t, \infty)^n} \det \{A(x_i, x_j)\}_{i,j \in [n]} dx_{1,n}. \end{aligned}$$

The first task is to show that the series is well-defined.

From (4.11) and Hadamard's inequality (B.9) we obtain an estimate for the terms of the Fredholm determinant:

$$\begin{aligned} \int_{[t, \infty)^n} \left| \det \{A(x_i, x_j)\}_{i,j \in [n]} \right| dx_{1,n} &\leq \int_{[t, \infty)^n} \prod_{i=1}^n A(x_i, x_i) dx_{1,n} \\ &\leq C(t)^n \int_{[t, \infty)^n} e^{-2 \sum_{i=1}^n x_i} dx_{1,n} = C_1(t)^n \end{aligned}$$

for a new constant $C_1(t)$ that is a function of t . Consequently the series in (4.16) converges absolutely for each fixed $t \in \mathbb{R}$.

We shall take for granted the basic fact that F_2 is indeed the cumulative distribution function of a probability distribution on \mathbb{R} .

F_2 has also the following characterization:

$$(4.17) \quad F_2(t) = \exp \left\{ - \int_t^{\infty} (x-t)u(x)^2 dx \right\}$$

where u is the unique solution of the Painlevé II equation

$$(4.18) \quad u''(x) = xu(x) + 2u(x)^3, \quad u(x) \sim \text{Ai}(x) \text{ as } x \rightarrow \infty.$$

Formula (4.17) together with the boundary condition $u(x) \sim \text{Ai}(x)$ as $x \rightarrow \infty$ and the tail bound (4.8) are enough to conclude that F_2 is a distribution function.

The distribution F_2 first arose as the limit distribution of the largest eigenvalue of the Gaussian Unitary Ensemble (GUE) [TW94]. We summarise this briefly.

A *standard real-valued normal* or *Gaussian* random variable ξ has probability distribution $(2\pi)^{-1/2} e^{-x^2/2} dx$ on \mathbb{R} . A *standard complex-valued normal* random variable is of the form $\zeta = (\xi + i\eta)/\sqrt{2}$ for two independent standard real normal random variables ξ and η . Hence as its real counterpart, the complex normal has $E\zeta = 0$ and $E|\zeta|^2 = 1$.

Let $\{\zeta_{i,i} : i \in \mathbb{N}\}$ be standard real-valued normal random variables and $\{\zeta_{i,j} : i < j \text{ in } \mathbb{N}\}$ standard complex-valued normal random variables. Construct $N \times N$ random Hermitian matrices $X^N = (X_{i,j}^N)_{i,j \in [N]}$ by putting

$$(4.19) \quad X_{i,j}^N = \begin{cases} \zeta_{i,i} & i = j \\ \zeta_{i,j} & i < j \\ \bar{\zeta}_{j,i} & i > j. \end{cases}$$

The distribution of X^N is the Gaussian unitary ensemble (GUE).

As a Hermitian matrix, X^N has N real eigenvalues $\lambda_1^N \leq \lambda_2^N \leq \dots \leq \lambda_N^N$. The Tracy-Widom F_2 distribution arises in this scaling limit: for $t \in \mathbb{R}$

$$(4.20) \quad \lim_{N \rightarrow \infty} \mathbf{P}\{N^{2/3}(N^{-1/2}\lambda_N^N - 2) \leq t\} = F_2(t).$$

Comments

The classic text by Olver [Olv97] is a source of information about the Airy function. The proof of (4.10) is from [TW94].

Distributional limit for the last-passage time

This chapter proves a distributional limit for fluctuations from the law of large numbers of the last passage time with geometric weights. As before the weights $\{Y_{i,j}\}_{(i,j) \in \mathbb{N}^2}$ are i.i.d. geometric with common distribution $\mathbf{P}\{Y_{i,j} = k\} = pq^k$ for $k \in \mathbb{Z}_+$. Parameter $0 < p < 1$ is fixed and $q = 1 - p$. As discovered in Theorem 2.2 the limit

$$\Psi(w, 1) = \lim_{N \rightarrow \infty} N^{-1} G(\lfloor Nw \rfloor, N) \quad \text{a.s.}$$

is explicitly given by

$$(5.1) \quad \Psi(w, 1) = p^{-1} (qw + q + 2\sqrt{qw}).$$

Set

$$(5.2) \quad \sigma = p^{-1} q^{1/6} w^{-1/6} (\sqrt{w} + \sqrt{q})^{2/3} (\sqrt{wq} + 1)^{2/3}.$$

The Tracy-Widom GUE distribution was defined by a Fredholm determinant of the Airy kernel:

$$(5.3) \quad F_2(t) = 1 + \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \int_{[t, \infty)^n} \det_{i,j \in [n]} \{A(x_i, x_j)\} dx_{1,n}.$$

THEOREM 5.1. *For $w \geq 1$ and $t \in \mathbb{R}$*

$$(5.4) \quad \lim_{N \rightarrow \infty} \mathbf{P} \left\{ \frac{G(\lfloor Nw \rfloor, N) - N\Psi(w, 1)}{\sigma N^{1/3}} \leq t \right\} = F_2(t).$$

Proof of Theorem 5.1 takes off from the orthogonal polynomial ensemble representation for the distribution function of $G(M, N)$ obtained in Theorem 3.3 in Chapter 3: for $M \geq N \geq 1$

$$(5.5) \quad \mathbf{P}\{G(M, N) \leq s\} = \frac{1}{Z_{M,N}} \sum_{\substack{x \in \mathbb{Z}_+^N: \\ \forall i \in [N]: x_i \leq s + N - 1}} \Delta_N(x)^2 \prod_{i=1}^N \left\{ \binom{x_i + M - N}{x_i} q^{x_i} \right\}$$

with normalization

$$(5.6) \quad Z_{M,N} = \sum_{x \in \mathbb{Z}_+^N} \Delta_N(x)^2 \prod_{i=1}^N \left\{ \binom{x_i + M - N}{x_i} q^{x_i} \right\}.$$

In the next section another formula for this probability is developed, in terms of the Fredholm determinant of the Meixner kernel. In the limiting regime described in Theorem 5.1 this kernel converges to the Airy kernel. With some additional bounds the Fredholm determinants converge also.

5.1. From orthogonal polynomial ensemble to Meixner kernel

In this section we turn (5.5) into a Fredholm determinant of the Meixner kernel. We have restricted the parameters so that $M \geq N \geq 1$. Throughout this chapter we abbreviate

$$K = M - N + 1.$$

Define the measure μ^K on \mathbb{Z}_+ by

$$(5.7) \quad \mu^K(x) = \binom{x+K-1}{x} q^x, \quad x \in \mathbb{Z}_+.$$

Rewrite (5.5) as

$$(5.8) \quad \mathbf{P}\{G(M, N) \leq s\} = \frac{1}{Z_{M,N}} \sum_{\substack{x \in \mathbb{Z}_+^N \\ \forall i \in [N]: x_i \leq s+N-1}} \Delta_n(x)^2 \prod_{i=1}^N \mu^K(x_i)$$

where the normalization is given by

$$(5.9) \quad Z_{M,N} = \sum_{x \in \mathbb{Z}_+^N} \Delta_n(x)^2 \prod_{i=1}^N \mu^K(x_i).$$

After bringing in some orthogonal polynomials we can give $Z_{M,N}$ an explicit formula (5.32) below.

Define the probability measure $\mathcal{Q}_{M,N}$ on \mathbb{Z}_+^N by

$$(5.10) \quad \mathcal{Q}_{M,N}(A) = \frac{1}{Z_{M,N}} \sum_{x \in \mathbb{Z}_+^N: x \in A} \Delta_n(x)^2 \prod_{j=1}^N \mu^K(x_j)$$

Then

$$(5.11) \quad \begin{aligned} \mathbf{P}\{G(M, N) \leq s\} &= \int \prod_{j=1}^N \mathbf{1}_{[0, s+N-1]}(x_j) \mathcal{Q}_{M,N}(dx_{1,N}) \\ &= \int \prod_{j=1}^N (1 - \mathbf{1}_{[s+N, \infty)}(x_j)) \mathcal{Q}_{M,N}(dx_{1,N}). \end{aligned}$$

The probability measure $\mathcal{Q}_{M,N}$ is exactly of the type (D.14) defined in Appendix D, with the measure μ given by the weights $\{\mu^K(x) : x \in \mathbb{Z}_+\}$. To take advantage of equation (D.20), let $\{M_j^K(x) : j \in \mathbb{Z}_+\}$ denote the polynomials that are orthonormal under the weights $\mu^K(x)$:

$$(5.12) \quad \sum_{x \in \mathbb{Z}_+} M_i^K(x) M_j^K(x) \mu^K(x) = \delta_{i,j}, \quad i, j \in \mathbb{Z}_+$$

and have positive leading coefficient $\kappa_j > 0$: $M_j^K(x) = \kappa_j x^j + \dots$. These are the so-called *Meixner polynomials*. As in (D.19) define on $\mathbb{Z}_+ \times \mathbb{Z}_+$ the kernel

$$(5.13) \quad K_N(x, y) = \sum_{j=0}^{N-1} M_j^K(x) M_j^K(y) \mu^K(x)^{1/2} \mu^K(y)^{1/2}.$$

Proposition D.5, applied to the expectation (5.11) in the form (D.20) gives

$$(5.14) \quad \mathbf{P}\{G(M, N) \leq s\} = 1 + \sum_{k=1}^N \frac{(-1)^k}{k!} \sum_{x \in \mathbb{Z}_+^k: \forall i \in [k]: x_i \geq N+s} \det_{i,j \in [k]} [K_N(x_i, x_j)].$$

The result we want is for $M = \lfloor Nw \rfloor$ for a fixed $w \geq 1$. Introduce also the constant

$$(5.15) \quad \beta = 1 + \Psi(w, 1) = 1 + p^{-1}(qw + q + 2\sqrt{qw}) = p^{-1}(\sqrt{qw} + 1)^2.$$

With $s = \lfloor N\beta - N + \sigma t N^{1/3} \rfloor$ in (5.14) above,

$$(5.16) \quad \mathbf{P} \left\{ \frac{G(\lfloor Nw \rfloor, N) - N\Psi(w, 1)}{\sigma N^{1/3}} \leq t \right\} \\ = 1 + \sum_{k=1}^N \frac{(-1)^k}{k!} \sum_{x \in \mathbb{Z}_+^k} \det_{i,j \in [k]} [K_N(\lfloor N\beta + \sigma t N^{1/3} \rfloor + x_i, \lfloor N\beta + \sigma t N^{1/3} \rfloor + x_j)].$$

The above formula is the starting point for deriving asymptotics.

5.2. Meixner polynomials

We take up some preliminary work on the kernel $K_N(x, y)$ to prepare for convergence. In the literature the Meixner polynomials are normalized somewhat differently from the orthonormal $\{M_j^K(x)\}$. We denote these ‘‘standard’’ Meixner polynomials by $\{m_j^K(x) : j \in \mathbb{Z}_+\}$. One way to determine the polynomials $m_j^K(x)$ is by the normalization

$$(5.17) \quad \sum_{x \in \mathbb{Z}_+} m_i^K(x) m_j^K(x) \mu^K(x) = \delta_{i,j} d_j^2, \quad i, j \in \mathbb{Z}_+$$

with

$$(5.18) \quad d_j^2 = \frac{j!(j+K-1)!}{p^K q^j (K-1)!}$$

and by fixing the sign of the leading coefficient of m_j^K to be $(-1)^j$. Polynomials $\{m_j^K(x)\}$ are uniquely determined, for it follows from Theorem D.1 that a sequence of polynomials $\{f_k(x) : k \in \mathbb{Z}_+\}$ is uniquely determined by specifying that f_k have precisely degree k with a given nonzero leading coefficient, and by requiring orthogonality $\int f_k f_\ell d\mu = 0$ for $k \neq \ell$. The two types of Meixner polynomials are connected by

$$(5.19) \quad M_j^K(x) = \frac{(-1)^j}{d_j} m_j^K(x).$$

We derive some explicit formulas. In (5.22) below we use noninteger powers of complex numbers. By definition, $z^a = e^{a \log z}$ and throughout this text we take the principal branch for the logarithm: $\log z$ is the holomorphic function on the complement of the nonpositive real axis for which $\log z = \log|z| + i \arg z$ with $\arg z \in (-\pi, \pi)$. Thus also z^a is holomorphic for $z \in \mathbb{C} \setminus \mathbb{R}_-$ for any fixed $a \in \mathbb{C}$.

PROPOSITION 5.2. *For $x \in \mathbb{R}$ and $n \in \mathbb{Z}_+$ we have the formula*

$$(5.20) \quad m_n^K(x) = (-1)^n n! \sum_{k=0}^n \binom{x}{k} \binom{-x-K}{n-k} q^{-k}$$

and the leading term is given by

$$(5.21) \quad m_n^K(x) = \left(-\frac{p}{q}\right)^n x^n + [\text{terms of degree less than } n \text{ in } x].$$

The generating function is

$$(5.22) \quad \sum_{n=0}^{\infty} \frac{z^n}{n!} m_n^K(x) = \left(1 - \frac{z}{q}\right)^x (1-z)^{-x-K}$$

for $x \in \mathbb{R}$ and complex z such that $|z| < q$.

PROOF. Let us begin by taking (5.20) as the definition of $m_n^K(x)$. Then $m_n^K(x)$ is a polynomial of degree n with leading coefficient given by (5.21). We derive a bound on $|m_n^K(x)|$ to justify convergence of the series in (5.22). In formula (5.20) apply bound (C.10) to estimate the binomial coefficients:

$$(5.23) \quad \begin{aligned} |m_n^K(x)| &\leq n!(ne)^{|x+K|} + n! \sum_{k=1}^{n-1} q^{-k} (ke)^{|x|} ((n-k)e)^{|x+K|} + n!q^{-n} (ne)^{|x|} \\ &\leq (n+1)!q^{-n} (ne)^{2|x+K|}. \end{aligned}$$

Thus for each fixed $x \in \mathbb{R}$, $|z| < q$ is within the radius of convergence of the series in (5.22).

Having justified the convergence we compute the generating function.

$$(5.24) \quad \begin{aligned} \sum_{n=0}^{\infty} \frac{z^n}{n!} m_n^K(x) &= \sum_{n=0}^{\infty} z^n \sum_{k=0}^n \binom{x}{k} (-q)^{-k} \binom{-x-K}{n-k} (-1)^{n-k} \\ &= \sum_{k=0}^{\infty} \binom{x}{k} (-z/q)^k \cdot \sum_{j=0}^{\infty} \binom{-x-K}{j} (-z)^j \\ &= (1 - (z/q))^x (1 - z)^{-x-K}. \end{aligned}$$

Since both z/q and z lie in the open unit disk, the two series on the second-last line are holomorphic functions and equal the functions on the last line. (See the discussion following equation (C.9) in Appendix C.2.)

Thus the proposition is true, provided the polynomials $m_n^K(x)$ defined by (5.20) are the right ones. Since we already verified the sign of the leading coefficient in (5.21), it only remains to verify the orthogonality relation (5.17).

We shall check the orthogonality with the help of the series identity

$$(5.25) \quad \sum_{k \geq 0, n \geq 0} \frac{u^k v^n}{k!n!} \sum_{x \in \mathbb{Z}_+} \binom{x+K-1}{x} q^x m_k^K(x) m_n^K(x) = \sum_{k \geq 0, n \geq 0} \frac{u^k v^n}{k!n!} \delta_{k,n} d_n^2.$$

We first check absolute convergence of the series. There is no convergence issue on the right-hand side of (5.25). To check absolute convergence on the left-hand side, use formula (5.20) for the polynomials and note that for integers $x, K \geq 0$ and $n \geq j$:

$$\begin{aligned} \left| \binom{-x-K}{n-j} \right| &= \left| \frac{(-x-K)(-x-K-1) \cdots (-x-K-n+j+1)}{(n-j)!} \right| \\ &= \binom{x+K+n-j-1}{n-j}. \end{aligned}$$

Then the series of absolute values of the terms on the left-hand side of (5.25) is bounded by this series:

$$\begin{aligned} &\sum_{k,n,x \geq 0} |u|^k |v|^n \binom{x+K-1}{x} q^x \sum_{i=0}^k \binom{x}{i} \binom{x+K+k-i-1}{k-i} q^{-i} \\ &\quad \times \sum_{j=0}^n \binom{x}{j} \binom{x+K+n-j-1}{n-j} q^{-j} \\ &= \sum_{i,j,x \geq 0} \binom{x+K-1}{x} q^{x-i-j} \binom{x}{i} \binom{x}{j} \sum_{k=i}^{\infty} \binom{x+K+k-i-1}{k-i} |u|^k \end{aligned}$$

$$(5.26) \quad \times \sum_{n=j}^{\infty} \binom{x+K+n-j-1}{n-j} |v|^n$$

To sum up the last two series over k and n use this formula: for $a \in \mathbb{R}$ and complex ζ in the open unit disk

$$(5.27) \quad \begin{aligned} \sum_{n \geq 0} \binom{a+n}{n} \zeta^n &= \sum_{n \geq 0} \frac{(a+n)(a-1+n) \cdots (a+1)}{n!} \zeta^n \\ &= \sum_{n \geq 0} \frac{(-a-1)(-a-2) \cdots (-a-n)}{n!} (-\zeta)^n \\ &= \sum_{n \geq 0} \binom{-a-1}{n} (-\zeta)^n = (1-\zeta)^{-a-1}. \end{aligned}$$

Continuing then from line (5.26):

$$= \sum_{i,j,x \geq 0} \binom{x+K-1}{x} q^x \binom{x}{i} \binom{x}{j} (|u|/q)^i (|v|/q)^j (1-|u|)^{-x-K} (1-|v|)^{-x-K}$$

which after summing over i and j

$$= (1-|u|)^{-K} (1-|v|)^{-K} \sum_{x \geq 0} \binom{x+K-1}{x} \left(\frac{q(1+|u|/q)(1+|v|/q)}{(1-|u|)(1-|v|)} \right)^x.$$

We can guarantee absolute convergence of the last series by picking $s > 0$ so that $q \in (0, 1)$ satisfies

$$q < \frac{1}{(1+s)^2 + 2s}$$

and by then requiring that the complex u, v lie in the open disk of radius sq around the origin. We can now work with the series in (5.25).

The right-hand side of (5.25) develops into

$$(5.28) \quad \begin{aligned} \sum_{n \geq 0} \frac{u^n v^n}{(n!)^2} \frac{n!(n+K-1)!}{p^K q^n (K-1)!} &= \frac{1}{p^K} \sum_{n \geq 0} \binom{n+K-1}{n} \left(\frac{uv}{q} \right)^n \\ &= p^{-K} \left(1 - \frac{uv}{q} \right)^{-K}. \end{aligned}$$

The last equality is another instance of (5.27).

To transform the left-hand side of (5.25), change order of summation, use the generating function (5.22), then set

$$\zeta = \frac{q(1-u/q)(1-v/q)}{(1-u)(1-v)} \quad \text{with} \quad 1-\zeta = \frac{p(q-uv)}{q(1-u)(1-v)}$$

and get

$$\begin{aligned}
& \sum_{x \in \mathbb{Z}_+} \binom{x+K-1}{x} q^x \left\{ \sum_{k \geq 0} \frac{u^k}{k!} m_k^K(x) \right\} \left\{ \sum_{n \geq 0} \frac{v^n}{n!} m_n^K(x) \right\} \\
&= \sum_{x \in \mathbb{Z}_+} \binom{x+K-1}{x} q^x (1-(u/q))^x (1-u)^{-x-K} (1-(v/q))^x (1-v)^{-x-K} \\
&= (1-u)^{-K} (1-v)^{-K} \sum_{x \in \mathbb{Z}_+} \binom{x+K-1}{x} \zeta^x \\
&= (1-u)^{-K} (1-v)^{-K} (1-\zeta)^{-K} \\
(5.29) \quad &= \frac{q^K}{p^K (q-uv)^K}.
\end{aligned}$$

Lines (5.28) and (5.29) agree and thereby (5.25) has been verified.

The power series in (5.25) can be differentiated term by term with respect to u and v . Operating with $(\partial^k / \partial u^k)$ and $(\partial^n / \partial v^n)$ and setting $u = v = 0$ gives

$$(5.30) \quad \sum_{x \in \mathbb{Z}_+} \binom{x+K-1}{x} q^x m_k^K(x) m_n^K(x) = \delta_{k,n} d_n^2.$$

This is the desired orthogonality relation (5.17). To summarize, by verifying that the polynomials defined by (5.20) satisfy the this orthogonality we have concluded the proof of the proposition. \square

Since κ_j was by definition the leading coefficient of the polynomial $M_j^K(x)$, comparison of (5.19) and (5.21) gives

$$(5.31) \quad \kappa_j = \frac{1}{d_j} \left(\frac{p}{q} \right)^j.$$

Now we are in a position to compute explicitly the normalization constant $Z_{M,N}$ in (5.9). Although, it will turn out that this is not needed for the limit.

