UNIVERSIDADE DE SÃO PAULO
Instituto de Ciências Matemáticas e de Computação

Random particle systems and Schur processes

Victor Guerreiro Vendruscollo

Dissertação de Mestrado do Programa de Pós-Graduação em Matemática (PPG-Mat)

ICMC USP SÃO CARLOS

Victor Guerreiro Vendruscollo

Random particle systems and Schur processes

Dissertation submitted to the Instituto de Ciências Matemáticas e de Computação – ICMC-USP – in accordance with the requirements of the Mathematics Graduate Program, for the degree of Master in Science. *EXAMINATION BOARD PRESENTATION COPY*

Concentration Area: Mathematics

Advisor: Prof. Dr. Guilherme Lima Ferreira da Silva

USP – São Carlos
January 2023

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

Vendruscollo, Victor Guerreiro
V453s Sistemas de partículas aleatórias e processos de Schur / Victor Guerreiro Vendruscollo; orientador Guilherme Lima Ferreira da Silva. -- São Carlos, 2023.
147 p.

Dissertação (Mestrado - Programa de Pós-Graduação em Matemática) -- Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, 2023.

1. Processos de Schur. 2. Probabilidade Integrável. 3. Funções Simétricas. 4. TASEP. 5. Análise Assintótica. I. Lima Ferreira da Silva, Guilherme, orient. II. Título.

Bibliotecários responsáveis pela estrutura de catalogação da publicação de acordo com a AACR2:
Gláucia Maria Saia Cristianini - CRB - 8/4938
Juliana de Souza Moraes - CRB - 8/6176

Victor Guerreiro Vendruscollo

Sistemas de partículas aleatórias e processos de Schur

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Mestre em Ciências – Matemática. *EXEMPLAR DE DEFESA*

Área de Concentração: Matemática

Orientador: Prof. Dr. Guilherme Lima Ferreira da Silva

USP – São Carlos
Janeiro de 2023

Dedico este trabalho aos meus pais Ademar e Tirsá.

ACKNOWLEDGEMENTS

Para começar, agradeço aos meus familiares por todo o suporte dado até então. Agradeço aos meus pais, Ademar e Tirsá, por me apoiarem incondicionalmente durante toda a graduação e principalmente na minha mudança para São Carlos, mesmo sendo extremamente difícil ver o filho menor indo morar a mais de 1000km de distância de casa. Obrigado pai e mãe por me alimentarem durante o primeiro ano de mestrado em meio à pandemia e por nunca terem deixado me faltar nada durante minha vida. Agradeço ao meu irmão Anderson, por ir me buscar de madrugada no aeroporto de Porto Alegre para eu aparecer de surpresa em casa no dia dos pais. Agradeço à minha tia Marinês, por ser praticamente uma coorientadora durante meus 6 anos na Matemática. Agradeço ao meu pequeno