LEMMA 5.3. *For $M \geq N \geq 1$ the constant $Z_{M,N}$ defined in (5.9) is given by*

$$(5.32) \quad Z_{M,N} = q^{\frac{N(N-1)}{2}} p^{-MN} \frac{N!}{((M-N)!)^N} \prod_{j=0}^{N-1} j!(M-N+j)!.$$

PROOF. By equation (D.18) in the Appendix

$$Z_{M,N} = N! \prod_{j=0}^{N-1} \kappa_j^{-2}.$$

Now substitute in (5.31) and simplify. \square

Our ultimate goal is to find asymptotics for the kernel K_N . By the Christoffel-Darboux formula (Theorem D.4 in Appendix D), for $x \neq y$ in \mathbb{Z}_+

$$(5.33) \quad K_N(x, y) = \frac{\kappa_{N-1}}{\kappa_N} \frac{M_N^K(x) M_{N-1}^K(y) - M_N^K(y) M_{N-1}^K(x)}{x-y} \mu^K(x)^{1/2} \mu^K(y)^{1/2}$$

$$(5.34) \quad = \frac{-q}{p d_{N-1}^2} \frac{m_N^K(x) m_{N-1}^K(y) - m_N^K(y) m_{N-1}^K(x)}{x-y} \mu^K(x)^{1/2} \mu^K(y)^{1/2}.$$

On the diagonal

$$(5.35) \quad K_N(x, x) = \frac{-q}{pd_{N-1}^2} [(m_N^K)'(x)m_{N-1}^K(x) - m_N^K(x)(m_{N-1}^K)'(x)]\mu^K(x).$$

As the last item of this section we derive integral representations for the polynomials $m_n^K(x)$ to be used in the analysis of asymptotics in formulas (5.34) and (5.35). Recall that $w \geq 1$ was fixed by the statement (5.4) that we are out to prove. In the generating function (5.22) replace z by $-z\sqrt{q/w}$ so that this formula is valid for $|z| < \sqrt{wq}$:

$$(5.36) \quad \sum_{n=0}^{\infty} \frac{(-1)^n z^n}{n!} \left(\frac{q}{w}\right)^{n/2} m_n^K(x) = \frac{w^{K/2}}{q^{x/2}} \cdot \frac{(\sqrt{wq} + z)^x}{(\sqrt{w} + z\sqrt{q})^{x+K}}$$

Multiply through above by $(2\pi i)^{-1} z^{-N-1}$ and integrate z over a circle Γ_r of radius $r < \sqrt{wq}$. Note that the denominator on the right does not vanish for z on this circle.

$$\frac{(-1)^N}{N!} \left(\frac{q}{w}\right)^{N/2} m_N^K(x) = \frac{w^{K/2}}{q^{x/2}} \cdot \frac{1}{2\pi i} \int_{\Gamma_r} \frac{(\sqrt{wq} + z)^x}{(\sqrt{w} + z\sqrt{q})^{x+K}} \frac{dz}{z^{N+1}}$$

from which we get the formula

$$(5.37) \quad m_N^K(x) = \frac{w^{\frac{N+K}{2}}}{q^{\frac{N+x}{2}}} \cdot \frac{(-1)^N N!}{2\pi i} \int_{\Gamma_r} \frac{(\sqrt{wq} + z)^x}{(\sqrt{w} + z\sqrt{q})^{x+K}} \frac{dz}{z^{N+1}}.$$

The integral representation (5.37) is valid for circles Γ_r with radius $r < \sqrt{wq}$. Eventually for the asymptotics we wish to let r approach 1 which is problematic if $\sqrt{wq} < 1$. However, we actually only need the kernel $K_N(x, y)$ for integer values $x, y \in \mathbb{Z}_+$. For $x \in \mathbb{Z}_+$ the integrand on the right in (5.37) is holomorphic on the open annulus $\{z : 0 < |z| < \sqrt{w/q}\}$. Thus by Cauchy's theorem [Rud87, Theorem 10.35] identity (5.37) is in fact valid for all radii $0 < r < \sqrt{w/q}$ and so letting $r \nearrow 1$ is not problematic. Consequently we will use (5.37) for $x \in \mathbb{Z}_+$ and radii r up to 1.

But we need to work more on the integral representation because $K_N(x, x)$ of (5.35) requires us to differentiate (5.37) and for this we need real x and not only integral x . When x is real we rely on

$$(\sqrt{wq} + z)^x = e^{x \log(\sqrt{wq} + z)}$$

which defines a holomorphic function only for $z \in \mathbb{C} \setminus (-\infty, -\sqrt{wq}]$. Thus in the case $\sqrt{wq} < 1$, for $x \in \mathbb{R}_+$ we cannot extend the validity of (5.37) to all radii $r < 1$ for free. An application of Lemma C.3 from Appendix C.2 to formula (5.37) yields the following more complicated formula, now valid for all real $x > -1$ and all radii $0 < r < 1$:

$$(5.38) \quad m_N^K(x) = \frac{w^{\frac{N+K}{2}}}{q^{\frac{N+x}{2}}} \cdot (-1)^N N! \left\{ \frac{1}{2\pi i} \int_{\Gamma_r} \frac{(\sqrt{wq} + z)^x}{(\sqrt{w} + z\sqrt{q})^{x+K}} \frac{dz}{z^{N+1}} \right.$$

$$(5.39) \quad \left. + \mathbf{1}\{r > \sqrt{wq}\} \cdot (-1)^{N+1} \cdot \frac{\sin \pi x}{\pi} \int_{\sqrt{wq}}^r \frac{|\sqrt{wq} - s|^x}{(\sqrt{w} - s\sqrt{q})^{x+K}} \frac{ds}{s^{N+1}} \right\}.$$

We also assume that $r \neq \sqrt{wq}$ so that $\log(\sqrt{wq} + z)$ is defined and bounded for all $z \in \Gamma_r \setminus \{-r\}$. If $\sqrt{wq} \geq 1$ this is not problematic since in any case we only consider $r < 1$. If $\sqrt{wq} < 1$ we assume $\sqrt{wq} < r < 1$. Since $q < 1 \leq w$ the denominators of the integrands cannot vanish. With the abbreviations

$$(5.40) \quad a(z) = \frac{\sqrt{wq} + z}{\sqrt{wq} + 1} \cdot \frac{\sqrt{w} + \sqrt{q}}{\sqrt{w} + z\sqrt{q}} \quad \text{and} \quad b(z) = \frac{\sqrt{w} + \sqrt{q}}{\sqrt{w} + z\sqrt{q}}$$

the formula becomes

$$(5.41) \quad m_N^K(x) = \frac{(\sqrt{wq} + 1)^x}{(\sqrt{w} + \sqrt{q})^{x+K}} \cdot \frac{w^{\frac{N+K}{2}}}{q^{\frac{N+x}{2}}} \cdot (-1)^N N! \left\{ \frac{1}{2\pi i} \int_{\Gamma_r} a(z)^x b(z)^K \frac{dz}{z^{N+1}} \right. \\ (5.42) \quad \left. + \mathbf{1}\{r > \sqrt{wq}\} \cdot (-1)^{N+1} \cdot \frac{\sin \pi x}{\pi} \int_{\sqrt{wq}}^r |a(-s)|^x b(-s)^K \frac{ds}{s^{N+1}} \right\}.$$

Recalling

$$(5.43) \quad \beta = p^{-1}(\sqrt{wq} + 1)^2$$

and defining

$$(5.44) \quad \alpha = \beta + w - 1 = p^{-1}(\sqrt{w} + \sqrt{q})^2$$

we rewrite the formula once more as

$$(5.45) \quad m_N^K(x) = \frac{w^{\frac{N+K}{2}} \beta^{\frac{x}{2}} (-1)^N N!}{q^{\frac{N+x}{2}} \alpha^{\frac{x+K}{2}} p^{\frac{K}{2}}} \cdot \left\{ \frac{1}{2\pi i} \int_{\Gamma_r} a(z)^x b(z)^K \frac{dz}{z^{N+1}} \right. \\ (5.46) \quad \left. + \mathbf{1}\{r > \sqrt{wq}\} \cdot (-1)^{N+1} \cdot \frac{\sin \pi x}{\pi} \int_{\sqrt{wq}}^r |a(-s)|^x b(-s)^K \frac{ds}{s^{N+1}} \right\}.$$

Note that when x is an integer the part on line (5.46) vanishes due to $\sin \pi x = 0$. With the temporary abbreviation

$$(5.47) \quad B_N(x) = \frac{w^{\frac{N+K}{2}} \beta^{\frac{x}{2}} (-1)^N N!}{q^{\frac{N+x}{2}} \alpha^{\frac{x+K}{2}} p^{\frac{K}{2}}}$$

the formula becomes

$$(5.48) \quad m_N^K(x) = B_N(x) \left\{ \frac{1}{2\pi i} \int_{\Gamma_r} a(z)^x b(z)^K \frac{dz}{z^{N+1}} \right. \\ (5.49) \quad \left. + \mathbf{1}\{r > \sqrt{wq}\} \cdot (-1)^{N+1} \cdot \frac{\sin \pi x}{\pi} \int_{\sqrt{wq}}^r |a(-s)|^x b(-s)^K \frac{ds}{s^{N+1}} \right\}.$$

We differentiate the above equation with respect to x . We need the answer only for integral x which simplifies the resulting formula. We also need it only for large x , so we restrict to $x \in \mathbb{N}$ to avoid considering negative x .

PROPOSITION 5.4. *For all $x \in \mathbb{N}$,*

$$(5.50) \quad (m_N^K)'(x) = B_N(x) \cdot \frac{1}{2} \log\left(\frac{\beta}{q\alpha}\right) \frac{1}{2\pi i} \int_{\Gamma_r} a(z)^x b(z)^K \frac{dz}{z^{N+1}} \\ + B_N(x) \left\{ \frac{1}{2\pi i} \int_{\Gamma_r} (\log a(z)) a(z)^x b(z)^K \frac{dz}{z^{N+1}} \right. \\ \left. + \mathbf{1}\{r > \sqrt{wq}\} \cdot (-1)^{x+N+1} \int_{\sqrt{wq}}^r |a(-s)|^x b(-s)^K \frac{ds}{s^{N+1}} \right\}.$$

PROOF. Formally (5.50) follows from the product rule in a straightforward manner. But we need to justify and explain some steps.

Let us first address the well-definedness and boundedness of $\log a(z)$. We can write

$$a(z) = \frac{c}{|\sqrt{w} + z\sqrt{q}|^2} (w\sqrt{q} + z\sqrt{w} + \bar{z}q\sqrt{w} + |z|^2\sqrt{q})$$

for a positive constant c . First note that an inadmissible value $a(z) \in \mathbb{R}_-$ can happen on Γ_r only if $z = -r$ (and q must be small enough relative to w). A single exceptional point $\{-r\}$ has no bearing on the well-definedness of integrals over Γ_r . The numerator above vanishes only at $z = -\sqrt{qw}$ and $z = -\sqrt{w/q}$ and the denominator only at $z = -\sqrt{w/q}$. Since we assume $0 < r < 1$ and $r \neq \sqrt{wq}$, $|a(z)|$ stays bounded and bounded away from zero on Γ_r . Consequently the function $\log a(z)$ is defined and bounded on $\Gamma_r \setminus \{-r\}$.

It follows from these considerations that, with the definition $a(z)^x = e^{x \log a(z)}$, the integral on line (5.48) can be differentiated in x to yield the integral on the second line of (5.50). More precisely, the operation d/dx can be taken inside the integral on line (5.48) because

$$(5.51) \quad \begin{aligned} \frac{a(z)^{x+h} - a(z)^x}{h} &= a(z)^x \cdot \frac{e^{h \log a(z)} - 1}{h} \\ &= a(z)^x \left\{ \log a(z) + \sum_{k=2}^{\infty} \frac{h^{k-1} (\log a(z))^k}{k!} \right\} \end{aligned}$$

converges boundedly as $h \rightarrow 0$, for a fixed $x \in \mathbb{R}$ and while z varies over $\Gamma_r \setminus \{-r\}$.

We turn to justify

$$(5.52) \quad \frac{d}{dx} \left\{ \frac{\sin \pi x}{\pi} \int_{\sqrt{wq}}^r |a(-s)|^x b(-s)^K \frac{ds}{s^{N+1}} \right\} = (-1)^x \int_{\sqrt{wq}}^r |a(-s)|^x b(-s)^K \frac{ds}{s^{N+1}}.$$

The identity is claimed only for points $x \in \mathbb{N}$.

The function $|a(-s)|^x$ is jointly continuous as a function of $(x, s) \in [1/2, \infty) \times [\sqrt{wq}, r]$, hence the integral is a continuous function of $x \geq 1/2$. (Here $x < 0$ would make $|a(-s)|^x$ unbounded and $(x, s) = (0, \sqrt{wq})$ is a discontinuity, so restricting to $x > 0$ is convenient.) Still we cannot blithely repeat argument (5.51) because $\log|a(-s)|$ blows up at $s = \sqrt{wq}$. However, we can avoid looking at the details with this simple lemma:

LEMMA 5.5. *Let $G(x) = f(x)g(x)$ on some interval of the real line. Suppose u is a point such that $f(u) = 0$, f is differentiable at u and g is continuous at u . Then $G'(u) = f'(u)g(u)$.*

PROOF. By the hypotheses,

$$\frac{G(u+h) - G(u)}{h} = \frac{f(u+h)}{h} g(u+h) \rightarrow f'(u)g(u) \quad \text{as } h \rightarrow 0. \quad \square$$

Identity (5.52) is verified by applying the lemma to

$$f(x) = (\sin \pi x)/\pi \quad \text{and} \quad g(x) = \int_{\sqrt{wq}}^r |a(-s)|^x b(-s)^K \frac{ds}{s^{N+1}}.$$

$(d/dx)(\sin \pi x)/\pi = \cos \pi x$ which equals $(-1)^x$ for $x \in \mathbb{Z}_+$. This completes the proof of the differentiation formula (5.50). \square

We are ready to start deriving the asymptotics to which the next section is devoted. The entire development is based on the Christoffel-Darboux formulas (5.34) and (5.35) for $K_N(x, y)$, and on the integral formulas (5.37) for $m_N^K(x)$ and (5.50) for $(m_N^K)'(x)$ for $x \in \mathbb{Z}_+$.

5.3. Airy asymptotics for the Meixner kernel

The estimates that imply the Airy limit are summarized in the next theorem. From definition (5.13) it is immediate that the kernel K_N is symmetric: $K_N(x, y) = K_N(y, x)$. Since it is real-valued,

it follows that K_N is *Hermitian*: $K_N(x, y) = \overline{K_N(y, x)}$. Furthermore, K_N is *nonnegative definite*: for any finitely supported function $f : \mathbb{Z}_+ \rightarrow \mathbb{C}$, abbreviating temporarily $\nu(x) = m_k^K(x)\mu^K(x)^{1/2}$,

$$(5.53) \quad \begin{aligned} \sum_{x, y \in \mathbb{Z}_+} \overline{f(x)} K_N(x, y) f(y) &= \sum_{k=0}^{N-1} \left(\sum_x \overline{f(x)} \nu_k(x) \right) \left(\sum_y f(y) \nu_k(y) \right) \\ &= \sum_{k=0}^{N-1} \overline{\left(\sum_x f(x) \nu_k(x) \right)} \left(\sum_y f(y) \nu_k(y) \right) = \sum_{k=0}^{N-1} \left| \sum_x f(x) \nu_k(x) \right|^2 \geq 0. \end{aligned}$$

In particular, since $K_N(x, x) \geq 0$ absolute values are not needed in the assumptions below.

THEOREM 5.6. *Let $\beta > 0$ be a constant. Assume given a scaling $\lambda_N \rightarrow \infty$ as $N \rightarrow \infty$ along positive integers and assume $\lambda_N = o(N)$. Let $K_N : \mathbb{Z}_+^2 \rightarrow \mathbb{R}$ be a sequence of Hermitian nonnegative definite kernels defined for all large enough $N \in \mathbb{Z}_+$. Assume that the following properties (i)–(iv) hold.*

(i) *For each $\tau \in \mathbb{R}$ there exist constants $C(\tau), N_0(\tau) < \infty$ such that*

$$(5.54) \quad \sup_{N \geq N_0(\tau)} \sum_{m=0}^{\infty} K_N(\lfloor N\beta + \lambda_N \tau \rfloor + m, \lfloor N\beta + \lambda_N \tau \rfloor + m) \leq C(\tau).$$

(ii) *For every $\varepsilon > 0$ there exist finite $L = L(\varepsilon), N_0(\varepsilon) < \infty$ such that*

$$(5.55) \quad \sup_{N \geq N_0(\varepsilon)} \sum_{m=0}^{\infty} K_N(\lfloor N\beta + \lambda_N L \rfloor + m, \lfloor N\beta + \lambda_N L \rfloor + m) \leq \varepsilon.$$

(iii) *For each $M < \infty$ there exists $N_0(M) < \infty$ such that*

$$(5.56) \quad \sup_{-M \leq \xi \leq M} \sup_{N \geq N_0(M)} \lambda_N K_N(\lfloor N\beta + \lambda_N \xi \rfloor, \lfloor N\beta + \lambda_N \xi \rfloor) < \infty.$$

(iv) *Scaled kernels converge pointwise to the Airy kernel defined in (4.9). Let $\xi, \eta \in \mathbb{R}$. Suppose $\ell_\xi = \ell_\xi(N)$ and $\ell_\eta = \ell_\eta(N)$ are quantities such that $x = N\beta + \ell_\xi$ and $y = N\beta + \ell_\eta$ are nonnegative integers for each N , and for some constant C ,*

$$(5.57) \quad |\ell_\xi - \lambda_N \xi| + |\ell_\eta - \lambda_N \eta| \leq C$$

for all large enough N . Then if $\xi \neq \eta$

$$(5.58) \quad \lim_{N \rightarrow \infty} \lambda_N K_N(N\beta + \ell_\xi, N\beta + \ell_\eta) = A(\xi, \eta)$$

and also on the diagonal

$$(5.59) \quad \lim_{N \rightarrow \infty} \lambda_N K_N(N\beta + \ell_\xi, N\beta + \ell_\xi) = A(\xi, \xi).$$

From these assumptions it follows that the Fredholm determinants converge: for each $t \in \mathbb{R}$,

$$(5.60) \quad 1 + \lim_{N \rightarrow \infty} \sum_{k=1}^N \frac{(-1)^k}{k!} \sum_{h \in \mathbb{Z}_+^k} \det_{i, j \in [k]} \{K_N(\lfloor N\beta + \lambda_N t \rfloor + h_i, \lfloor N\beta + \lambda_N t \rfloor + h_j)\}$$

$$(5.61) \quad = 1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_{[t, \infty)^k} \det_{i, j \in [k]} \{A(x_i, x_j)\} dx_{1, k}.$$

PROOF. Fix $t \in \mathbb{R}$. The strategy of the proof is to show the convergence (5.60)–(5.61) for k and h restricted to bounded sets, and then to show that the terms left out on both sides add up to very little. Consider $0 < \ell, L < \infty$ fixed for the moment. Let $\bar{1} = (1, 1, \dots, 1)$.

$$\begin{aligned}
(5.62) \quad & 1 + \sum_{k=1}^{\ell} \frac{(-1)^k}{k!} \sum_{h \in \{0, \dots, \lfloor \lambda_N L \rfloor - 1\}^k} \det_{i,j \in [k]} \{K_N(\lfloor N\beta + \lambda_N t \rfloor + h_i, \lfloor N\beta + \lambda_N t \rfloor + h_j)\} \\
& = 1 + \sum_{k=1}^{\ell} \frac{(-1)^k}{k!} \int_{[0, L]^k} \sum_{h \in \{0, \dots, \lfloor \lambda_N L \rfloor - 1\}^k} \mathbf{1}_{[\lambda_N^{-1}h, \lambda_N^{-1}(h+\bar{1})]}(x_{1,k}) \\
& \quad \times \det_{i,j \in [k]} \{\lambda_N K_N(\lfloor N\beta + \lambda_N t \rfloor + \lfloor \lambda_N x_i \rfloor, \lfloor N\beta + \lambda_N t \rfloor + \lfloor \lambda_N x_j \rfloor)\} dx_{1,k}.
\end{aligned}$$

By assumption (iv) for each fixed $k \geq 1$ and $x_{1,k} \in [0, L]^k$ the integrand converges to $\det_{i,j \in [k]} [A(t + x_i, t + x_j)]$. Hadamard's inequality (B.9) (Appendix B) and assumption (iii) give

$$\begin{aligned}
& \sup_{x \in [0, L]^k} \left| \det_{i,j \in [k]} \{\lambda_N K_N(\lfloor N\beta + \lambda_N t \rfloor + \lfloor \lambda_N x_i \rfloor, \lfloor N\beta + \lambda_N t \rfloor + \lfloor \lambda_N x_j \rfloor)\} \right| \\
& \leq \sup_{x \in [0, L]^k} \prod_{i=1}^k \left\{ \lambda_N K_N(\lfloor N\beta + \lambda_N t \rfloor + \lfloor \lambda_N x_i \rfloor, \lfloor N\beta + \lambda_N t \rfloor + \lfloor \lambda_N x_i \rfloor) \right\} \leq C^k
\end{aligned}$$

for a constant $C = C(L, t)$ that is independent of k . This bound is good enough for dominated convergence, and we can conclude that for all fixed $0 < \ell, L < \infty$ the sum on line (5.62) converges as $N \rightarrow \infty$ to

$$(5.63) \quad 1 + \sum_{k=1}^{\ell} \frac{(-1)^k}{k!} \int_{[t, t+L]^k} \det_{i,j \in [k]} [A(x_i, x_j)] dx_{1,k}.$$

The remainder of the proof consists in showing that the parts missing from (5.62) and (5.63) in comparison with their counterparts on lines (5.60) and (5.61) can be made arbitrarily small by choosing ℓ and L large enough.

Consider first the difference between (5.60) and (5.62) with $N > \ell$. Let us abbreviate

$$x_N = \lfloor N\beta + \lambda_N t \rfloor.$$

This difference comes as a sum of two parts:

$$(5.64) \quad \sum_{k=1}^{\ell} \frac{(-1)^k}{k!} \sum_{h \in \mathbb{Z}_+^k : \exists i \in [k] : h_i \geq \lfloor \lambda_N L \rfloor} \det_{i,j \in [k]} \{K_N(x_N + h_i, x_N + h_j)\}$$

$$(5.65) \quad + \sum_{k=\ell+1}^N \frac{(-1)^k}{k!} \sum_{h \in \mathbb{Z}_+^k} \det_{i,j \in [k]} \{K_N(x_N + h_i, x_N + h_j)\}$$

Applying Hadamard's inequality again, we bound the absolute value of the sum on line (5.64) by

$$\begin{aligned}
& \sum_{k=1}^{\ell} \frac{1}{k!} \sum_{i=1}^k \sum_{h \in \mathbb{Z}_+^k: h_i \geq \lfloor \lambda_N L \rfloor} \prod_{i=1}^k K_N(x_N + h_i, x_N + h_i) \\
&= \sum_{k=1}^{\ell} \frac{1}{(k-1)!} \left(\sum_{m=0}^{\infty} K_N(x_N + m, x_N + m) \right)^{k-1} \left(\sum_{m=\lfloor \lambda_N L \rfloor}^{\infty} K_N(x_N + m, x_N + m) \right) \\
(5.66) \quad & \leq \sum_{k=1}^{\ell} \frac{1}{(k-1)!} C(t)^{k-1} \varepsilon \leq e^{C(t)} \varepsilon.
\end{aligned}$$

Above we used assumptions (i) and (ii), choosing L large enough to get bound ε in (ii).

Next we bound the absolute value of the sum on line (5.65), again starting off with an application of Hadamard's inequality:

$$\begin{aligned}
& \sum_{k=\ell+1}^N \frac{1}{k!} \sum_{h \in \mathbb{Z}_+^k} \prod_{i=1}^k K_N(x_N + h_i, x_N + h_i) \leq \sum_{k=\ell+1}^N \frac{1}{k!} \left(\sum_{m=0}^{\infty} K_N(x_N + m, x_N + m) \right)^k \\
(5.67) \quad & \leq \sum_{k=\ell+1}^N \frac{1}{k!} C(t)^k \leq \varepsilon
\end{aligned}$$

if ℓ is chosen large enough. Since t is fixed, we have shown that the absolute value of the sum of lines (5.64) and (5.65) can be made arbitrarily small uniformly over N by choosing ℓ and L large enough.

Last we derive the analogous bounds for the Airy kernel, namely that by choosing ℓ and L large enough,

$$\begin{aligned}
(5.68) \quad & \left| \sum_{k=1}^{\ell} \frac{(-1)^k}{k!} \int_{[t, \infty)^k \setminus [t, t+L]^k} \det_{i,j \in [k]} [A(x_i, x_j)] dx_{1,k} \right| \\
(5.69) \quad & + \left| \sum_{k=\ell+1}^{\infty} \frac{(-1)^k}{k!} \int_{[t, \infty)^k} \det_{i,j \in [k]} [A(x_i, x_j)] dx_{1,k} \right| \leq \varepsilon.
\end{aligned}$$

Both sums are controlled by Hadamard's inequality and estimate (4.11). By (4.10) $A(x, x) \geq 0$ and hence absolute values are not needed below. The sum on line (5.68) is bounded by

$$\begin{aligned}
& \sum_{k=1}^{\ell} \frac{1}{k!} \int_{[t, \infty)^k \setminus [t, t+L]^k} \prod_{i=1}^k A(x_i, x_i) dx_{1,k} \\
& \leq \sum_{k=1}^{\ell} \frac{1}{(k-1)!} \left(\int_t^{\infty} A(s, s) ds \right)^{k-1} \left(\int_{t+L}^{\infty} A(s, s) ds \right) \\
(5.70) \quad & \leq \sum_{k=1}^{\ell} \frac{1}{(k-1)!} C(t)^{k-1} \varepsilon \leq e^{C(t)} \varepsilon.
\end{aligned}$$

Above $C(t) = \int_t^\infty A(s, s) ds < \infty$ and L is chosen so that $\int_{t+L}^\infty A(s, s) ds < \varepsilon$, both by (4.11). Finally, utilizing the same $C(t)$,

$$\text{line (5.69)} \leq \sum_{k=\ell+1}^{\infty} \frac{1}{k!} \int_{[t, \infty)^k} \prod_{i=1}^k A(x_i, x_i) dx_{1,k} \leq \sum_{k=\ell+1}^{\infty} \frac{1}{k!} C(t)^k \leq \varepsilon \quad (5.71)$$

by choosing ℓ large enough.

To summarize, the convergence of line (5.60) to line (5.61) has been proved for a fixed $t \in \mathbb{R}$, by showing the convergence of (5.62) to (5.63) and by showing that for large L and ℓ the sum (5.62) is close to the sum (5.60) uniformly in N , and sum (5.63) is close to (5.61). \square

Theorem 5.1 is now proved by applying Theorem 5.6 to the determinantal representation (5.16) of the scaled distribution function. The remainder of this section checks hypotheses (i)–(iv) of Theorem 5.6, with scaling $\lambda_N = \sigma N^{1/3}$. We begin with hypothesis (iv), the Airy limits (5.58)–(5.59) for the point values of the scaled kernel. This needs the most work. At the end of the chapter (Proposition 5.17) the estimates derived for the proof of (iv) will be used to verify (i)–(iii).

The following definitions serve to organize some calculations. For $x \in \mathbb{Z}_+$ define the positive constants

$$A_N(x) = \frac{\beta^x}{\alpha^{x+K}} \cdot \frac{(x+K-1)!N!}{x!(N+K-2)!} \cdot \frac{w^{N+K}}{p} \cdot \frac{\sqrt{q}}{\sqrt{w}}. \quad (5.72)$$

For $x \in \mathbb{Z}_+$ and a function g that is bounded and measurable on the sets where the integrations take place define

$$D_N(x, g) = \frac{1}{2\pi i} \int_{\Gamma_r} g(z) a(z)^x b(z)^K z^{-N-1} dz. \quad (5.73)$$

and

$$F_N(x, g) = \mathbf{1}\{r > \sqrt{wq}\} \cdot (-1)^{x+N+1} \int_{\sqrt{wq}}^r g(-s) |a(-s)|^x b(-s)^K \frac{ds}{s^{N+1}}. \quad (5.74)$$

In our derivation of the asymptotics, the following four functions will appear as g :

$$\begin{aligned} g_1(z) &= 1 & g_3(z) &= a(z) \log a(z) \\ g_2(z) &= z - 1 & g_4(z) &= g_2(z)g_3(z) = (z-1)a(z) \log a(z). \end{aligned} \quad (5.75)$$

As always, we take the principal branch of $\log z$.

The asymptotics of the kernel K_N is based on the following representations.

LEMMA 5.7. *For integers $x \neq y$ in \mathbb{N}*

$$K_N(x, y) = A_N(x)^{1/2} A_N(y)^{1/2} \frac{D_N(x, g_1)D_N(y, g_2) - D_N(x, g_2)D_N(y, g_1)}{x - y} \quad (5.76)$$

and on the diagonal

$$\begin{aligned} K_N(x, x) &= A_N(x) [D_N(x-1, g_3)D_N(x, g_2) - D_N(x, g_1)D_N(x-1, g_4) \\ &\quad + F_N(x, g_1)D_N(x, g_2) - F_N(x, g_2)D_N(x, g_1)]. \end{aligned} \quad (5.77)$$

PROOF. Let us temporarily use the shorthand $\int \cdots dz$ for $(2\pi i)^{-1} \int \cdots dz$, and simplify notation in other obvious ways. Begin with

$$\begin{aligned}
& D_N(x, g_1)D_N(y, g_2) - D_N(x, g_2)D_N(y, g_1) \\
&= \int a^x b^K z^{-N-1} dz \int (z-1) a^y b^K z^{-N-1} dz - \int (z-1) a^x b^K z^{-N-1} dz \int a^y b^K z^{-N-1} dz \\
&= \int a^x b^K z^{-N-1} dz \int z a^y b^K z^{-N-1} dz - \int z a^x b^K z^{-N-1} dz \int a^y b^K z^{-N-1} dz \\
&= \int a^x b^K z^{-N-1} dz \int a^y b^K z^{-N} dz - \int a^x b^K z^{-N} dz \int a^y b^K z^{-N-1} dz \\
(5.78) \quad &= -\frac{q^{N-\frac{1}{2}+\frac{x+y}{2}} \alpha^{K+\frac{x+y}{2}} p^K}{w^{N+K-\frac{1}{2}} \beta^{\frac{x+y}{2}} (N-1)! N!} (m_N^K(x) m_{N-1}^K(y) - m_{N-1}^K(x) m_N^K(y)).
\end{aligned}$$

The last equality above came from (5.45). (For $x \in \mathbb{Z}$ the second line (5.46) vanishes.) Next, the constant in front of line (5.78) multiplied by $A_N(x)^{1/2} A_N(y)^{1/2}$ from (5.72) equals

$$\begin{aligned}
& -\frac{q^{N+\frac{x+y}{2}} p^{K-1}}{(N-1)!(N+K-2)!} \left\{ \frac{(x+K-1)!}{x!} \right\}^{1/2} \left\{ \frac{(y+K-1)!}{y!} \right\}^{1/2} \\
&= -\frac{q}{p} \cdot \frac{p^K q^{N-1} (K-1)!}{(N-1)!(N+K-2)!} \cdot \binom{x+K-1}{x}^{1/2} q^{x/2} \binom{x+K-1}{y}^{1/2} q^{y/2} \\
(5.79) \quad &= -\frac{q}{p} \cdot d_{N-1}^{-2} \cdot \mu^K(x)^{1/2} \mu^K(y)^{1/2}.
\end{aligned}$$

In summary, we have shown that

$$\begin{aligned}
& A_N(x)^{1/2} A_N(y)^{1/2} [D_N(x, g_1)D_N(y, g_2) - D_N(x, g_2)D_N(y, g_1)] \\
&= -\frac{q}{pd_{N-1}^2} \mu^K(x)^{1/2} \mu^K(y)^{1/2} [m_N^K(x) m_{N-1}^K(y) - m_{N-1}^K(x) m_N^K(y)].
\end{aligned}$$

Comparison with (5.34) shows that this is exactly (5.76) without the denominator $x-y$.

To prove (5.77) use definitions (5.7) of μ^K and (5.18) of d_N , and formulas (5.48) and (5.50) for m_N and m'_N . Abbreviate temporarily $\chi = \mathbf{1}\{r > \sqrt{wq}\}$.

$$\begin{aligned}
& \frac{q \mu^K(x)}{pd_{N-1}^2} [m'_N(x) m_{N-1}(x) - m_N(x) m'_{N-1}(x)] \\
&= -\frac{q}{p} \cdot \frac{p^K q^{N-1} (K-1)!}{(N-1)!(N+K-2)!} \cdot \binom{x+K-1}{x} q^x \cdot \frac{w^{N+K-\frac{1}{2}} \beta^x (N-1)! N!}{q^{N-\frac{1}{2}+x} \alpha^{x+K} p^K} \\
&\quad \times \left(\left\{ \int \left(\frac{1}{2} \log(\beta/q\alpha) + \log a \right) a^x b^K \frac{dz}{z^{N+1}} + F_N(x, g_1) \right\} \int a^x b^K \frac{dz}{z^N} \right. \\
&\quad \quad \left. - \int a^x b^K \frac{dz}{z^{N+1}} \left\{ \int \left(\frac{1}{2} \log(\beta/q\alpha) + \log a \right) a^x b^K \frac{dz}{z^N} + F_{N-1}(x, g_1) \right\} \right) \\
&= \frac{\beta^x}{\alpha^{x+K}} \cdot \frac{(x+K-1)! N!}{(N+K-2)! x!} \cdot \frac{w^{N+K}}{p} \cdot \frac{\sqrt{q}}{\sqrt{w}} \\
&\quad \times \left(\int (\log a) a^x b^K \frac{dz}{z^{N+1}} \int z a^x b^K \frac{dz}{z^{N+1}} - \int a^x b^K \frac{dz}{z^{N+1}} \int z (\log a) a^x b^K \frac{dz}{z^{N+1}} \right)
\end{aligned}$$

$$\begin{aligned}
& + \chi \cdot (-1)^{x+N+1} \int_{\sqrt{wq}}^r |a(-s)|^x b(-s)^K \frac{ds}{s^{N+1}} \cdot \int z a^x b^K \frac{dz}{z^{N+1}} \\
& \quad - \int a^x b^K \frac{dz}{z^{N+1}} \cdot \chi \cdot (-1)^{x+N} \int_{\sqrt{wq}}^r s |a(-s)|^x b(-s)^K \frac{ds}{s^{N+1}} \\
& = A_N(x) \left(\int (a \log a) a^{x-1} b^K \frac{dz}{z^{N+1}} \int (z-1) a^x b^K \frac{dz}{z^{N+1}} \right. \\
& \quad \left. - \int a^x b^K \frac{dz}{z^{N+1}} \int (z-1) (a \log a) a^{x-1} b^K \frac{dz}{z^{N+1}} \right. \\
& \quad + \chi \cdot (-1)^{x+N+1} \int_{\sqrt{wq}}^r |a(-s)|^x b(-s)^K \frac{ds}{s^{N+1}} \cdot \int (z-1) a^x b^K \frac{dz}{z^{N+1}} \\
& \quad \left. - \int a^x b^K \frac{dz}{z^{N+1}} \cdot \chi \cdot (-1)^{x+N+1} \int_{\sqrt{wq}}^r (-s-1) |a(-s)|^x b(-s)^K \frac{ds}{s^{N+1}} \right) \\
& = A_N(x) [D_N(x-1, g_3) D_N(x, g_2) - D_N(x, g_1) D_N(x-1, g_4) \\
& \quad + F_N(x, g_1) D_N(x, g_2) - F_N(x, g_2) D_N(x, g_1)].
\end{aligned}$$

This proves (5.77). \square

Continuing to set up preliminaries, a frequently appearing constant is

$$(5.80) \quad \rho = \frac{w\sqrt{q}}{(\sqrt{wq}+1)(\sqrt{w}+\sqrt{q})} = \frac{w\sqrt{q}}{p\sqrt{\alpha\beta}}.$$

In terms of ρ ,

$$(5.81) \quad \sigma = \frac{\sqrt{wq}}{p\rho^{2/3}}.$$

At various times it will be convenient to treat some discrete variables as continuous. Introduce the correction $\omega_N'' \in (0, 1]$ such that

$$(5.82) \quad K = M - N + 1 = \lfloor Nw \rfloor - N + 1 = Nw - N + \omega_N'' = N(\alpha - \beta) + \omega_N''.$$

Also write

$$(5.83) \quad x = N\beta + \ell_x$$

where x is a positive integer. We need to allow negative ℓ_x , so we always consider large enough N so that $x > 0$.

The task is to deduce asymptotics for the kernel $K_N(x, y)$ from formulas (5.76) and (5.77). We begin with analysis of $A_N(x)$ and then turn to the more involved analysis of $D_N(x, g)$.

LEMMA 5.8. *For each $-\infty < h_0 < \infty$ there exist finite constants $C = C(h_0)$ and $N_0 = N_0(h_0)$ such that*

$$(5.84) \quad 0 < A_N(N\beta + \ell_x) \leq C N e^{CN^{-1}|\ell_x|} \quad \text{for } N \geq N_0 \text{ and } \ell_x \geq N^{1/3} h_0.$$

Furthermore, for any $h_1 < \infty$, we have the uniform limit

$$(5.85) \quad \lim_{N \rightarrow \infty} \sup_{-h_1 \leq \xi \leq h_1} |N^{-1} A_N(\lfloor N\beta + N^{1/3} \sigma \xi \rfloor) - \rho| = 0.$$

In both statements we consider N large enough so that the argument of A_N is positive.

PROOF. Write $x = N\beta + \ell_x$ as in (5.83) above. Start by rearranging and then apply Stirling's formula:

$$\begin{aligned} A_N(x) &= \frac{\beta^x w^{N+K}}{\alpha^{x+K}} \cdot \frac{(x+K)!N!}{x!(N+K)!} \cdot \frac{(N+K-1)(N+K)}{(x+K)} \cdot \frac{\sqrt{q}}{p\sqrt{w}} \\ &= \left\{ \frac{\beta^x w^{N+K}}{\alpha^{x+K}} \cdot \frac{(x+K)^{x+K} N^N}{x^x (N+K)^{N+K}} \right\} \left\{ \frac{(N+K-1)(N+K)^{1/2} N^{1/2}}{(x+K)^{1/2} x^{1/2}} \cdot \frac{\sqrt{q} e^{\varepsilon(x,N)}}{p\sqrt{w}} \right\} \\ &\equiv A_N^{(1)}(x) \cdot A_N^{(2)}(x). \end{aligned}$$

Above we named the first expression in large brackets $A_N^{(1)}(x)$, the second $A_N^{(2)}(x)$, and $\varepsilon(x, N)$ is the error from Stirling's formula that tends to 0 as both $x, N \rightarrow \infty$. The parameter ranges in (5.84) and (5.85) are such that $x \rightarrow \infty$ follows from $N \rightarrow \infty$.

From (5.82) and (5.83)

$$\begin{aligned} A_N^{(2)}(x) &= \frac{(Nw-1+\omega''_N)(Nw+\omega''_N)^{1/2} N^{1/2}}{(N(\beta+w-1)+\ell_x+\omega''_N)^{1/2} (N\beta+\ell_x)^{1/2}} \cdot \frac{\sqrt{q} e^{\varepsilon(x,N)}}{p\sqrt{w}} \\ &= N \cdot \frac{\left(w - \frac{1-\omega''_N}{N}\right) \left(w + \frac{\omega''_N}{N}\right)^{1/2}}{\left(\beta+w-1 + \frac{\ell_x+\omega''_N}{N}\right)^{1/2} \left(\beta + \frac{\ell_x}{N}\right)^{1/2}} \cdot \frac{\sqrt{q} e^{\varepsilon(x,N)}}{p\sqrt{w}}. \end{aligned}$$

This shows two things: for large enough N ,

$$(5.86) \quad A_N^{(2)}(x) \leq CN \text{ for } \ell_x \geq N^{1/3} h_0 \text{ as required for (5.84),}$$

and

$$(5.87) \quad N^{-1} A_N^{(2)}(x) \rightarrow \frac{w^{3/2}}{(\beta+w-1)^{1/2} \beta^{1/2}} \cdot \frac{\sqrt{q}}{p\sqrt{w}} = \frac{w\sqrt{q}}{p\sqrt{\alpha\beta}} = \rho$$

uniformly over bounded $N^{-1/3} \ell_x$ (as required in (5.85)).

Turning to $A_N^{(1)}(x)$, rewrite it as

$$\begin{aligned} A_N^{(1)}(x) &= \left(\frac{N\beta}{x}\right)^x \left(\frac{x+K}{N\alpha}\right)^{x+K} \cdot \left(\frac{Nw}{N+K}\right)^{N+K} \\ &= \left(\frac{N\beta}{x}\right)^x \left(\frac{N\alpha+\ell_x}{N\alpha}\right)^{x+K} \left(\frac{x+K}{N\alpha+\ell_x}\right)^{x+K} \cdot \left(\frac{Nw}{N+K}\right)^{N+K} \\ (5.88) \quad &= \left(\frac{N\beta}{x}\right)^x \left(\frac{N\alpha+\ell_x}{N\alpha}\right)^{N\alpha+\ell_x} \left(\frac{N\alpha+\ell_x}{N\alpha}\right)^{\omega''_N} \cdot \left(\frac{x+K}{N\alpha+\ell_x}\right)^{x+K} \cdot \left(\frac{Nw}{N+K}\right)^{N+K} \end{aligned}$$

For the last two factors factor on line (5.88), utilizing the expansion $\log(1+y) = y + O(y^2)$ for small real y :

$$\begin{aligned} \left(\frac{x+K}{N\alpha+\ell_x}\right)^{x+K} \left(\frac{Nw}{N+K}\right)^{N+K} &= \left(\frac{N\alpha+\ell_x+\omega''_N}{N\alpha+\ell_x}\right)^{N\alpha+\ell_x+\omega''_N} \left(\frac{Nw}{Nw+\omega''_N}\right)^{Nw+\omega''_N} \\ &= \left(1 + \frac{\omega''_N}{N\alpha+\ell_x}\right)^{N\alpha+\ell_x+\omega''_N} \left(1 - \frac{\omega''_N}{Nw+\omega''_N}\right)^{Nw+\omega''_N} \\ &= e^{\omega''_N + O(N^{-1})} \cdot e^{-\omega''_N + O(N^{-1})} = e^{O(N^{-1})}. \end{aligned}$$

Thus this quantity is irrelevant for the statements (5.84) and (5.85) we are in the process of proving.

For the third last factor on line (5.88):

$$(5.89) \quad \left(\frac{N\alpha+\ell_x}{N\alpha}\right)^{\omega''_N} = e^{\omega''_N \log(1+(N\alpha)^{-1}\ell_x)} \leq e^{(N\alpha)^{-1}|\ell_x|}.$$

The last expression shows that this factor is bounded by $e^{CN^{-1}|\ell_x|}$ in the parameter range of (5.84). The middle expression above shows that this factor converges to 1 uniformly in the parameter range of (5.85).

It remains to treat the first two factors on line (5.88). Introduce the variable $\zeta = \ell_x/N$ and write

$$(5.90) \quad \left(\frac{N\beta}{x}\right)^x \left(\frac{N\alpha + \ell_x}{N\alpha}\right)^{N\alpha + \ell_x} = \left(\frac{\beta}{\beta + \zeta}\right)^{N(\beta + \zeta)} \left(\frac{\alpha + \zeta}{\alpha}\right)^{N(\alpha + \zeta)} = e^{Nf(\zeta)}$$

with

$$f(\zeta) = (\beta + \zeta) \log \frac{\beta}{\beta + \zeta} + (\alpha + \zeta) \log \frac{\alpha + \zeta}{\alpha}.$$

The admissible values are $\zeta > -\beta$ which corresponds to $x > 0$.

First check that, due to $\alpha \geq \beta$, $f'(\zeta) \leq 0$ iff $\zeta \geq 0$. Thus $f(0) = 0$ is a global maximum and consequently the factors in (5.90) are irrelevant for the bound claimed in (5.84). Statement (5.84) is thereby proved, its right-hand side coming from (5.86) and (5.89).

As $\zeta \rightarrow 0$, $f(0) = f'(0) = 0$ implies that $f(\zeta) \geq -C\zeta^2$ for a constant C . Thus in the parameter range of (5.85),

$$1 \geq e^{Nf(\zeta)} \geq e^{-CN\zeta^2} \geq e^{-CN^{-1}\ell_x^2} \geq e^{-CN^{-1/3}}.$$

This together with the previous estimates shows that $A_N^{(1)}(x) \rightarrow 1$ uniformly in the range required by (5.85). In combination with (5.87) this proves (5.85). \square

We turn to work on $D_N(x, g)$. Define

$$(5.91) \quad u(z) = \beta \log(\sqrt{wq} + z) - \alpha \log(\sqrt{w} + z\sqrt{q}) - \log z$$

where the logarithms take their principal value. This function is holomorphic for $z \in \mathbb{C} \setminus \mathbb{R}_-$ (the complex plane minus the nonpositive real axis). For a fixed $r \in (0, 1)$ such that $r \neq \sqrt{wq}$, the arguments of the logarithms are bounded and bounded away from zero for z on Γ_r . Thus $u(z)$ is well-defined and bounded on the set $\Gamma_r \setminus \{-r\}$. Write

$$(5.92) \quad \begin{aligned} D_N(x, g) &= \frac{1}{2\pi i} \int_{\Gamma_r} g(z) a(z)^{N\beta + \ell_x} b(z)^{N(\alpha - \beta) + \omega'_N} \frac{dz}{z^{N+1}} \\ &= \frac{1}{2\pi i} \int_{\Gamma_r} g(z) e^{N(u(z) - u(1)) + \ell_x \log a(z) + \omega'_N \log b(z)} z^{-1} dz. \end{aligned}$$

For the part of integral (5.92) close to $z = r$ we develop an expansion for the function u of (5.91). The precise value of the radius $1/4$ around $z = 1$ taken in the next lemma is immaterial, any fixed small radius would do.

LEMMA 5.9. *For z such that $|z - 1| \leq 1/4$,*

$$(5.93) \quad u(z) - u(1) = \frac{1}{3}\rho(1 - z)^3 + \rho(1 - z)^4 v(z)$$

for a function v that satisfies $|v(z)| \leq B_u$ with a fixed constant B_u . The constant B_u can even be taken independent of w and q .

PROOF. Differentiation and algebraic manipulation lead to

$$\begin{aligned}
u'(z) &= \frac{\beta}{\sqrt{wq} + z} - \frac{\alpha\sqrt{q}}{\sqrt{w} + z\sqrt{q}} - \frac{1}{z} = \frac{-w\sqrt{q}(1-z)^2}{z(\sqrt{wq} + z)(\sqrt{w} + z\sqrt{q})} \\
&= -\rho(1-z)^2 + (1-z)^2 \left(\rho - \frac{w\sqrt{q}}{z(\sqrt{wq} + z)(\sqrt{w} + z\sqrt{q})} \right) \\
&= -\rho(1-z)^2 - \rho(1-z)^3 \cdot \frac{z^2\sqrt{q} + (\sqrt{w} + q\sqrt{w} + \sqrt{q})z + (\sqrt{wq} + 1)(\sqrt{w} + \sqrt{q})}{z(\sqrt{wq} + z)(\sqrt{w} + z\sqrt{q})} \\
&\equiv -\rho(1-z)^2 - \rho(1-z)^3 f(z).
\end{aligned}$$

The last equality defines the function $f(z)$. It is evident that the denominator of f is bounded away from 0 for example under any restriction of the type $\Re z \geq h$ for any $h > 0$, while its numerator is bounded on any bounded set. There is a bound on f that is independent of $w \geq 1$ and $q \in (0, 1)$ since for large w the dominant term in both numerator and denominator is w .

Given z such that $0 < |1-z| \leq 1/4$, integrate along the line segment \mathcal{L} from 1 to z , parametrized by $\zeta(s) = 1 + se^{i\lambda}$, $0 \leq s \leq |1-z|$, for the appropriate argument λ :

$$u(z) - u(1) = \int_{\mathcal{L}} u'(\zeta) d\zeta = \frac{1}{3}\rho(1-z)^3 + \rho e^{i4\lambda} \int_0^{|1-z|} s^3 f(\zeta(s)) ds.$$

Let $B_u = \frac{1}{4} \sup |f(z)|$ over the radius 1/4 disk centered at 1. Then the function

$$G(z) = e^{i4\lambda} \int_0^{|1-z|} s^3 f(\zeta(s)) ds$$

satisfies

$$\left| \frac{G(z)}{(1-z)^4} \right| \leq \frac{4B_u}{|1-z|^4} \int_0^{|1-z|} s^3 ds = B_u.$$

We can satisfy (5.93) by taking $v(z) = (1-z)^{-4}G(z)$.

Regularity of v is not of consequence to us. But (5.93) does show that v is holomorphic in a punctured neighborhood around 1, and by boundedness of v the point $z = 1$ is a removable singularity [Rud87, Theorem 10.21]. Thus v can be assumed holomorphic in the open disk of radius 1/4 around 1. \square

As we take the limit $N \nearrow \infty$, we also take the radius $r \nearrow 1$. So we set $r = 1 - \delta$ where $\delta \searrow 0$ as $N \nearrow \infty$. The main contribution to the integral (5.92) will come from a small neighborhood around the point $z = 1$. To this end we split it in two parts

$$(5.94) \quad D_N(x, g) = I_1 + I_2$$

where

$$(5.95) \quad I_1 = \frac{1}{2\pi} \int_{-\varepsilon}^{\varepsilon} g(re^{i\theta}) e^{N(u(re^{i\theta}) - u(1)) + \ell_x \log a(re^{i\theta}) + \omega_N'' \log b(re^{i\theta})} d\theta$$

and

$$(5.96) \quad I_2 = \frac{1}{2\pi} \int_{(-\pi, \pi) \setminus (-\varepsilon, \varepsilon)} g(re^{i\theta}) a(re^{i\theta})^x b(re^{i\theta})^K r^{-N} e^{-iN\theta} d\theta.$$

We introduce scaled variables (η, ν) such that

$$(5.97) \quad \delta = \frac{\eta}{(\rho N)^{1/3}} \quad \text{and} \quad \theta = \frac{\nu}{(\rho N)^{1/3}}.$$

Here $\eta > 0$ is a constant. The integration limit above is

$$(5.98) \quad \varepsilon = (\rho N)^{-7/24}$$

and so the new integration variable ν has range $-(\rho N)^{1/24} \leq \nu \leq (\rho N)^{1/24}$.

We expand the parts that appear in the exponent of the integrand of I_1 . We use repeatedly these expansions, valid for small real or complex z :

$$(5.99) \quad \log(1+z) = z + O(|z|^2)$$

$$(5.100) \quad \cos z = 1 - \frac{1}{2}z^2 + O(|z|^4) \quad \text{and} \quad \sin z = z + O(|z|^3)$$

In the estimates that follow the O -terms are uniform over the range $0 \leq \eta$, $|\nu| \leq (\rho N)^{1/24}$. First we utilize the scaling to find the main part and order of magnitude of the complex value $1-z$ that appears in expansion (5.93). The first $O(N^{-7/12})$ term below comes from $\theta^2 \leq \varepsilon^2$.

$$(5.101) \quad \begin{aligned} 1 - re^{i\theta} &= 1 - (1 - \delta)(\cos \theta + i \sin \theta) \\ &= 1 - \left(1 - \frac{\eta}{(\rho N)^{1/3}}\right) \left(1 + O(N^{-7/12}) + i \frac{\nu}{(\rho N)^{1/3}}\right) \\ &= \frac{\eta - i\nu}{(\rho N)^{1/3}} + O(N^{-7/12}) = O(N^{-7/24}). \end{aligned}$$

Then we approximate (5.93):

$$(5.102) \quad \begin{aligned} N(u(re^{i\theta}) - u(1)) &= \frac{1}{3}N\rho(1 - re^{i\theta})^3 + N\rho(1 - re^{i\theta})^4\nu(re^{i\theta}) \\ &= \frac{1}{3}(\eta - i\nu)^3 + O(N^{-1/6}). \end{aligned}$$

Next the $\log a(re^{i\theta})$ function in the exponent in the integral (5.95). The principal branch of the logarithm does satisfy $\log z_1 z_2 = \log z_1 + \log z_2$ when z_1 and z_2 lie in the right half plane because then $\arg z_1 + \arg z_2$ stays in $(-\pi, \pi)$. This situation we have on the first line below as soon as N is large enough to make $|\theta|$ small enough.

$$(5.103) \quad \begin{aligned} \log a(re^{i\theta}) &= \log \frac{\sqrt{wq} + re^{i\theta}}{\sqrt{wq} + 1} - \log \frac{\sqrt{w} + re^{i\theta}\sqrt{q}}{\sqrt{w} + \sqrt{q}} \\ &= \log \left(1 - \frac{1 - re^{i\theta}}{\sqrt{wq} + 1}\right) - \log \left(1 - \frac{\sqrt{q} - re^{i\theta}\sqrt{q}}{\sqrt{w} + \sqrt{q}}\right) \\ &= -(1 - re^{i\theta}) \left(\frac{1}{\sqrt{wq} + 1} - \frac{\sqrt{q}}{\sqrt{w} + \sqrt{q}}\right) + O(|1 - re^{i\theta}|^2) \\ &= \frac{-\eta + i\nu}{(\rho N)^{1/3}} \cdot \frac{p\rho}{\sqrt{wq}} + O(N^{-7/12}) \\ &= N^{-1/3}\sigma^{-1}(-\eta + i\nu) + O(N^{-7/12}). \end{aligned}$$

The last equality used (5.81). For the b -function we only record an error:

$$(5.104) \quad \begin{aligned} \log b(re^{i\theta}) &= -\log \frac{\sqrt{w} + re^{i\theta}\sqrt{q}}{\sqrt{w} + \sqrt{q}} = -\log \left(1 - \frac{\sqrt{q} - re^{i\theta}\sqrt{q}}{\sqrt{w} + \sqrt{q}}\right) \\ &= O(|1 - re^{i\theta}|) = O(N^{-7/24}). \end{aligned}$$

For a function g in (5.95) that is holomorphic in a neighborhood of $z = 1$, let $k = k(g)$ be the order of the zero at $z = 1$, so that for z in some open disk around 1

$$g(z) = (k!)^{-1}g^{(k)}(1)(z-1)^k + O((z-1)^{k+1}) \quad \text{with } g^{(k)}(1) \neq 0.$$

Order $k = 0$ means that $g(1) \neq 0$. Consequently the g -factor inside the integral (5.95) satisfies

$$(5.105) \quad \begin{aligned} g(re^{i\theta}) &= g\left(1 - \frac{\eta - i\nu}{(\rho N)^{1/3}} + O(N^{-7/12})\right) \\ &= (\rho N)^{-k/3} \frac{1}{k!} g^{(k)}(1) (-\eta + i\nu)^k + O(N^{-7(k+1)/24}). \end{aligned}$$

To see the order of the error term above, consider separately the cases $k = 0$, $k = 1$ and $k \geq 2$.

The functions g_i in (5.75) are holomorphic in a neighborhood around $z = 1$, and we have the following values for the orders $k(i) = k(g_i)$ and the derivatives:

$$(5.106) \quad \begin{aligned} k(1) &= 0, \quad g_1(1) = 1 & k(3) &= 1, \quad g'_3(1) = \rho p / \sqrt{wq} \\ k(2) &= 1, \quad g'_2(1) = 1 & k(4) &= 2, \quad g''_4(1) = 2\rho p / \sqrt{wq}. \end{aligned}$$

Now we transform the integral I_1 . We begin with the definition (5.95), perform the change of variables $\theta = \nu(\rho N)^{-1/3}$, and then substitute in (5.102), (5.103), (5.104) and (5.105):

$$(5.107) \quad \begin{aligned} I_1(g) &= \frac{1}{2\pi} \int_{-\varepsilon}^{\varepsilon} g(re^{i\theta}) e^{N(u(re^{i\theta}) - u(1)) + \ell_x \log a(re^{i\theta}) + \omega'_N \log b(re^{i\theta})} d\theta \\ &= (\rho N)^{-\frac{1}{3}} \cdot \frac{1}{2\pi} \int_{-(\rho N)^{1/24}}^{(\rho N)^{1/24}} \left\{ (\rho N)^{-\frac{k}{3}} \cdot \frac{g^{(k)}(1)}{k!} (-\eta + i\nu)^k + O(N^{-\frac{7(k+1)}{24}}) \right\} \\ &\quad \times \exp\left\{ \frac{1}{3}(\eta - i\nu)^3 + O(N^{-1/6}) + \ell_x \left(\frac{-\eta + i\nu}{N^{1/3}\sigma} + O(N^{-\frac{7}{12}}) \right) + O(N^{-\frac{7}{24}}) \right\} d\nu \end{aligned}$$

$$(5.108) \quad = (\rho N)^{-\frac{k+1}{3}} \cdot \frac{g^{(k)}(1)}{k!} \cdot I_{11}(k) + O(N^{-\frac{7k+15}{24}}) \cdot I_{11}(0)$$

where the integral $I_{11}(k)$, $k \in \mathbb{Z}_+$, is defined by

$$(5.109) \quad \begin{aligned} I_{11}(k) &= \frac{1}{2\pi} \int_{-(\rho N)^{1/24}}^{(\rho N)^{1/24}} (-\eta + i\nu)^k \\ &\quad \times \exp\left\{ \frac{1}{3}(\eta - i\nu)^3 + \ell_x \left(\frac{-\eta + i\nu}{N^{1/3}\sigma} + O(N^{-\frac{7}{12}}) \right) + O(N^{-1/6}) \right\} d\nu \end{aligned}$$

We are ready to see the Airy functions arise as limits. We remind ourselves of the key definitions. One way to write the Airy function $\text{Ai}(\xi)$ is

$$(5.110) \quad \text{Ai}(\xi) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{\frac{1}{3}(\eta - i\nu)^3 + \xi(-\eta + i\nu)} d\nu$$

where $\eta > 0$ is fixed but arbitrary. (Except for the change of variable $\nu \mapsto -\nu$, this is (4.4).) For a real ξ , the modulus of the integrand is

$$(5.111) \quad e^{\Re\{\frac{1}{3}(\eta - i\nu)^3 + \xi(-\eta + i\nu)\}} = e^{\eta^3/3 - \eta\nu^2 - \xi\eta}$$

which shows that the integral converges, and also shows that, by dominated convergence, differentiation can be taken inside the integral to compute all the derivatives:

$$(5.112) \quad \text{Ai}^{(k)}(\xi) = \frac{1}{2\pi} \int_{-\infty}^{\infty} (-\eta + i\nu)^k e^{\frac{1}{3}(\eta - i\nu)^3 + \xi(-\eta + i\nu)} d\nu.$$

The Airy kernel limit (5.54) for the kernel K_N requires us to scale the variables as $x = \lfloor N\beta + N^{1/3}\sigma\xi \rfloor$, where $\xi \in \mathbb{R}$ is the new variable. Since (5.83) we have represented the positive integer variable x as

$x = N\beta + \ell_x$. To have some room to maneuver with integer parts we make the following assumption that connects the variables x and ξ :

$$(5.113) \quad \begin{aligned} &\ell_x \text{ depends on } N \text{ and } \xi \text{ so that for a fixed constant } C \text{ and all large enough } N, \\ &|\ell_x - N^{1/3}\sigma\xi| \leq C. \end{aligned}$$

LEMMA 5.10. *Fix $\xi \in \mathbb{R}$ and $\eta > 0$. Assume (5.113). Then for $k \in \mathbb{Z}_+$*

$$(5.114) \quad \lim_{N \rightarrow \infty} I_{11}(k) = \text{Ai}^{(k)}(\xi).$$

PROOF. With $\ell_x = N^{1/3}\sigma\xi + O(1)$ the integral $I_{11}(k)$ becomes

$$I_{11}(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{1}_{(-(\rho N)^{1/24}, (\rho N)^{1/24})}(\nu) \cdot (-\eta + i\nu)^k e^{\frac{1}{3}(\eta - i\nu)^3 + \xi(-\eta + i\nu) + O(N^{-1/6})} d\nu$$

Equation (5.111) shows that dominated convergence gives the limit. \square

Returning to line (5.108) we can state the following limit.

COROLLARY 5.11. *Let g be holomorphic in a neighborhood around 1 with a zero of order k at $z = 1$, with $0 \leq k < 7$. Fix $\xi \in \mathbb{R}$ and $\eta > 0$ and assume ℓ_x satisfies (5.113). Then*

$$(5.115) \quad \begin{aligned} \lim_{N \rightarrow \infty} N^{\frac{k+1}{3}} I_1(g) &= \lim_{N \rightarrow \infty} \left\{ \rho^{-\frac{k+1}{3}} \frac{g^{(k)}(1)}{k!} \cdot I_{11}(k) + O(N^{\frac{k-7}{24}}) \cdot I_{11}(0) \right\} \\ &= \rho^{-\frac{k+1}{3}} \frac{1}{k!} g^{(k)}(1) \cdot \text{Ai}^{(k)}(\xi). \end{aligned}$$

The values of k we need are $0 \leq k \leq 2$ from (5.106). The reader may wonder why the order k works against us in the limit (5.115). The reason is that in the expansion (5.105) we bounded the error uniformly by the maximal modulus of $-\eta + i\nu$. For a better bound we could have included the powers of $|-\eta + i\nu|$ in the second integral on line (5.108).

The next step is to extend the limit to the full integral $D_N(x, g)$ by showing that the part I_2 in (5.96) vanishes in the limit. To control the integral I_2 we show that the magnitude of the integrand is dominated by the value taken closest to $z = r$. This next lemma contains a technical step in that direction.

LEMMA 5.12. *Let $0 < r \leq 1$, $w \geq 1$, $0 < \kappa \leq \lambda$ such that*

$$(5.116) \quad \frac{\sqrt{wq} + r}{\sqrt{w} + r\sqrt{q}} \leq \sqrt{\kappa/\lambda} \leq 1,$$

and define the function

$$(5.117) \quad f(\theta) = \kappa \log|\sqrt{wq} + re^{i\theta}| - \lambda \log|\sqrt{w} + re^{i\theta}\sqrt{q}| \quad \text{for } \theta \in [-\pi, \pi].$$

Then for any $\theta_0 \in [0, \pi]$, $f(\theta_0) = f(-\theta_0) \geq f(\theta)$ for $\theta_0 \leq |\theta| \leq \pi$.

PROOF. Rewrite the function as

$$\begin{aligned} f(\theta) &= \frac{\kappa}{2} \log((\sqrt{wq} + r \cos \theta)^2 + r^2 \sin^2 \theta) - \frac{\lambda}{2} \log((\sqrt{w} + r\sqrt{q} \cos \theta)^2 + r^2 q \sin^2 \theta) \\ &= \frac{\kappa}{2} \log(wq + 2rt\sqrt{wq} + r^2) - \frac{\lambda}{2} \log(w + 2rt\sqrt{wq} + r^2q) \equiv h(t) \end{aligned}$$

where we set $t = \cos \theta$ and renamed the function as $h(t)$. The symmetry $f(\theta) = f(-\theta)$ comes from the symmetry of the cosine. The inequality claimed in the lemma follows from showing that $h'(t) \geq 0$ for $-1 \leq t \leq 1$.

$$h'(t) = \frac{\kappa r \sqrt{wq}}{wq + 2rt\sqrt{wq} + r^2} - \frac{\lambda r \sqrt{wq}}{w + 2rt\sqrt{wq} + r^2q}$$

and so

$$h'(t) \geq 0 \iff k(t) \equiv \kappa(w + 2rt\sqrt{wq} + r^2) - \lambda(wq + 2rt\sqrt{wq} + r^2) \geq 0.$$

Now $k'(t) = 2r\sqrt{wq}(\kappa - \lambda) \leq 0$ by the assumption $\kappa \leq \lambda$. Hence it is enough to check $k(1) \geq 0$ which is equivalent to the left-hand inequality of the assumption (5.116). \square

Now the estimate on I_2 .

LEMMA 5.13. *Assume that $\ell_x \geq -C_0 N^{1/3}$ for some constant $0 < C_0 < \infty$. Then for large enough N*

$$(5.118) \quad |I_2(g)| \leq C e^{-c_1 N^{1/12} - c_{2,3} \ell_x N^{-1/3}}$$

for constants $0 < C, c_1 < \infty$ and where $c_{2,3}$ takes two distinct positive values: one if $\ell_x > 0$, the other if $\ell_x < 0$.

PROOF. From the definition (5.96), bounding g by a constant,

$$(5.119) \quad |I_2(g)| \leq C \int_{(-\pi, \pi) \setminus (-\varepsilon, \varepsilon)} |a(re^{i\theta})|^x |b(re^{i\theta})|^K r^{-N} d\theta.$$

With $x = N\beta + \ell_x$ and $K = N(\alpha - \beta) + \omega''_N$ the integrand is

$$|a(re^{i\theta})|^x |b(re^{i\theta})|^K r^{-N} = \frac{|\sqrt{wq} + re^{i\theta}|^{N\beta + \ell_x}}{|\sqrt{w} + re^{i\theta} \sqrt{q}|^{N\alpha + \ell_x}} \cdot \frac{|\sqrt{w} + \sqrt{q}|^{N\alpha + \ell_x}}{|\sqrt{wq} + 1|^{N\beta + \ell_x}} \cdot |b(re^{i\theta})|^{\omega''_N} r^{-N}.$$

We wish to apply Lemma 5.12 to the first quotient above to claim that, over the range $\theta \in (-\pi, -\varepsilon) \cup (\varepsilon, \pi)$, it is maximized at $\theta = \varepsilon$. Now $\kappa = N\beta + \ell_x \leq N\alpha + \ell_x = \lambda$ so condition (5.116) is

$$\frac{\beta + N^{-1}\ell_x}{\alpha + N^{-1}\ell_x} \geq \frac{(\sqrt{wq} + r)^2}{(\sqrt{w} + r\sqrt{q})^2}.$$

Rearranging this and using $r = 1 - \delta$ leads to

$$\frac{\ell_x}{N} \geq -\delta \cdot \frac{w(1+r)(1+q) + 2\sqrt{wq}(w+r)}{(w-r^2)}.$$

Since $\delta = \eta(\rho N)^{-1/3}$, this implies the existence of a constant $0 < c < \infty$ such that (5.116) is satisfied if $\ell_x \geq -cN^{2/3}$. Thus for large enough N Lemma 5.12 applies to give

$$|a(re^{i\theta})|^x |b(re^{i\theta})|^K r^{-N} \leq |a(re^{i\varepsilon})|^{N\beta + \ell_x} |b(re^{i\varepsilon})|^{N(\alpha - \beta)} \cdot |b(re^{i\theta})|^{\omega''_N} r^{-N}.$$

Bound the function $|b(re^{i\theta})|^{\omega''_N}$ by a constant, drop the integration in (5.119) altogether, and recall definition (5.91) of $u(z)$ to get

$$(5.120) \quad \begin{aligned} |I_2(g)| &\leq C |a(re^{i\varepsilon})|^{N\beta + \ell_x} |b(re^{i\varepsilon})|^{N(\alpha - \beta)} |r^{-N}| \\ &= C |a(re^{i\varepsilon})|^{N\beta} |b(re^{i\varepsilon})|^{N(\alpha - \beta)} (re^{i\varepsilon})^{-N} \cdot a(re^{i\varepsilon})^{\ell_x} \\ &= C |e^{N(u(re^{i\varepsilon}) - u(1)) + \ell_x \log a(re^{i\varepsilon})}| \\ &= C e^{N \Re(u(re^{i\varepsilon}) - u(1)) + \ell_x \Re \log a(re^{i\varepsilon})}. \end{aligned}$$

For small ε the logarithms of $a(re^{i\varepsilon})$ and $b(re^{i\varepsilon})$ are well-defined, and hence so are their powers above.

From (5.102)

$$\begin{aligned} N \Re(u(re^{i\theta}) - u(1)) &= \frac{1}{3} \Re\{(\eta - i\nu)^3\} + O(N^{-1/6}) \\ &= \frac{1}{3}\eta^3 - \eta\nu^2 + O(N^{-1/6}). \end{aligned}$$

Take $\theta = \varepsilon = (\rho N)^{-7/24}$ so that $\nu = \theta(\rho N)^{-1/3} = (\rho N)^{1/24}$. Recall that η is a fixed positive constant. Consequently for large enough N and a constant $0 < c_1 < \infty$

$$(5.121) \quad N \Re(u(re^{i\varepsilon}) - u(1)) \leq -c_1 N^{1/12}.$$

Similarly from (5.103)

$$\Re \log a(re^{i\varepsilon}) = \frac{-\eta}{\sigma N^{1/3}} + O(N^{-7/12})$$

and thereby

$$(5.122) \quad -c_2 N^{-1/3} \leq \Re \log a(re^{i\varepsilon}) \leq -c_3 N^{-1/3}$$

for constants $0 < c_2, c_3 < \infty$.

Substituting these back into (5.120) gives

$$|I_2(g)| \leq C e^{-c_1 N^{1/12} - c_{2,3} \ell_x N^{-1/3}}$$

where $c_{2,3}$ takes two distinct positive values: one if $\ell_x > 0$, the other if $\ell_x < 0$. \square

We can obtain the Airy convergence of the full integrals $D_N(x, g)$ defined in (5.92) and earlier in (5.73).

LEMMA 5.14. *Let g be holomorphic in a neighborhood around 1 with a zero of order k at $z = 1$, with $0 \leq k < 7$. Fix $\xi \in \mathbb{R}$ and $\eta > 0$. Let $x = N\beta + \ell_x$ with ℓ_x satisfying (5.113). Then*

$$(5.123) \quad \lim_{N \rightarrow \infty} N^{\frac{k+1}{3}} D_N(x, g) = \rho^{-\frac{k+1}{3}} \frac{1}{k!} g^{(k)}(1) \cdot \text{Ai}^{(k)}(\xi).$$

PROOF. We only need to combine the decomposition $D_N(x, g) = I_1 + I_2$ from (5.94) with limit (5.115) and bound (5.118). \square

The next item is to show that the second part of formula (5.77) for $K_N(x, x)$ that includes the $F_N(x, g)$ -terms is irrelevant for the asymptotics. The argument is basically the same as for (5.118).

LEMMA 5.15. *Suppose $\ell_x \geq -C_1 N^{1/3}$ for some constant $0 < C_1 < \infty$. Then for $x = N\beta + \ell_x \in \mathbb{N}$, if N is large enough,*

$$(5.124) \quad |F_N(x, g)| \leq C e^{-c_1 N^{1/12} - c_{2,3} \ell_x N^{-1/3}}$$

for constants $0 < C, c_1 < \infty$ and where $c_{2,3}$ takes two distinct positive values: one if $\ell_x > 0$, the other if $\ell_x < 0$.

PROOF. Assume $\sqrt{wq} < 1$ and $\delta = 1 - r = \eta(\rho N)^{-1/3}$ small enough so that $\sqrt{wq} < r < 1$, for otherwise $F_N(x, g)$ vanishes. From the definition (5.74), with a bounded function g in the integral, begin with

$$(5.125) \quad |F_N(x, g)| \leq C \int_{\sqrt{wq}}^r |a(-s)|^x b(-s)^K s^{-N} ds.$$

Set temporarily $\zeta = \ell_x/N$ and

$$u_1(s) = (\beta + \zeta) \log(s - \sqrt{wq}) - (\alpha + \zeta) \log(\sqrt{w} - s\sqrt{q}) - \log s, \quad s \in (\sqrt{wq}, 1]$$

and rewrite the integrand as

$$\begin{aligned}
|a(-s)|^x b(-s)^K s^{-N} &= |a(-s)|^{N(\beta+\zeta)} b(-s)^{N(\alpha-\beta)} s^{-N} \cdot b(-s)^{\omega''_N} \\
&= \frac{(s - \sqrt{wq})^{N(\beta+\zeta)}}{(\sqrt{w} - s\sqrt{q})^{N(\alpha+\zeta)} s^N} \cdot \frac{(\sqrt{w} + \sqrt{q})^{N(\alpha+\zeta)}}{(\sqrt{wq} + 1)^{N(\beta+\zeta)}} \cdot b(-s)^{\omega''_N} \\
(5.126) \qquad &= e^{Nu_1(s)} \cdot \frac{(\sqrt{w} + \sqrt{q})^{N(\alpha+\zeta)}}{(\sqrt{wq} + 1)^{N(\beta+\zeta)}} \cdot b(-s)^{\omega''_N}.
\end{aligned}$$

Looking at the factors on the last line, $b(-s)$ is bounded by a constant. One computes

$$u'_1(s) = \frac{w\sqrt{q}(s+1)^2 + \zeta ps\sqrt{w}}{s(s - \sqrt{wq})(\sqrt{w} - s\sqrt{q})}.$$

Thus there is some constant $0 < c < \infty$ such that if $\zeta > -c$ then $u'_1(s) > 0$ for all $s \in (\sqrt{wq}, 1]$. The condition $\zeta > -c$ is the same as $\ell_x > -cN$ which is satisfied under the hypothesis $\ell_x \geq -C_1 N^{1/3}$ if N is large enough. Hence $u_1(s) \leq u_1(r)$ for the range of s in the integral (5.125).

Combine (5.125), (5.126) and the uniform bound $u_1(s) \leq u_1(r)$:

$$\begin{aligned}
|F_N(x, g)| &\leq C e^{Nu_1(r)} \cdot \frac{(\sqrt{w} + \sqrt{q})^{N(\alpha+\zeta)}}{(\sqrt{wq} + 1)^{N(\beta+\zeta)}} \\
&= C \frac{|\sqrt{wq} - r|^{N\beta+\ell_x}}{|\sqrt{w} - r\sqrt{q}|^{N\alpha+\ell_x}} \cdot \frac{(\sqrt{w} + \sqrt{q})^{N\alpha+\ell_x}}{(\sqrt{wq} + 1)^{N\beta+\ell_x}} \cdot r^{-N}.
\end{aligned}$$

Think of $-r$ as $re^{-i\pi}$. Apply Lemma 5.12 with $\kappa = N\beta + \ell_x \leq N\alpha + \ell_x = \lambda$ to the first quotient on the line above, to show that this quotient can only increase if $-r$ is replaced by $re^{i\varepsilon}$. Hypothesis (5.116) of Lemma 5.12 follows from

$$\frac{\sqrt{wq} + r}{\sqrt{w} + r\sqrt{q}} \leq \frac{\sqrt{wq} + 1}{\sqrt{w} + \sqrt{q}} = \sqrt{\frac{\beta}{\alpha}} \leq \sqrt{\frac{N\beta + \ell_x}{N\alpha + \ell_x}} \leq 1.$$

After this step we have

$$\begin{aligned}
|F_N(x, g)| &\leq C \frac{|\sqrt{wq} + re^{i\varepsilon}|^{N\beta+\ell_x}}{|\sqrt{w} + re^{i\varepsilon}\sqrt{q}|^{N\alpha+\ell_x}} \cdot \frac{(\sqrt{w} + \sqrt{q})^{N\alpha+\ell_x}}{(\sqrt{wq} + 1)^{N\beta+\ell_x}} \cdot r^{-N} \\
&= C |a(re^{i\varepsilon})|^{N\beta+\ell_x} |b(re^{i\varepsilon})|^{N(\alpha-\beta)} |re^{i\varepsilon}|^{-N}
\end{aligned}$$

by the definition (5.91) of the function u

$$\begin{aligned}
&= C |e^{N(u(re^{i\varepsilon}) - u(1)) + \ell_x \log a(re^{i\varepsilon})}| \\
&= C e^{N \Re(u(re^{i\varepsilon}) - u(1)) + \ell_x \Re \log a(re^{i\varepsilon})} \\
&\leq C e^{-c_1 N^{1/12} - c_{2,3} \ell_x N^{-1/3}}.
\end{aligned}$$

In the last step we used (5.121) and (5.122). \square

We can prove the main point, namely the convergence of the properly scaled kernel $K_N(x, y)$ to the Airy kernel. Recall that the Airy kernel is defined for $(x, y) \in \mathbb{R}^2$ by

$$(5.127) \qquad A(x, y) = \begin{cases} \frac{\text{Ai}(x)\text{Ai}'(y) - \text{Ai}'(x)\text{Ai}(y)}{x - y}, & x \neq y \\ \text{Ai}'(x)\text{Ai}'(x) - \text{Ai}(x)\text{Ai}''(x), & x = y. \end{cases}$$

PROPOSITION 5.16. *Let $\xi, \eta \in \mathbb{R}$. Let $\ell_\xi = \ell_\xi(N)$ and $\ell_\eta = \ell_\eta(N)$ be quantities such that $x = N\beta + \ell_\xi$ and $y = N\beta + \ell_\eta$ are positive integers for each N , and for some constant C ,*

$$(5.128) \quad |\ell_\xi - N^{1/3}\sigma\xi| + |\ell_\eta - N^{1/3}\sigma\eta| \leq C$$

for all large enough N . Then if $\xi \neq \eta$

$$(5.129) \quad \lim_{N \rightarrow \infty} \sigma N^{1/3} K_N(x, y) = A(\xi, \eta),$$

while on the diagonal

$$(5.130) \quad \lim_{N \rightarrow \infty} \sigma N^{1/3} K_N(x, x) = A(\xi, \xi).$$

PROOF. Suppose first $\xi \neq \eta$. Then

$$x - y = \ell_\xi - \ell_\eta = N^{1/3}\sigma(\xi - \eta) + O(1)$$

and in particular $x \neq y$ for large enough N . Start with formula (5.76), use limits (5.85) and (5.123) as $N \rightarrow \infty$, and note that the denominator $x - y$ takes up one factor of $N^{1/3}$:

$$\begin{aligned} \sigma N^{1/3} K_N(x, y) &= \sigma N^{1/3} A_N(x)^{1/2} A_N(y)^{1/2} \frac{D_N(x, g_1) D_N(y, g_2) - D_N(x, g_2) D_N(y, g_1)}{x - y} \\ &= \left(\frac{A_N(x)}{N} \right)^{1/2} \left(\frac{A_N(y)}{N} \right)^{1/2} \frac{N^{1/3} D_N(x, g_1) N^{2/3} D_N(y, g_2) - N^{2/3} D_N(x, g_2) N^{1/3} D_N(y, g_1)}{\xi - \eta + O(N^{-1/3})} \\ &\longrightarrow \rho^{1/2} \cdot \rho^{1/2} \cdot \frac{\rho^{-1/3} \text{Ai}(\xi) \rho^{-2/3} \text{Ai}'(\eta) - \rho^{-2/3} \text{Ai}'(\xi) \rho^{-1/3} \text{Ai}(\eta)}{\xi - \eta} = A(\xi, \eta). \end{aligned}$$

The orders $k(i)$ and values $g_1(1) = g_2'(1) = 1$ came from (5.106).

Same argument for the diagonal. Note that x can be replaced by $x - 1$ without any change in the limit (5.123). Apply bound (5.124) to show that the terms in $K_N(x, x)$ that involve F_N vanish in the limit.

$$\begin{aligned} \sigma N^{1/3} K_N(x, x) &= \sigma N^{1/3} A_N(x) [D_N(x - 1, g_3) D_N(x, g_2) - D_N(x, g_1) D_N(x - 1, g_4) \\ &\quad + F_N(x, g_1) D_N(x, g_2) - F_N(x, g_2) D_N(x, g_1)] \\ &= \sigma \cdot \frac{A_N(x)}{N} \cdot [N^{2/3} D_N(x - 1, g_3) \cdot N^{2/3} D_N(x, g_2) - N^{1/3} D_N(x, g_1) \cdot N D_N(x - 1, g_4) \\ &\quad + N^{2/3} F_N(x, g_1) \cdot N^{2/3} D_N(x, g_2) - N F_N(x, g_2) \cdot N^{1/3} D_N(x, g_1)] \\ &\longrightarrow \sigma \rho \left[\frac{\rho^{1/3} p}{\sqrt{wq}} \text{Ai}'(\xi) \cdot \rho^{-2/3} \text{Ai}'(\xi) - \rho^{-1/3} \text{Ai}(\xi) \cdot \frac{1}{2} \cdot \frac{2p}{\sqrt{wq}} \text{Ai}''(\xi) \right] \\ &= A(\xi, \xi). \end{aligned}$$

Recall (5.81) to cancel away the unnecessary constants for the last equality. \square

As the last item in the proof of the distributional limit Theorem 5.1 we turn to the auxiliary bounds needed for the convergence of the Fredholm determinant as summarized in Theorem 5.6. In (5.118) and (5.124) we have good bounds for the integrals $I_2(g)$ and $F_N(x, g)$. To derive a bound for $I_1(g)$ we revisit the steps that led to (5.108). Let k be again the order of the zero of g at $z = 1$,

so that $g(z) = (z - 1)^k h(z)$ for a function h holomorphic in a neighborhood of 1. Then utilize expansion (5.101) and recall that η is fixed:

$$(5.131) \quad \begin{aligned} |g(re^{i\theta})| &\leq C |re^{i\theta} - 1|^k \leq C \left| \frac{\eta - i\nu}{(\rho N)^{1/3}} + O(N^{-7/12}) \right|^k \\ &\leq CN^{-k/3} (1 + |\nu|^k). \end{aligned}$$

In the next calculation take the development for the exponent from line (5.107) but retain only its real part, and subsume everything bounded into a constant C that depends on all the parameters, including k :

$$(5.132) \quad \begin{aligned} |I_1(g)| &\leq \frac{1}{2\pi} \int_{-\varepsilon}^{\varepsilon} |g(re^{i\theta})| \cdot e^{\Re\{N(u(re^{i\theta}) - u(1)) + \ell_x \log a(re^{i\theta}) + \omega'_N \log b(re^{i\theta})\}} d\theta \\ &\leq CN^{-\frac{k+1}{3}} \int_{-(\rho N)^{1/24}}^{(\rho N)^{1/24}} (1 + |\nu|^k) e^{-\eta\nu^2 - c_{2,3}N^{-1/3}\ell_x} d\nu \\ &\leq CN^{-\frac{k+1}{3}} e^{-c_{2,3}N^{-1/3}\ell_x}. \end{aligned}$$

Above $c_{2,3}$ is a positive constant that takes one of two values depending on whether ℓ_x is positive or negative. This little complication arises from the (totally irrelevant) term $O(N^{-7/12})$ that multiplies ℓ_x on line (5.107).

For our purposes the order k of the zero is bounded, and so we can combine (5.118) and (5.132) to get the bound

$$(5.133) \quad |D(x, g)| \leq CN^{-\frac{k+1}{3}} e^{-c_{2,3}N^{-1/3}\ell_x}$$

for $x = N\beta + \ell_x \in \mathbb{Z}_+$, as long as $\ell_x \geq -C_0 N^{1/3}$ for some constant C_0 . We can collect the estimates we need for the kernel.

PROPOSITION 5.17. *The following bounds hold.*

(i) *For each $\tau \in \mathbb{R}$ there exist constants $C(\tau), N_0(\tau) < \infty$ such that*

$$\sup_{N \geq N_0(\tau)} \sum_{m=0}^{\infty} K_N(\lfloor N\beta + N^{1/3}\sigma\tau \rfloor + m, \lfloor N\beta + N^{1/3}\sigma\tau \rfloor + m) \leq C(\tau).$$

(ii) *For every $\varepsilon > 0$ there exist finite $L = L(\varepsilon), N_0(\varepsilon) < \infty$ such that*

$$\sup_{N \geq N_0(\varepsilon)} \sum_{m=0}^{\infty} K_N(\lfloor N\beta + N^{1/3}\sigma L \rfloor + m, \lfloor N\beta + N^{1/3}\sigma L \rfloor + m) \leq \varepsilon.$$

(iii) *For each $M < \infty$ there exists $N_0(M) < \infty$,*

$$\sup_{-M \leq \xi \leq M} \sup_N N^{1/3} K_N(\lfloor N\beta + N^{1/3}\sigma\xi \rfloor, \lfloor N\beta + N^{1/3}\sigma\xi \rfloor) < \infty.$$

PROOF. Combine bounds (5.84), (5.124) and (5.133) with formula (5.77) for K_N on the diagonal and the values $k(i)$ from (5.106). Then for $x = N\beta + \ell_x$ with $\ell_x \geq -C_0N^{1/3}$,

$$\begin{aligned}
 K_N(x, x) &\leq A_N(x) [|D_N(x-1, g_3)| \cdot |D_N(x, g_2)| + |D_N(x, g_1)| \cdot |D_N(x-1, g_4)| \\
 &\quad + |F_N(x, g_1)| \cdot |D_N(x, g_2)| + |F_N(x, g_2)| \cdot |D_N(x, g_1)|] \\
 &\leq CN e^{CN^{-1}|\ell_x|} [(N^{-2/3} \cdot N^{-2/3} + N^{-1/3} \cdot N^{-1}) e^{-c_{2,3}N^{-1/3}\ell_x} \\
 &\quad + (e^{-c_1N^{1/12}}N^{-2/3} + e^{-c_1N^{1/12}}N^{-1/3}) e^{-c_{2,3}N^{-1/3}\ell_x}] \\
 (5.134) \quad &\leq CN^{-1/3} e^{-c_{2,3}N^{-1/3}\ell_x}.
 \end{aligned}$$

Above we took N large enough so that $e^{-c_1N^{1/12}} \leq N^{-1}$ and $CN^{-2/3} < c_{2,3}/2$ (conditions independent of ℓ_x) and then redefined the constant $c_{2,3}$.

The bound claimed in (iii) is now immediate.

For the bound claimed in (i), let $x = \lfloor N\beta + N^{1/3}\sigma\tau \rfloor + m$. Then for some constant $0 < c_4 < \infty$,

$$-\ell_x = -\lfloor N\beta + N^{1/3}\sigma\tau \rfloor + N\beta - m \leq c_4N^{1/3} - m.$$

Substitute this back into line (5.134), and then

$$\begin{aligned}
 &\sum_{m=0}^{\infty} K_N(\lfloor N\beta + N^{1/3}\sigma\tau \rfloor + m, \lfloor N\beta + N^{1/3}\sigma\tau \rfloor + m) \\
 &\leq CN^{-1/3} \sum_{m=0}^{\infty} e^{-c_{2,3}N^{-1/3}m} \leq \frac{CN^{-1/3}}{1 - e^{-c_{2,3}N^{-1/3}}} \leq C.
 \end{aligned}$$

Bound (ii) follows by a similar summation. □

Propositions 5.16 and 5.17 verify the hypotheses of Theorem 5.6 and thereby establish the Airy kernel limit for the Meixner kernel. Theorem 5.6 applied to the representation (5.16) proves the Tracy-Widom limit of Theorem 5.1.

Comments

This chapter follows closely Johansson's paper [Joh00].

APPENDIX A

Probability theory

This appendix collects some ideas and results from probability theory for easy reference.

Convergence in probability. The following situation occurs in the text. There is a sequence of probability spaces $(\Omega_N, \mathcal{F}_N, \mathbf{P}_N)$ and on each Ω_N a random variable X_N . Let c be a constant. Then we say that $X_N \rightarrow c$ in probability if $\mathbf{P}_N\{|X_N - c| \geq \varepsilon\} \rightarrow 0$ as $N \rightarrow \infty$, for each $\varepsilon > 0$.

In textbook formulations of convergence in probability $X_N \xrightarrow{P} X$ the random variables X_N and the limit variable X are on the same probability space. Since the limit we have is a constant, we can achieve this by defining the product space

$$(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P}) = \left(\prod_N \Omega_N, \bigotimes_N \mathcal{F}_N, \bigotimes_N \mathbf{P}_N \right)$$

with generic sample point $\tilde{\omega} = (\omega_N)$, and the random variables $\tilde{X}_N(\tilde{\omega}) = X_N(\omega_N)$ and $c(\tilde{\omega}) = c$.

Equality in distribution. Let S be a measurable space, X an S -valued random variable defined on a probability space (Ω, \mathcal{F}, P) , and Y another S -valued random variable defined on a possibly different probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$. Equality in distribution $X \stackrel{d}{=} Y$ means that the probability distributions agree: $P\{X \in B\} = \tilde{P}\{Y \in B\}$ for all measurable subsets B of S .

LEMMA A.1. *Suppose $X_n \rightarrow \gamma \in [-\infty, \infty]$ a.s. and the limit γ is a deterministic constant. Let $\{Y_n\}$ be random variables all defined on the same probability space, and such that for each n $Y_n \stackrel{d}{=} X_n$. Then there exists a subsequence n_j such that $Y_{n_j} \rightarrow \gamma \in [-\infty, \infty]$ a.s.*

PROOF. If $\gamma > -\infty$ pick real $c_j < \gamma$ so that $c_j \nearrow \gamma$. By the convergence of X_n we can pick an increasing subsequence $n_1 < n_2 < n_3 < \dots$ so that for each $j \in \mathbb{N}$,

$$P\{Y_{n_j} \leq c_j\} = P\{X_{n_j} \leq c_j\} \leq 2^{-j}.$$

Then by the Borel-Cantelli lemma $\underline{\lim} Y_{n_j} \geq \gamma$. If $\gamma = \infty$ this is enough for the conclusion. If $\gamma < \infty$, repeat the argument from above γ . \square

Subadditive ergodic theory. Here is a version of the subadditive ergodic theorem, originally due to Kingman. This improved version was proved by Liggett. Proofs can be found in textbooks, for example [Dur04] and [Kal02].

Let $\{X_{m,n} : m, n \in \mathbb{Z}_+, 0 \leq m < n\}$ be a real-valued process that satisfies these assumptions.

- (i) $X_{0,n} \leq X_{0,m} + X_{m,n}$ for $0 \leq m < n$.
- (ii) For each $k \in \mathbb{N}$, the process $\{X_{nk, (n+1)k} : n \in \mathbb{Z}_+\}$ is stationary.
- (iii) The probability distribution of the process $\{X_{m, m+j} : j \in \mathbb{N}\}$ is the same for all $m \in \mathbb{Z}_+$.
- (iv) $E(X_{0,1}^+) < \infty$ and for some $\gamma_0 > -\infty$, $E(X_{0,n}) \geq \gamma_0 n$ for all $n \in \mathbb{N}$.

THEOREM A.2. *There is a limit*

$$X = \lim_{n \rightarrow \infty} \frac{X_{0,n}}{n} \quad \text{almost surely and in } L^1.$$

The expectation of X exists and satisfies

$$E(X) = \lim_{n \rightarrow \infty} \frac{E(X_{0,n})}{n} = \inf_{n \rightarrow \infty} \frac{E(X_{0,n})}{n}.$$

If all the stationary processes in assumption (ii) above are ergodic, then X is constant: $P\{X = EX\} = 1$.

At times it is convenient to know that for nonnegative superadditive processes moment assumptions are not needed for almost sure convergence. Let $\{Z_{m,n} : m, n \in \mathbb{Z}_+, 0 \leq m < n\}$ be a process that satisfies $0 \leq Z_{m,n} \leq \infty$, assumptions (ii) and (iii) from above, and superadditivity: $Z_{0,n} \geq Z_{0,m} + Z_{m,n}$ for $0 \leq m < n$. Assume also that the processes $\{Z_{nk,(n+1)k} : n \in \mathbb{Z}_+\}$ are ergodic in addition to stationary.

COROLLARY A.3. *There exists a constant $\gamma \in [0, \infty]$ such that $n^{-1}Z_{0,n} \rightarrow \gamma$ almost surely.*

PROOF. For $K \in \mathbb{N}$, the process $Z_{m,n}^{(K)} = Z_{m,n} \wedge K(m-n)$ is superadditive, and $X_{m,n} = -Z_{m,n}^{(K)}$ satisfies all the assumptions of Theorem A.2, including the ergodicity of the processes in assumption (ii). Thus there are constants $\gamma^{(K)}$ such that $n^{-1}Z_{0,n}^{(K)} \rightarrow \gamma^{(K)}$ almost surely. Since we are considering countably many $K \in \mathbb{N}$, there is a probability one event Ω_0 on which this convergence holds for all $K \in \mathbb{N}$. Let $\gamma = \sup_K \gamma^{(K)}$. We claim that $n^{-1}Z_{0,n} \rightarrow \gamma$ on Ω_0 .

Since $Z_{0,n} \geq Z_{0,n}^{(K)}$ for all K , by letting $n \rightarrow \infty$ along a suitable subsequence and then $K \nearrow \infty$ gives $\liminf_{n \rightarrow \infty} n^{-1}Z_{0,n} \geq \gamma$.

If $\gamma = \infty$ this already gives the limit. Suppose $\gamma < \infty$. If $\overline{\lim}_{n \rightarrow \infty} n^{-1}Z_{0,n} > \gamma$ then pick $\varepsilon > 0$ and a subsequence n_j such that $n_j^{-1}Z_{0,n_j} > \gamma + \varepsilon$ for all j . Pick $K > \gamma + \varepsilon$. Then on the one hand

$$n_j^{-1}Z_{0,n_j}^{(K)} = (n_j^{-1}Z_{0,n_j}) \wedge K > \gamma + \varepsilon \quad \text{for all } j$$

but on the other hand $n_j^{-1}Z_{0,n_j}^{(K)} \rightarrow \gamma^{(K)} \leq \gamma$. This contradiction implies that $\overline{\lim}_{n \rightarrow \infty} n^{-1}Z_{0,n} \leq \gamma$. \square

Large deviations. Let $S_n = X_1 + \dots + X_n$ be a sum of i.i.d. mean zero ($EX_i = 0$) random variables. Assume the existence of some exponential moment, that is, the existence of $\theta_0 > 0$ such that

$$(A.1) \quad \phi(t) = E(e^{tX}) < \infty \quad \text{for } t \in [-\theta_0, \theta_0].$$

Here X has the same distribution as the X_i . Estimate $\phi(t)$ for $|t| \leq \theta_0$ by expanding the exponential:

$$\begin{aligned} |\phi(t) - 1 - \frac{1}{2}t^2E(X^2)| &= \left| t^2 \sum_{k=3}^{\infty} \frac{t^{k-2}E(X^k)}{k!} \right| \\ &\leq t^2 \cdot \theta_0^{-2} \sum_{k=3}^{\infty} \frac{\theta_0^k E(|X|^k)}{k!} \leq t^2 \theta_0^{-2} E(e^{\theta_0|X|}) \\ &\leq t^2 \theta_0^{-2} (E(e^{\theta_0 X}) + E(e^{-\theta_0 X})) \leq Ct^2. \end{aligned}$$

Here is a useful estimate.

LEMMA A.4. *For $0 < x < \infty$ there exists a function $A(x) > 0$ such that $A(x) \rightarrow \infty$ as $x \rightarrow \infty$ and for all $n \geq 1$ and $\alpha \in (0, 1/2]$,*

$$(A.2) \quad P\{|S_n| \geq xn^{1/2+\alpha}\} \leq 2e^{-A(x)n^{2\alpha}}.$$

PROOF. We derive first an upper tail bound for S_n . Let $0 < t \leq \theta_0$, apply an exponential Chebyshev inequality and then $\phi(t) \leq 1 + Ct^2$:

$$P\{S_n \geq xn^{1/2+\alpha}\} \leq \exp[-txn^{1/2+\alpha} + n \log \phi(t)] \leq \exp[-txn^{1/2+\alpha} + Cnt^2].$$

Let $t^* = xn^{\alpha-1/2}/(2C)$. If $t^* \leq \theta_0$, then take $t = t^*$ above to get

$$P\{S_n \geq xn^{1/2+\alpha}\} \leq \exp[-x^2n^{2\alpha}/(4C)].$$

If $t^* > \theta_0$, then take $t = \theta_0$ to get

$$\begin{aligned} P\{S_n \geq xn^{1/2+\alpha}\} &\leq \exp[-\theta_0xn^{1/2+\alpha} + Cn\theta_0^2] \leq \exp[-\theta_0xn^{1/2+\alpha} + Cn\theta_0 \cdot xn^{\alpha-1/2}/(2C)] \\ &= \exp[-\theta_0xn^{1/2+\alpha}/2] \leq \exp[-\theta_0xn^{2\alpha}/2]. \end{aligned}$$

The same argument can be applied to $-X_i$ to get the complementary lower tail bound. \square

APPENDIX B

Linear algebra and determinants

Bijections of the set $[N] = \{1, 2, \dots, N\}$ are called *permutations* and they form the symmetric group $\mathbf{S}(N)$. A permutation τ is a *transposition* if it interchanges 2 elements of $[N]$ and fixes the rest: for some $k \neq \ell$ in $[N]$, $\tau(k) = \ell$, $\tau(\ell) = k$ and $\tau(i) = i$ for $i \in [N] \setminus \{k, \ell\}$.

Every permutation can be expressed as a composition of transpositions. The *signum* (*sign*, *parity*) of a permutation is $\text{sgn}(\sigma) = (-1)^m$ where m is any integer such that σ can be expressed as a composition of m transpositions. The number m is not unique but its parity (even or odd) is. Here is an alternative way to express this same definition: the value $\text{sgn}(\sigma) \in \{\pm 1\}$ is determined uniquely by the following identity, valid for arbitrary variables x_1, \dots, x_N :

$$(B.1) \quad \prod_{1 \leq i < j \leq N} (x_{\sigma(j)} - x_{\sigma(i)}) = \text{sgn}(\sigma) \prod_{1 \leq i < j \leq N} (x_j - x_i).$$

The determinant above is the Vandermonde determinant

$$(B.2) \quad \det_{i,j \in [n]} (x_j^{i-1}) = \prod_{1 \leq i < j \leq n} (x_j - x_i).$$

This identity is proved by induction on n . Subtract x_1 times row $n-1$ from row n , then x_1 times row $n-2$ from row $n-1$, and so on, until the determinant has turned into

$$\begin{vmatrix} 1 & 1 & \cdots & 1 & \cdots & 1 \\ 0 & x_2 - x_1 & \cdots & x_j - x_1 & \cdots & x_n - x_1 \\ 0 & x_2^2 - x_1 x_2 & \cdots & x_j^2 - x_1 x_j & \cdots & x_n^2 - x_1 x_n \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & x_2^{n-1} - x_1 x_2^{n-2} & \cdots & x_j^{n-1} - x_1 x_j^{n-2} & \cdots & x_n^{n-1} - x_1 x_n^{n-2} \end{vmatrix}$$

Expand along the first column and note that the remaining determinant equals $\prod_{j=2}^n (x_j - x_1)$ times a Vandermonde of smaller order. For a vector $x = (x_1, \dots, x_n)$ we use the shorthand

$$(B.3) \quad \Delta_n(x) = \prod_{1 \leq i < j \leq n} (x_j - x_i).$$

In the sequel all matrices have complex entries unless otherwise indicated.

Let $A = [a_{i,j}]_{i,j \in [N]}$ be an $N \times N$ matrix. For subsets $\alpha, \beta \subseteq [N]$ the submatrix $A(\alpha, \beta)$ of A is obtained by deleting from A row k and column ℓ for all $k \notin \alpha$ and $\ell \notin \beta$. The notation is

$$A(\alpha, \beta) = [a_{i,j}]_{i \in \alpha, j \in \beta}.$$

If $\alpha = \beta$ then $A(\beta, \beta)$ is called a *principal submatrix* of A . The *principal minors* are the determinants $\det[A(\beta, \beta)]$ of principal submatrices.

LEMMA B.1. *For an $N \times N$ matrix A and an identity matrix I of the same dimensions,*

$$(B.4) \quad \det(I + A) = 1 + \sum_{\emptyset \neq \beta \subseteq [N]} \det A(\beta, \beta)$$

$$(B.5) \quad = 1 + \sum_{m=1}^N \frac{1}{m!} \sum_{(i_1, \dots, i_m) \in [N]^m} \det_{k, \ell \in [m]} [a_{i_k, i_\ell}].$$

PROOF. Let $A = [a_1, a_2, \dots, a_N]$ express A decomposed into columns. By the multilinearity of the determinant as a function of the columns,

$$(B.6) \quad \det(I + A) = \det[e_1 + a_1, e_2 + a_2, \dots, e_N + a_N] = \sum_{\beta \subseteq [N]} \det B_\beta$$

where $B_\beta = [b_1, b_2, \dots, b_N]$ has columns

$$b_j = \begin{cases} a_j, & j \in \beta \\ e_j, & j \notin \beta. \end{cases}$$

Since the e_j columns are zero except at the diagonal,

$$\det B_\beta = \sum_{\sigma \in \mathbf{S}(N)} \operatorname{sgn}(\sigma) \prod_{j=1}^N b_{\sigma(j), j} = \sum_{\sigma|_{\beta^c} = \operatorname{id}} \operatorname{sgn}(\sigma) \prod_{j \in \beta} a_{\sigma(j), j}.$$

Permutations of the set β are in one-to-one correspondence with permutations $\sigma \in \mathbf{S}(N)$ that fix the set β^c , and parity is preserved by this identification since the same transpositions appear in the compositions. Consequently $\det B_\beta = \det A(\beta, \beta)$ for $\beta \neq \emptyset$, while $\det B_\emptyset = 1$. The right-hand sides of (B.4) and (B.6) coincide.

To go from (B.4) to (B.5), write each set β of cardinality m in each of its $m!$ permutations, and note that $\det_{k, \ell \in [m]} [a_{i_k, i_\ell}] = 0$ if the vector (i_1, \dots, i_m) has any repetitions. \square

Here is a generalized form of the Cauchy-Binet identity.

THEOREM B.2. *Let (X, \mathcal{F}, μ) be a σ -finite measure space, $N \in \mathbb{N}$, and $f_1, \dots, f_N, g_1, \dots, g_N$ measurable functions on X such that $f_i g_j \in L^1(\mu)$ for all pairs i, j . Then*

$$(B.7) \quad \det_{i, j \in [N]} \left[\int_X f_i(x) g_j(x) \mu(dx) \right] = \frac{1}{N!} \int_{X^N} \det_{i, j \in [N]} [f_i(x_j)] \det_{i, j \in [N]} [g_i(x_j)] \mu^{\otimes N}(dx_{1, N}).$$

The integral on the right is over the N -fold product measure $\mu^{\otimes N}$, and the integrand is integrable.

PROOF. The absolute value of the integrand on the right-hand side of (B.7) is bounded by a sum of terms of the type

$$\prod_{i=1}^N |f_{\sigma(i)}(x_i) g_{\tau(i)}(x_i)|$$

where $\sigma, \tau \in \mathbf{S}(N)$ are permutations. These products are integrable under $\mu^{\otimes N}$ by the assumption $f_i g_j \in L^1(\mu)$.

To prove (B.7), start by expanding the right-hand side via the definition of the determinant. Note that $\operatorname{sgn}(\tau) = \operatorname{sgn}(\tau^{-1})$ for a permutation τ . Use the product structure of the multidimensional

integral, and rename suitably multiplication and summation indices.

$$\begin{aligned}
 & \frac{1}{N!} \sum_{\sigma, \tau \in \mathbf{S}(N)} \operatorname{sgn}(\sigma) \operatorname{sgn}(\tau) \int_{X^N} \prod_{i=1}^N f_{\sigma(i)}(x_i) \cdot \prod_{i=1}^N g_{\tau^{-1}(i)}(x_i) \mu^{\otimes N}(dx_{1,N}) \\
 &= \frac{1}{N!} \sum_{\sigma, \tau \in \mathbf{S}(N)} \operatorname{sgn}(\sigma) \operatorname{sgn}(\tau) \prod_{i=1}^N \int_X f_{\sigma(i)}(x) g_{\tau^{-1}(i)}(x) \mu(dx) \\
 &= \frac{1}{N!} \sum_{\sigma \in \mathbf{S}(N)} \sum_{\tau \in \mathbf{S}(N)} \operatorname{sgn}(\sigma \circ \tau) \prod_{j=1}^N \int_X f_{\sigma \circ \tau(j)}(x) g_j(x) \mu(dx) \\
 &= \frac{1}{N!} \sum_{\sigma \in \mathbf{S}(N)} \sum_{\rho \in \mathbf{S}(N)} \operatorname{sgn}(\rho) \prod_{j=1}^N \int_X f_{\rho(j)}(x) g_j(x) \mu(dx) \\
 &= \det_{i,j \in [N]} \left[\int_X f_i(x) g_j(x) \mu(dx) \right].
 \end{aligned}$$

On the third-last line, for each fixed σ the composite $\rho = \sigma \circ \tau$ ranges over the symmetric group $\mathbf{S}(N)$. Dependence on σ disappears from the terms and so the averaging over σ can be dropped. \square

Let us also state the Cauchy-Binet identity for matrices. Let A be an $M \times K$ matrix, B a $K \times N$ matrix, and $C = AB$ the $M \times N$ product.

COROLLARY B.3. *Let $r \leq M \wedge N$, and $\alpha \subseteq [M]$ and $\beta \subseteq [N]$ subsets of size r . Then*

$$(B.8) \quad \det C(\alpha, \beta) = \sum_{\gamma \subseteq [K]: |\gamma|=r} \det A(\alpha, \gamma) \det B(\gamma, \beta).$$

The equality is true also if $|K| < r$ in which case the sum on the right is empty and returns the value zero.

PROOF. Enumerate the sets as $\alpha = \{k_1 < \dots < k_r\}$ and $\beta = \{\ell_1 < \dots < \ell_r\}$. Then

$$\det C(\alpha, \beta) = \det_{i,j \in [r]} [c(k_i, \ell_j)] = \det_{i,j \in [r]} \left[\sum_{x=1}^K a(k_i, x) b(x, \ell_j) \right].$$

Apply (B.7) with $X = [K]$ and μ counting measure to get

$$\det C(\alpha, \beta) = \frac{1}{r!} \sum_{x_1, \dots, x_r \in [K]} \det_{i,j \in [r]} [a(k_i, x_j)] \det_{i,j \in [r]} [b(x_i, \ell_j)].$$

Any repetition among the x_1, \dots, x_r leads to repeated rows and columns and thereby to zero determinants on the right-hand side. Thus $\gamma = \{x_1, \dots, x_r\}$ is a set of cardinality r , and each such set appears in the sum in all its permutations, $r!$ times. Permuting the x_j 's to order them leaves the term on the right-hand side unchanged because the same permutation operates on both determinants and thus multiplies by the square of the signum. We can rewrite the above as

$$\det C(\alpha, \beta) = \sum_{\gamma = \{x_1 < \dots < x_r\} \subseteq [K]} \det_{i,j \in [r]} [a(k_i, x_j)] \det_{i,j \in [r]} [b(x_i, \ell_j)]$$

which equals the right-hand side of (B.8). \square

Next Hadamard's inequality. An $n \times n$ Hermitian matrix ($A = A^*$) is *positive definite* if $\mathbf{x}^* A \mathbf{x} > 0$ for all nonzero vectors $\mathbf{x} \in \mathbb{C}^n$, and *positive semidefinite* (or *nonnegative definite*) if $\mathbf{x}^* A \mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{C}^n$. Equivalently, a Hermitian matrix is positive definite if all its eigenvalues

are strictly positive, and positive semidefinite if all its eigenvalues are nonnegative. In particular, a positive semidefinite matrix is positive definite iff it is nonsingular.

THEOREM B.4. (a) For a Hermitian, positive semidefinite $n \times n$ matrix A ,

$$(B.9) \quad \det A \leq \prod_{i=1}^n a_{i,i}.$$

(b) For any $n \times n$ matrix B ,

$$(B.10) \quad |\det B| \leq \prod_{i=1}^n \left(\sum_{j=1}^n |b_{i,j}|^2 \right)^{1/2}.$$

PROOF. (a) If A is singular then $\det A = 0$ and (B.9) holds because $a_{i,i} = \mathbf{e}_i^* A \mathbf{e}_i \geq 0$.

Suppose A is nonsingular. Then A is positive definite, and each diagonal element is strictly positive: $a_{i,i} = \mathbf{e}_i^* A \mathbf{e}_i > 0$.

We reduce the proof to the case where all diagonal elements are equal to 1. Let D be the diagonal matrix with entries $d_{i,i} = 1/\sqrt{a_{i,i}}$. Then DAD has diagonal elements all 1, and $\det A \leq \prod a_{i,i}$ iff $\det DAD \leq 1$.

Lastly the case $a_{1,1} = a_{2,2} = \dots = a_{n,n} = 1$. Let $\lambda_1, \dots, \lambda_n$ denote the eigenvalues which are positive now. By Jensen's inequality (in the guise of the arithmetic-geometric mean inequality)

$$\det A = \prod \lambda_i \leq \left(\frac{1}{n} \sum \lambda_i \right)^n = \left(\frac{1}{n} \operatorname{tr} A \right)^n = \left(\frac{1}{n} \sum a_{i,i} \right)^n = 1.$$

(b) Part (a) applied to the positive semidefinite matrix $A = BB^*$ gives

$$|\det B|^2 = \det(BB^*) \leq \prod_i (BB^*)_{i,i} = \prod_i \sum_j b_{i,j} \bar{b}_{i,j} = \prod_i \sum_j |b_{i,j}|^2. \quad \square$$

The Fredholm determinant of a finite-rank operator on $\ell^2(S)$ for some countable set S . Let a_m and b_m , $m \in [M]$, be functions on S and the kernel given by

$$(B.11) \quad K(x, y) = \sum_{m=1}^M a_m(x) b_m(y).$$

If $a_m, b_m \in \ell^2$ then K is an operator on ℓ^2 . Assume at least

$$(B.12) \quad \sum_{x \in S} |a_\ell(x) b_m(x)| < \infty \quad \forall \ell, m \in [M].$$

PROPOSITION B.5. Let the operator K on $\ell^2(S)$ be given by (B.11) and assume (B.12). Then the Fredholm determinant satisfies

$$(B.13) \quad \det(I - K)_{\ell^2(S)} = \det_{\ell, m \in [M]} \left[\delta_{\ell, m} - \sum_{x \in S} a_\ell(x) b_m(x) \right].$$

PROOF. By the expansion of the Fredholm determinant,

$$(B.14) \quad \det(I - K) = 1 + \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \sum_{x_1, \dots, x_n \in S} \det_{i, j \in [n]} [K(x_i, x_j)].$$

In the determinant

$$\det_{i, j \in [n]} [K(x_i, x_j)] = \det_{i, j \in [n]} \left[\sum_{m=1}^M a_m(x_i) b_m(x_j) \right]$$

column j is a linear combination of M vectors:

$$\sum_{m=1}^M b_m(x_j) \begin{bmatrix} a_m(x_1) \\ a_m(x_2) \\ \vdots \\ a_m(x_n) \end{bmatrix}.$$

Thus the rank is at most M , and in the sum on line (B.14) all terms for $n > M$ vanish. Apply Cauchy-Binet (B.7) with $X = [M]$ to rewrite (B.14) as

$$\det(I - K) = 1 + \sum_{n=1}^M \frac{(-1)^n}{n!} \sum_{x_1, \dots, x_n \in S} \frac{1}{n!} \sum_{m_1, \dots, m_n \in [M]} \det_{i,j \in [n]} [a_{m_j}(x_i)] \det_{i,j \in [n]} [b_{m_j}(x_i)].$$

Take the sum over (x_1, \dots, x_n) inside and apply Cauchy-Binet (B.7) again, this time with $X = S$ and counting measure:

$$\begin{aligned} \det(I - K) &= 1 + \sum_{n=1}^M \frac{(-1)^n}{n!} \sum_{m_1, \dots, m_n \in [M]} \frac{1}{n!} \sum_{x_1, \dots, x_n \in S} \det_{i,j \in [n]} [a_{m_j}(x_i)] \det_{i,j \in [n]} [b_{m_j}(x_i)] \\ &= 1 + \sum_{n=1}^M \frac{(-1)^n}{n!} \sum_{m_1, \dots, m_n \in [M]} \det_{i,j \in [n]} \left[\sum_{x \in S} a_{m_i}(x) b_{m_j}(x) \right]. \end{aligned}$$

The last determinant vanishes unless the m_1, \dots, m_n are distinct. The factor $n!$ in the denominator takes care of the repeated permutations of a given ordered set $\{m_1 < \dots < m_n\}$. Thus the last line rewrites as

$$\det(I - K) = 1 + \sum_{\emptyset \neq \alpha \subseteq [M]} \det_{\ell, m \in \alpha} \left[- \sum_{x \in S} a_\ell(x) b_m(x) \right] = \det_{\ell, m \in [M]} \left[\delta_{\ell, m} - \sum_{x \in S} a_\ell(x) b_m(x) \right].$$

The last equality is by (B.4). □

APPENDIX C

Analysis

C.1. Convex functions

This section discusses the duality of convex functions in one dimension. A function $f : \mathbb{R} \rightarrow (-\infty, \infty]$ is *convex* if

$$(C.1) \quad f(sx + (1-s)w) \leq sf(x) + (1-s)f(w) \quad \text{for } x, w \in \mathbb{R} \text{ and } 0 < s < 1.$$

Here are some basic facts. For any convex f , the set $\{f < \infty\}$ is an interval. If f is convex and finite on an open interval I then f is continuous on I , and at each $x \in I$ the left and right derivatives exist:

$$f'(x-) = \lim_{h \searrow 0} \frac{f(x-h) - f(x)}{-h} \quad \text{and} \quad f'(x+) = \lim_{h \searrow 0} \frac{f(x+h) - f(x)}{h}.$$

The derivatives are nondecreasing: $f'(x-) \leq f'(x+) \leq f'(w-)$ for $x < w$ in I . The graph of f lies at or above the tangent lines: if $c \in [f'(x-), f'(x+)]$ then

$$(C.2) \quad f(w) \geq f(x) + c(w-x) \quad \text{for all } w \in \mathbb{R}.$$

A form of regularity that arises naturally in the theory of convex functions is lower semicontinuity: a function f from a metric space S into $[-\infty, \infty]$ is *lower semicontinuous* if the sets $\{f > t\}$ are open subsets of S for all $t \in \mathbb{R}$, or, equivalently, if $\liminf_{w \rightarrow x} f(w) \geq f(x)$ for each $x \in S$. A supremum of a collection of continuous functions is lower semicontinuous.

The *convex dual* of a function $f : \mathbb{R} \rightarrow [-\infty, \infty]$ is

$$(C.3) \quad f^*(y) = \sup_{x \in \mathbb{R}} \{xy - f(x)\}, \quad y \in \mathbb{R}.$$

Repeating the step gives the *convex double dual* of f :

$$(C.4) \quad f^{**}(x) = \sup_{y \in \mathbb{R}} \{xy - f^*(y)\}, \quad x \in \mathbb{R}.$$

THEOREM C.1. *Let $f : \mathbb{R} \rightarrow (-\infty, \infty]$ be a lower semicontinuous convex function. Then $f^{**} = f$.*

PROOF. Let us first dispose of the case $f \equiv \infty$. In this case $f^* \equiv -\infty$ and then $f^{**} \equiv \infty$. For the remainder assume that $\{f < \infty\}$ is nonempty, and hence a nonempty interval but possibly a singleton.

For all x ,

$$f^{**}(x) = \sup_{y \in \mathbb{R}} \{xy - f^*(y)\} \leq \sup_{y \in \mathbb{R}} \{xy - (xy - f(x))\} = f(x).$$

We argue the converse $f^{**}(x) \geq f(x)$ by cases.

Case 1: x is an isolated point of $\{f < \infty\}$. Since $\{f < \infty\}$ is an interval, the singleton $\{x\}$ must be all of $\{f < \infty\}$. Then $f^*(y) = xy - f(x)$, and for all $w \in \mathbb{R}$,

$$f^{**}(w) = \sup_y \{(w-x)y + f(x)\} = f(x) \cdot \mathbf{1}\{w=x\} + \infty \cdot \mathbf{1}\{w \neq x\} = f(w).$$

For the remainder of the proof we can assume $\{f < \infty\}$ has nonempty interior.

Case 2: x is an interior point of $\{f < \infty\}$. Pick any $c \in [f'(x-), f'(x+)]$. Then (C.2) implies that $f^*(c) = cx - f(x)$ and from this

$$f^{**}(x) \geq xc - f^*(c) = f(x).$$

Case 3: x is a boundary point of $\{f < \infty\}$ and $\{f < \infty\}$ has nonempty interior. Pick a sequence $\{x_j\}$ of interior points of $\{f < \infty\}$ such that $x_j \rightarrow x$ monotonically. We claim that

$$f^{**}(x) \geq \overline{\lim}_{j \rightarrow \infty} f^{**}(x_j) \geq \underline{\lim}_{j \rightarrow \infty} f(x_j) \geq f(x).$$

The middle inequality follows from Case 2. By lower semicontinuity, $f(x) \leq \underline{\lim} f(x_j)$. Pick a subsequence $\{x_{j_k}\}$ so that $f^{**}(x_{j_k}) \rightarrow \overline{\lim} f^{**}(x_j)$. Fix an index ℓ . By the assumption of monotone convergence $x_{j_k} \rightarrow x$, for each $k > \ell$ there exists $s_k \in [0, 1)$ such that $x_{j_k} = (1 - s_k)x_{j_\ell} + s_k x$ and $s_k \nearrow 1$. Then by convexity

$$f^{**}(x_{j_k}) \leq (1 - s_k)f^{**}(x_{j_\ell}) + s_k f^{**}(x)$$

and by letting $k \rightarrow \infty$, $\overline{\lim} f^{**}(x_j) \leq f^{**}(x)$.

Case 4: x is an interior point of $\{f = \infty\}$. Suppose x is to the right of $\{f < \infty\}$. (An analogous argument works for x to the left of $\{f < \infty\}$.) Pick a point x_1 between x and $\{f < \infty\}$, so that $x_1 < x$ and $f(x_1) = \infty$. Also, pick x_0 in the interior of $\{f < \infty\}$ and a slope c so that $f(w) \geq f(x_0) + c(w - x_0) \equiv b + cw$ for all $w \in \mathbb{R}$. Now observe that for any $\lambda > 0$,

$$f(w) \geq \lambda(w - x_1) + b + cw.$$

The reason is that $f(w) < \infty$ implies $w - x_1 < 0$. From the above inequality

$$f^*(c + \lambda) \leq -b + \lambda x_1$$

and then

$$f^{**}(x) \geq (c + \lambda)x - f^*(c + \lambda) \geq cx + \lambda(x - x_1) + b.$$

Letting $\lambda \nearrow \infty$ shows $f^{**}(x) = \infty$.

We have covered all the cases and the proof is complete. \square

This theorem is equivalent to the statement that a lower semicontinuous convex $f : \mathbb{R} \rightarrow (-\infty, \infty]$ is equal to the supremum of its affine minorants:

$$(C.5) \quad f(x) = \sup\{ax + b : a, b \in \mathbb{R}; \forall z \in \mathbb{R}, az + b \leq f(z)\}.$$

In Section 2.2 we apply duality to concave functions, so we state a corollary for concave functions. A function $g : \mathbb{R} \rightarrow [-\infty, \infty]$ is *concave* if $-g$ is convex. A function g from a metric space S into $[-\infty, \infty]$ is *upper semicontinuous* if $-g$ is lower semicontinuous. The *concave dual* and *concave double dual* of a function $g : \mathbb{R} \rightarrow [-\infty, \infty]$ are

$$(C.6) \quad g^*(y) = \inf_{x \in \mathbb{R}} \{xy - g(x)\}, \quad y \in \mathbb{R},$$

and

$$(C.7) \quad g^{**}(x) = \inf_{y \in \mathbb{R}} \{xy - g^*(y)\}, \quad x \in \mathbb{R}.$$

COROLLARY C.2. *Let $g : \mathbb{R} \rightarrow [-\infty, \infty)$ be an upper semicontinuous concave function and g^{**} its concave double dual defined by (C.7). Then $g^{**} = g$.*

Analogues of these facts hold in higher dimensions, including infinite dimensions. The reader is referred to the monographs of Rockafellar [Roc70] and Ekeland and Temam [ET99].

C.2. Complex variables

The principal branch of the logarithm is the holomorphic function

$$\log z = \log|z| + i \arg z \quad \text{for } z \in \mathbb{C} \setminus \mathbb{R}_- \text{ and with } -\pi < \arg z < \pi.$$

We can write

$$\log|z| + i \arg z = \int_1^{|z|} \frac{dt}{t} + \int_0^{\arg z} \frac{i|z|e^{is}}{|z|e^{is}} ds = \int_1^z \frac{d\zeta}{\zeta}$$

where we integrate first along the line segment from 1 to $|z|$ and then along the circular arc from $|z|$ to $|z|e^{i \arg z}$. That the function $f(z) = \int_1^z \zeta^{-1} d\zeta$ is holomorphic can be seen from the definition of the integral:

$$\frac{f(z) - f(w)}{z - w} = \frac{1}{z - w} \int_w^z \frac{d\zeta}{\zeta} = \int_0^1 \frac{dt}{w + t(z - w)} \xrightarrow{z \rightarrow w} \int_0^1 \frac{dt}{w} = \frac{1}{w}.$$

Any holomorphic branch of the logarithm, that is, a holomorphic f in a domain G such that $e^{f(z)} = z$, satisfies $f'(z) = 1/z$. Note that $0 \notin G$ because $e^w \neq 0$ for all $w \in \mathbb{C}$.

Let γ be the circle of radius $r > 0$ around the origin and f a continuous function on γ^* . We follow the convention of [Rud87] that $\gamma^* \subseteq \mathbb{C}$ denotes the image set of a path $\gamma : [a, b] \rightarrow \mathbb{C}$. Then for any $n \in \mathbb{Z}$ it is legitimate to use the principal branch to write

$$(C.8) \quad \int_{\gamma} f(z) z^n dz = \int_{\gamma} f(z) e^{n \log z} dz$$

simply because with the principal branch $e^{n \log z} = r^n e^{int}$ for $z = re^{it}$ with $-\pi < t < \pi$. The fact that the integrand is not well-defined at the single point $z = -1$ of γ^* is immaterial for the integral.

For any real α the function $(1+z)^\alpha = e^{\alpha \log(1+z)}$ is holomorphic in the open unit disk $\{|z| < 1\}$. The logarithm is the principal branch. On the open unit disk this function has the series

$$(C.9) \quad (1+z)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} z^n.$$

The binomial coefficient above is defined by

$$\binom{\alpha}{0} = 1 \quad \text{and} \quad \binom{\alpha}{n} = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!} \quad \text{for } n \in \mathbb{N}.$$

We record some simple bounds on these coefficients. For $\alpha \in \mathbb{R}$ and $n \geq 1$,

$$(C.10) \quad \left| \binom{\alpha}{n} \right| = \prod_{j=1}^n \left| \frac{\alpha}{j} - \frac{j-1}{j} \right| \leq \prod_{j=1}^n \left(1 + \frac{|\alpha|}{j} \right) = e^{\sum_{j=1}^n \log(1+|\alpha|/j)} \leq e^{|\alpha| \sum_{j=1}^n 1/j} \\ \leq e^{|\alpha|(1+\log n)} = (ne)^{|\alpha|}.$$

Alternatively, one can reverse the product in the numerator and follow similar steps to get this bound:

$$(C.11) \quad \left| \binom{\alpha}{n} \right| = \prod_{j=1}^n \left| \frac{\alpha-n}{j} + 1 \right| \leq \prod_{j=1}^n \left(1 + \frac{|\alpha-n|}{j} \right) \leq (ne)^{|\alpha-n|}.$$

To verify (C.9) let $f(z)$ denote the series on the right. Bound (C.10) shows that the radius of convergence is at least 1. Inside the radius of convergence the series can be differentiated term by term, resulting in the equation $(1+z)f'(z) = \alpha f(z)$. Multiply through this equation by $e^{-\alpha \log(1+z)}$, where \log is the principal branch, to derive

$$\frac{d}{dz} (f(z) e^{-\alpha \log(1+z)}) = 0$$

from which follows $f(z) = ce^{\alpha \log(1+z)}$ for some constant. The value $f(0) = 1$ chooses $c = 1$.

In the next lemma we write explicitly the correction term to an integral that comes from integrating across a branch cut of the logarithm. Let Γ_r denote the counterclockwise circle of radius r around the origin. For a real ξ let $z^\xi = e^{\xi \log z}$ denote the principal value defined for $z \in \mathbb{C} \setminus \mathbb{R}_-$.

LEMMA C.3. *Let $0 < r_1 < \alpha < r_2$. Let f be holomorphic in an annulus $\{z : s_1 < |z| < s_2\}$ that contains the circles of radii r_1 and r_2 . Let $\xi > -1$ be a real number. Then*

$$(C.12) \quad \int_{\Gamma_{r_1}} f(z)(\alpha+z)^\xi dz = \int_{\Gamma_{r_2}} f(z)(\alpha+z)^\xi dz + 2i \sin(\pi\xi) \int_\alpha^{r_2} f(-s)|\alpha-s|^\xi ds.$$

Note that $(\alpha+z)^\xi$ is holomorphic for $z \in \mathbb{C} \setminus (-\infty, -\alpha]$, so in particular in an open region that contains the path Γ_{r_1} of the integral on the left. In the Γ_{r_2} -integral on the right $(\alpha+z)^\xi = e^{\xi \log|\alpha+z| + i\xi \arg(\alpha+z)}$ is well-defined and bounded on $\Gamma_{r_2} \setminus \{-r_2\}$, hence the existence and finiteness of the integral is not in doubt. The last integral on the right is an ordinary Lebesgue or Riemann integral over the interval $[\alpha, r_2]$.

PROOF OF LEMMA C.3. Let $\varepsilon > 0$ be small and let $\tilde{\varepsilon} \in (0, \pi/2)$ be the small number such that $r_2 \sin \tilde{\varepsilon} = \varepsilon$. Let $\tilde{r}_2 = r_2 \cos \tilde{\varepsilon}$. By Cauchy's theorem

$$\int_{\Gamma_{r_1}} f(z)(\alpha+z)^\xi dz = \left(\int_{\Gamma_{r_2}^\varepsilon} + \int_{\gamma_1} + \int_{\mathcal{C}} + \int_{\gamma_2} \right) f(z)(\alpha+z)^\xi dz$$

where

- (i) $\Gamma_{r_2}^\varepsilon$ is the circular path $r_2 e^{i\theta}$ for $-\pi + \tilde{\varepsilon} < \theta < \pi - \tilde{\varepsilon}$,
- (ii) γ_1 is the line segment from $-\tilde{r}_2 + i\varepsilon$ to $-\alpha + i\varepsilon$,
- (iii) \mathcal{C} is the small circular arc $-\alpha + \varepsilon e^{i(\pi/2-\theta)}$ for $0 \leq \theta \leq \pi$,
- (iv) γ_2 is the line segment from $-\alpha - i\varepsilon$ to $-\tilde{r}_2 - i\varepsilon$.

The boundedness of the integrands implies that as $\varepsilon \rightarrow 0$, the integral over $\Gamma_{r_2}^\varepsilon$ converges to the integral over Γ_{r_2} . The integral over \mathcal{C} tends to zero because the length of the path is $\pi\varepsilon$ and the integrand is bounded by $C\varepsilon^{-\xi} \leq C\varepsilon^{-1+\delta}$ for some $\delta > 0$.

The inverse of path γ_1 is $s \mapsto -s + i\varepsilon$ for $\alpha \leq s \leq \tilde{r}_2$, and so

$$\int_{\gamma_1} f(z)(\alpha+z)^\xi dz = \int_\alpha^{\tilde{r}_2} f(-s+i\varepsilon)|\alpha-s+i\varepsilon|^\xi \exp\{i\xi \arg(\alpha-s+i\varepsilon)\} ds.$$

Let $\varepsilon \rightarrow 0$ and apply dominated convergence. If $\xi \geq 0$ the integrand is bounded. If $-1 < \xi < 0$ then for some $\delta > 0$, $|\alpha-s+i\varepsilon|^\xi \leq |\alpha-s|^{-1+\delta}$ which is integrable for $s \in (\alpha, r_2)$. In the limit

$$\lim_{\varepsilon \rightarrow 0} \int_{\gamma_1} f(z)(\alpha+z)^\xi dz = \int_\alpha^{r_2} f(-s)|\alpha-s|^\xi e^{i\xi\pi} ds.$$

Similarly with $\gamma_2(s) = -s - i\varepsilon$ for $\alpha \leq s \leq \tilde{r}_2$,

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0} \int_{\gamma_2} f(z)(\alpha+z)^\xi dz \\ &= - \lim_{\varepsilon \rightarrow 0} \int_\alpha^{\tilde{r}_2} f(-s-i\varepsilon)|\alpha-s-i\varepsilon|^\xi \exp\{i\xi \arg(\alpha-s-i\varepsilon)\} ds \\ &= - \int_\alpha^{r_2} f(-s)|\alpha-s|^\xi e^{-i\xi\pi} ds. \end{aligned}$$

Add up the limits:

$$\int_{\alpha}^{r_2} f(-s)|\alpha - s|^{\xi} (e^{i\xi\pi} - e^{-i\xi\pi}) ds = 2i \sin(\xi\pi) \int_{\alpha}^{r_2} f(-s)|\alpha - s|^{\xi} ds. \quad \square$$

APPENDIX D

Orthogonal Polynomials

Let μ be a finite positive Borel measure on \mathbb{R} , not supported on finitely many points. Assume that

$$(D.1) \quad \int_{\mathbb{R}} |x|^n \mu(dx) < \infty \quad \text{for all } n \geq 0.$$

Let $L^2(\mu)$ be the space of real-valued square-integrable Borel functions:

$$f \in L^2(\mu) \iff \int_{\mathbb{R}} f^2 d\mu < \infty.$$

$L^2(\mu)$ is a Hilbert space with inner product

$$\langle f, g \rangle = \int_{\mathbb{R}} fg d\mu.$$

Next we orthogonalize the monomials $1, x, x^2, \dots$ and then normalize to generate a countable orthonormal basis of polynomials for $L^2(\mu)$.

THEOREM D.1. (i) *There exists a unique sequence $\{p_n(x)\}_{n \in \mathbb{Z}_+}$ of real polynomials on \mathbb{R} with these properties: each $p_n(x)$ is precisely of degree n , its leading coefficient is positive, and the condition of orthonormality under μ is satisfied:*

$$(D.2) \quad \int_{\mathbb{R}} p_m(x)p_n(x) \mu(dx) = \delta_{m,n} \quad \text{for } m, n \in \mathbb{Z}_+.$$

(ii) *Any polynomial of degree n is a linear combination of p_0, \dots, p_n . For any polynomial f of degree strictly less than n , $\int f(x)p_n(x) \mu(dx) = 0$.*

PROOF. (i) We start by showing that for each n there is a uniquely determined sequence $\{\phi_0, \dots, \phi_n\}$ of polynomials such that each $\phi_k(x)$ is precisely of degree k , its leading coefficient is 1, and orthogonality is satisfied:

$$(D.3) \quad \int \phi_k \phi_\ell d\mu = 0 \quad \text{for } k \neq \ell.$$

Denote the moments of the measure μ by

$$(D.4) \quad c_n = \int x^n \mu(dx), \quad n \in \mathbb{Z}_+.$$

It is clear that if $\{\phi_0, \phi_1\}$ is to satisfy the requirements, then these must be $\phi_0(x) = 1$ and $\phi_1(x) = x - c_1/c_0$. Assume that polynomials $\{\phi_0(x), \dots, \phi_{n-1}(x)\}$ uniquely meet the requirements. Let

$$(D.5) \quad \phi_n(x) = x^n + b_{n,n-1}\phi_{n-1}(x) + b_{n,n-2}\phi_{n-2}(x) + \dots + b_{n,0}\phi_0(x)$$

with coefficients

$$(D.6) \quad b_{n,k} = - \int x^n \phi_k(x) \mu(dx) \cdot \left\{ \int \phi_k(x)^2 \mu(dx) \right\}^{-1}, \quad 0 \leq k \leq n-1.$$

The integral $\int \phi_k^2 d\mu$ cannot vanish because if it did μ would have to be supported on the set of at most k zeroes of ϕ_k , a situation we ruled out at the very outset. The definition shows that ϕ_n is strictly of order n , has leading coefficient 1, and the assumed orthogonality of $\phi_0(x), \dots, \phi_{n-1}(x)$ gives for $0 \leq k \leq n-1$

$$\int \phi_n \phi_k d\mu = \int x^n \phi_k(x) \mu(dx) + b_{n,k} \int \phi_k(x)^2 \mu(dx) = 0.$$

Since each ϕ_k is precisely of degree k , any polynomial of degree $n-1$ can be expressed as a linear combination of $\phi_0, \dots, \phi_{n-1}$. (This can be checked by induction too.) Hence ϕ_n must be given by a formula of the type (D.5), and then the orthogonality requirement forces (D.6). Thus ϕ_n is also uniquely determined.

For $n \geq 0$ let

$$\kappa_n = \left\{ \int \phi_n(x)^2 \mu(dx) \right\}^{-1/2} \in (0, \infty)$$

so that we get orthonormal polynomials with positive leading coefficients by setting

$$(D.7) \quad p_n(x) = \kappa_n \phi_n(x) = \kappa_n x^n + \dots$$

Via (D.7) polynomials $\{p_n\}$ and $\{\phi_n\}$ determine each other with the required properties, hence the uniqueness of $\{\phi_n\}$ implies the uniqueness of $\{p_n\}$.

(ii) The argument for spanning degree n polynomials is the same as already given above. If g is a polynomial of degree n , find the coefficient λ_n so that $g - \lambda_n p_n$ is of degree $n-1$, and use induction. Since a polynomial f of degree at most $n-1$ can be written as a linear combination of p_0, \dots, p_{n-1} , orthogonality of f and p_n follows from the orthogonality (D.2). \square

When p_{-1} is needed for some formulas, set $p_{-1}(x) = 0$. The notation κ_n for the leading coefficient of p_n in (D.7) will be used in the sequel.

EXAMPLE D.2. The three classical orthogonal polynomials are the following. In each case the measure has a density w with respect to Lebesgue measure; that is, $\int f(x) \mu(dx) = \int f(x) w(x) dx$. The exact normalizations in the literature vary somewhat.

(i) *Hermite polynomials* go together with the Gaussian measure with density $w(x) = e^{-x^2}$. In the classic treatise [Sze75] these polynomials $\{H_n\}$ are defined by requiring that leading coefficients be positive and

$$\int_{\mathbb{R}} H_m(x) H_n(x) e^{-x^2} dx = 2^n n! \sqrt{\pi} \delta_{m,n}.$$

(ii) If $w(x) = (1-x)^\alpha (1+x)^\beta$ supported on the interval $[-1, 1]$ with constants $\alpha, \beta > -1$, the orthogonal polynomials are known as *Jacobi polynomials*. The special case $\alpha = \beta = 0$ of uniform measure is the *Legendre polynomials*.

(iii) *Laguerre polynomials* come from weight $w(x) = x^\alpha e^{-x}$ on $[0, \infty)$, with $\alpha > -1$.

If the measure μ is supported on a bounded set then the orthogonal polynomials $\{p_n\}$ form a countable orthonormal basis in $L^2(\mu)$. This is also true for the Hermite and Laguerre polynomials.

We continue with another basic fact about orthogonal polynomials.

PROPOSITION D.3. *Let $\{p_n\}$ be the orthonormal polynomials constructed in Theorem D.1. Then this three term recursion holds for $n \geq 1$:*

$$(D.8) \quad p_n(x) = (A_n x + B_n) p_{n-1}(x) - C_n p_{n-2}(x)$$

with

$$(D.9) \quad A_n = \frac{\kappa_n}{\kappa_{n-1}} \text{ for } n \geq 1 \text{ and } C_n = \frac{\kappa_n \kappa_{n-2}}{\kappa_{n-1}^2} \text{ for } n \geq 2.$$

For C_1 we can take any value.

PROOF. The above choice for A_n makes $p_n(x) - A_n x p_{n-1}(x)$ a polynomial of degree $n-1$, and hence there are coefficients $c_{n,k}$ such that

$$(D.10) \quad p_n(x) - A_n x p_{n-1}(x) = c_{n,n-1} p_{n-1}(x) + c_{n,n-2} p_{n-2}(x) + \cdots + c_{n,0} p_0(x).$$

The case $n=1$ follows already and the value of C_1 is immaterial since $p_{-1} = 0$. Let $n \geq 2$ in the sequel.

Multiply through (D.10) by $p_k(x)$ for $0 \leq k \leq n-3$ and integrate.

$$0 = \int p_k(x) p_n(x) \mu(dx) - A_n \int x p_k(x) p_{n-1}(x) \mu(dx) = c_{n,k} \int p_k(x)^2 \mu(dx) \implies c_{n,k} = 0,$$

utilizing all the orthogonality properties, in particular that, as a polynomial of degree at most $n-2$, $x p_k(x)$ is orthogonal with respect to $p_{n-1}(x)$. Now we know that the recursion (D.8) holds with some B_n and C_n .

Multiply through (D.8) by $p_{n-2}(x)$, integrate, and use orthogonality and orthonormality:

$$0 = \int p_n(x) p_{n-2}(x) \mu(dx) = A_n \int p_{n-1}(x) x p_{n-2}(x) \mu(dx) - C_n.$$

Next, since a polynomial of degree $n-2$ is a linear combination of p_0, \dots, p_{n-2} ,

$$x p_{n-2}(x) = \kappa_{n-2} x^{n-1} + \cdots = \frac{\kappa_{n-2}}{\kappa_{n-1}} p_{n-1}(x) + \sum_{k=0}^{n-2} \lambda_k p_k(x)$$

for some coefficients $\lambda_0, \dots, \lambda_{n-2}$. Substitute this into the previous display and use orthonormality again:

$$0 = A_n \frac{\kappa_{n-2}}{\kappa_{n-1}} - C_n$$

from which we can solve for C_n . □

Define kernels K_n on $\mathbb{R} \times \mathbb{R}$ associated to the orthogonal polynomials $\{p_n\}$ by

$$(D.11) \quad K_n(x, y) = \sum_{k=0}^{n-1} p_k(x) p_k(y).$$

If we define the Fourier coefficients of a function $f \in L^2(\mu)$ with respect to the polynomials $\{p_n\}$ by

$$f_n = \int f(x) p_n(x) \mu(dx)$$

and the partial sums of the Fourier expansion by

$$s_n(x) = f_0 p_0(x) + f_1 p_1(x) + \cdots + f_{n-1} p_{n-1}(x),$$

then

$$s_n(x) = K_n f(x) = \int K_n(x, y) f(y) \mu(dy).$$

In particular, if f is a polynomial of degree $n-1$ then $K_n f = f$.

The next representation for the kernel is known as the *Christoffel-Darboux formula*.

THEOREM D.4. For $n \geq 1$ and $x \neq y$ on \mathbb{R}

$$(D.12) \quad K_n(x, y) = \frac{\kappa_{n-1}}{\kappa_n} \cdot \frac{p_n(x) p_{n-1}(y) - p_{n-1}(x) p_n(y)}{x - y}$$

and on the diagonal

$$(D.13) \quad K_n(x, x) = \frac{\kappa_{n-1}}{\kappa_n} (p'_n(x) p_{n-1}(x) - p'_{n-1}(x) p_n(x)).$$

PROOF. (D.13) follows from (D.12) by letting $x \rightarrow y$. To prove (D.12), use (D.8) for $k \geq 1$:

$$\begin{aligned} p_k(x)p_{k-1}(y) - p_{k-1}(x)p_k(y) &= [(A_k x + B_k)p_{k-1}(x) - C_k p_{k-2}(x)]p_{k-1}(y) \\ &\quad - p_{k-1}(x)[(A_k y + B_k)p_{k-1}(y) - C_k p_{k-2}(y)] \\ &= A_k(x-y)p_{k-1}(x)p_{k-1}(y) + C_k[p_{k-1}(x)p_{k-2}(y) - p_{k-2}(x)p_{k-1}(y)]. \end{aligned}$$

Now use the constants A_k and C_k from (D.9). For $k = 1$ the last expression in brackets vanishes and the identity reads

$$\frac{\kappa_0}{\kappa_1} \cdot \frac{p_1(x)p_0(y) - p_0(x)p_1(y)}{x-y} = p_0(x)p_0(y).$$

For $k \geq 2$

$$\begin{aligned} \frac{\kappa_{k-1}}{\kappa_k} \cdot \frac{p_k(x)p_{k-1}(y) - p_{k-1}(x)p_k(y)}{x-y} &= p_{k-1}(x)p_{k-1}(y) \\ &\quad + \frac{\kappa_{k-2}}{\kappa_{k-1}} \cdot \frac{p_{k-1}(x)p_{k-2}(y) - p_{k-2}(x)p_{k-1}(y)}{x-y}. \end{aligned}$$

Adding these identities for $k = 1, \dots, n$ gives (D.12). \square

We derive an integral formula. Introduce a probability measure ν on \mathbb{R}^n by weighting a product of μ -measures with the square of the Vandermonde and then normalizing so that $\nu(\mathbb{R}^n) = 1$:

$$(D.14) \quad \int_{\mathbb{R}^n} f(x_{1,n}) \nu(dx_{1,n}) = \frac{\int_{\mathbb{R}^n} f(x_{1,n}) \Delta_n(x_{1,n})^2 \mu^{\otimes n}(dx_{1,n})}{\int_{\mathbb{R}^n} \Delta_n(x_{1,n})^2 \mu^{\otimes n}(dx_{1,n})}$$

The test function f above is a bounded Borel function on \mathbb{R}^n .

PROPOSITION D.5. *For any measurable functions f and g on \mathbb{R} for which the integrals below are well-defined,*

$$(D.15) \quad \int_{\mathbb{R}^n} \prod_{j=1}^n f(x_j) \nu(dx_{1,n}) = \det_{i,j \in [n]} \left[\int_{\mathbb{R}} p_{i-1}(x) p_{j-1}(x) f(x) \mu(dx) \right]$$

and

$$(D.16) \quad \begin{aligned} \int_{\mathbb{R}^n} \prod_{j=1}^n (1 + g(x_j)) \nu(dx_{1,n}) \\ = 1 + \sum_{m=1}^n \frac{1}{m!} \int_{\mathbb{R}^m} \left(\prod_{\ell=1}^m g(x_\ell) \right) \det_{k,\ell \in [m]} [K_n(x_k, x_\ell)] \mu^{\otimes m}(dx_{1,m}). \end{aligned}$$

PROOF. First transform the Vandermondes into determinants of orthogonal polynomials p_n . Unless otherwise indicated, indices i, j in all the determinants below range over $[n]$.

$$\Delta_n(x_{1,n}) = \det[x_j^{i-1}] = \prod_{\ell=0}^{n-1} \kappa_\ell^{-1} \cdot \det[\kappa_{i-1} x_j^{i-1}] = \prod_{\ell=0}^{n-1} \kappa_\ell^{-1} \cdot \det[p_{i-1}(x_j)]$$

where the last equality comes by adding suitable multiples of rows $1, \dots, i-1$ to row i , for each $i = 2, \dots, n$.

First work on the unnormalized integral.

$$\begin{aligned}
& \int_{\mathbb{R}^n} \left(\prod_{j=1}^n f(x_j) \right) \Delta_n(x_{1,n})^2 \mu^{\otimes n}(dx_{1,n}) \\
&= \left(\prod_{\ell=0}^{n-1} \kappa_\ell^{-2} \right) \int_{\mathbb{R}^n} \left(\prod_{j=1}^n f(x_j) \right) \det[p_{i-1}(x_j)]^2 \mu^{\otimes n}(dx_{1,n}) \\
&= \left(\prod_{\ell=0}^{n-1} \kappa_\ell^{-2} \right) \int_{\mathbb{R}^n} \det[p_{i-1}(x_j) f(x_j)] \det[p_{i-1}(x_j)] \mu^{\otimes n}(dx_{1,n})
\end{aligned}$$

by the generalized Cauchy-Binet identity (B.7)

$$(D.17) \quad = n! \left(\prod_{\ell=0}^{n-1} \kappa_\ell^{-2} \right) \det_{i,j \in [n]} \left[\int_{\mathbb{R}} p_{i-1}(x) p_{j-1}(x) f(x) \mu(dx) \right].$$

Taking $f = 1$ gives, by the orthonormality of the polynomials $\{p_k\}$

$$(D.18) \quad \int_{\mathbb{R}^n} \Delta_n(x_{1,n})^2 \mu^{\otimes n}(dx_{1,n}) = n! \left(\prod_{\ell=0}^{n-1} \kappa_\ell^{-2} \right).$$

Dividing line (D.17) by (D.18) gives (D.15).

Take $f = 1 + g$ and continue from (D.15) and apply (B.5):

$$\begin{aligned}
& \int_{\mathbb{R}^n} \prod_{j=1}^n (1 + g(x_j)) \nu(dx_{1,n}) = \det \left[\delta_{i,j} + \int_{\mathbb{R}} p_{i-1}(x) p_{j-1}(x) g(x) \mu(dx) \right] \\
&= 1 + \sum_{m=1}^n \frac{1}{m!} \sum_{(i_1, \dots, i_m) \in [n]^m} \det_{k, \ell \in [m]} \left[\int_{\mathbb{R}} p_{i_k-1}(x) p_{i_\ell-1}(x) g(x) \mu(dx) \right]
\end{aligned}$$

by the generalized Cauchy-Binet identity (B.7)

$$= 1 + \sum_{m=1}^n \frac{1}{m!} \sum_{(i_1, \dots, i_m) \in [n]^m} \frac{1}{m!} \int_{\mathbb{R}^m} \det_{k, \ell \in [m]} [p_{i_k-1}(x_\ell)] \det_{k, \ell \in [m]} [p_{i_k-1}(x_\ell) g(x_\ell)] \mu^{\otimes m}(dx_{1,m})$$

by removing the g -factors from the determinant and rearranging

$$= 1 + \sum_{m=1}^n \frac{1}{m!} \int_{\mathbb{R}^m} \left(\prod_{\ell=1}^m g(x_\ell) \right) \frac{1}{m!} \sum_{(i_1, \dots, i_m) \in [n]^m} \det_{k, \ell \in [m]} [p_{i_k-1}(x_\ell)]^2 \mu^{\otimes m}(dx_{1,m})$$

by (B.7) again, this time with summation over $1 \leq i \leq n$ playing the role of integration and x_ℓ indexing the “functions” $p_{i-1}(x_\ell)$ of i

$$= 1 + \sum_{m=1}^n \frac{1}{m!} \int_{\mathbb{R}^m} \left(\prod_{\ell=1}^m g(x_\ell) \right) \det_{k, \ell \in [m]} \left[\sum_{i=0}^{n-1} p_i(x_k) p_i(x_\ell) \right] \mu^{\otimes m}(dx_{1,m}).$$

This is the right-hand side of (D.16). □

REMARK D.6. As in the examples above, often the measure μ appears in the form $\mu(dx) = w(x)\lambda(dx)$, that is, μ is given by a nonnegative *weight function*, or Radon-Nikodym derivative, $0 \leq w(x) < \infty$ relative to a nonnegative, σ -finite background measure λ :

$$\int f(x) \mu(dx) = \int f(x) w(x) \lambda(dx) \quad \text{for all bounded Borel functions } f.$$

Natural choices for λ are Lebesgue measure on the real line and counting measure on the integers.

In this same vein we can define an alternative form of the kernel by including extra factors coming from the weight:

$$(D.19) \quad K_n^{(w)}(x, y) = K_n(x, y) \sqrt{w(x)w(y)} = \sum_{k=0}^{n-1} p_k(x) p_k(y) w(x)^{1/2} w(y)^{1/2}.$$

Kernel $K_n^{(w)}(x, y)$ is not necessarily defined for all x, y but only for those points at which the weights are defined. (For example, the integers if λ is counting measure.) By including the factor $w(x)^{1/2} w(y)^{1/2}$ on the right-hand sides of (D.12) and (D.13) we can write the Christoffel-Darboux formula for $K^{(w)}(x, y)$. Also, since

$$\mu^{\otimes m}(dx_{1,m}) = \left(\prod_{k=1}^m w(x_k)^{1/2} \right) \left(\prod_{\ell=1}^m w(x_\ell)^{1/2} \right) \lambda^{\otimes m}(dx_{1,m}),$$

equation (D.16) can be rewritten by multiplying each row and column of the determinant by a $w^{1/2}(x_k)$ -factor:

$$(D.20) \quad \int_{\mathbb{R}^n} \prod_{j=1}^n (1 + g(x_j)) \nu(dx_{1,n}) \\ = 1 + \sum_{m=1}^n \frac{1}{m!} \int_{\mathbb{R}^m} \left(\prod_{\ell=1}^m g(x_\ell) \right) \det_{k,\ell \in [m]} [K_n^{(w)}(x_k, x_\ell)] \lambda^{\otimes m}(dx_{1,m}).$$

Bibliography

- [BCS06] M. Balázs, E. Cator, and T. Seppäläinen, *Cube root fluctuations for the corner growth model associated to the exclusion process*, Electron. J. Probab. **11** (2006), no. 42, 1094–1132 (electronic). MR MR2268539
- [CEL84] M. G. Crandall, L. C. Evans, and P.-L. Lions, *Some properties of viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc. **282** (1984), no. 2, 487–502. MR 86a:35031
- [CL83] Michael G. Crandall and Pierre-Louis Lions, *Viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc. **277** (1983), no. 1, 1–42. MR 85g:35029
- [DMP91] Anna De Masi and Errico Presutti, *Mathematical methods for hydrodynamic limits*, Lecture Notes in Mathematics, vol. 1501, Springer-Verlag, Berlin, 1991. MR 93k:60234
- [Dur04] Richard Durrett, *Probability: theory and examples*, third ed., Duxbury Advanced Series, Brooks/Cole-Thomson, Belmont, CA, 2004.
- [ET99] Ivar Ekeland and Roger Témam, *Convex analysis and variational problems*, english ed., Classics in Applied Mathematics, vol. 28, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1999, Translated from the French. MR MR1727362 (2000j:49001)
- [Eva98] Lawrence C. Evans, *Partial differential equations*, Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 1998. MR 99e:35001
- [GGK00] Israel Gohberg, Seymour Goldberg, and Nahum Krupnik, *Traces and determinants of linear operators*, Operator Theory: Advances and Applications, vol. 116, Birkhäuser Verlag, Basel, 2000. MR MR1744872 (2001b:47035)
- [GW91] Peter W. Glynn and Ward Whitt, *Departures from many queues in series*, Ann. Appl. Probab. **1** (1991), no. 4, 546–572. MR 92i:60162
- [Ish84] Hitoshi Ishii, *Uniqueness of unbounded viscosity solution of Hamilton-Jacobi equations*, Indiana Univ. Math. J. **33** (1984), no. 5, 721–748. MR 85h:35057
- [Joh00] Kurt Johansson, *Shape fluctuations and random matrices*, Comm. Math. Phys. **209** (2000), no. 2, 437–476. MR 2001h:60177
- [Joh07] ———, *A multi-dimensional Markov chain and the Meixner ensemble*, [arXiv:0707.0098](https://arxiv.org/abs/0707.0098).
- [Kal02] Olav Kallenberg, *Foundations of modern probability*, second ed., Probability and its Applications (New York), Springer-Verlag, New York, 2002. MR MR1876169 (2002m:60002)
- [Kel79] Frank P. Kelly, *Reversibility and stochastic networks*, John Wiley & Sons Ltd., Chichester, 1979, Wiley Series in Probability and Mathematical Statistics. MR MR554920 (81j:60105)
- [KL99] Claude Kipnis and Claudio Landim, *Scaling limits of interacting particle systems*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 320, Springer-Verlag, Berlin, 1999. MR 2000i:60001
- [Mar04] James B. Martin, *Limiting shape for directed percolation models*, Ann. Probab. **32** (2004), no. 4, 2908–2937. MR MR2094434 (2005i:60198)
- [Olv97] Frank W. J. Olver, *Asymptotics and special functions*, AKP Classics, A K Peters Ltd., Wellesley, MA, 1997, Reprint of the 1974 original [Academic Press, New York; MR0435697 (55 #8655)]. MR MR1429619 (97i:41001)
- [Rez91] Fraydoun Rezakhanlou, *Hydrodynamic limit for attractive particle systems on \mathbf{Z}^d* , Comm. Math. Phys. **140** (1991), no. 3, 417–448. MR 93f:82035
- [Roc70] R. Tyrrell Rockafellar, *Convex analysis*, Princeton Mathematical Series, No. 28, Princeton University Press, Princeton, N.J., 1970. MR 43 #445
- [Ros81] H. Rost, *Nonequilibrium behaviour of a many particle process: density profile and local equilibria*, Z. Wahrsch. Verw. Gebiete **58** (1981), no. 1, 41–53. MR 83a:60176
- [Rud87] Walter Rudin, *Real and complex analysis*, third ed., McGraw-Hill Book Co., New York, 1987. MR MR924157 (88k:00002)
- [Sep97] Timo Seppäläinen, *Increasing sequences of independent points on the planar lattice*, Ann. Appl. Probab. **7** (1997), no. 4, 886–898. MR 99a:60116

- [Sep98] T. Seppäläinen, *Hydrodynamic scaling, convex duality and asymptotic shapes of growth models*, Markov Process. Related Fields **4** (1998), no. 1, 1–26. MR 99e:60221
- [Sep99] Timo Seppäläinen, *Existence of hydrodynamics for the totally asymmetric simple K -exclusion process*, Ann. Probab. **27** (1999), no. 1, 361–415. MR 2000i:60116
- [Sep00] ———, *Strong law of large numbers for the interface in ballistic deposition*, Ann. Inst. H. Poincaré Probab. Statist. **36** (2000), no. 6, 691–736. MR 2002j:60186
- [Spi70] Frank Spitzer, *Interaction of Markov processes*, Advances in Math. **5** (1970), 246–290 (1970). MR 42 #3856
- [Spo91] Herbert Spohn, *Large scale dynamics of interacting particles*, Springer-Verlag, Berlin, 1991.
- [Sze75] Gábor Szegő, *Orthogonal polynomials*, fourth ed., American Mathematical Society, Providence, R.I., 1975, American Mathematical Society, Colloquium Publications, Vol. XXIII. MR MR0372517 (51 #8724)
- [TW94] Craig A. Tracy and Harold Widom, *Level-spacing distributions and the Airy kernel*, Comm. Math. Phys. **159** (1994), no. 1, 151–174. MR MR1257246 (95e:82003)

Index

- Airy
 - function, 49, 74
 - kernel, 50
- Cauchy-Binet identity, 88
- Christoffel-Darboux formula, 60, 101
- cluster, 1
- conservation law, 23
- dual
 - concave, 22, 94
 - convex, 93
- entropy solution, 23
- envelope property, 24
- first-in-first-out (FIFO), 6
- flux, 22
- Fredholm determinant, 52, 90
- Gaussian Unitary Ensemble (GUE), 53
- Hadamard's inequality, 52, 65, 89
- Hamilton-Jacobi equation, 22
- Hopf-Lax formula, 23
- logarithm
 - principal branch, 95
- M/M/1 queue, 17
- Meixner polynomials, 56
 - generating function, 57
 - integral representation, 61, 62
 - standard, 57
- normal distribution
 - standard complex, 52
- orthogonal polynomials
 - Christoffel-Darboux formula, 101
 - kernel, 101
 - three term recursion, 100
- queue, 6
 - M/M/1, 17
- shift, 15
- signum of a permutation, 87
- subadditive ergodic theorem, 83
- superadditivity, 8
- TASEP
 - continuous time, 5
 - discrete time, 4
- Tracy-Widom distribution, 52
- traffic jam, 5
- Vandermonde determinant, 87
- viscosity solution, 23
- weak solution, 23
- zero range process, 20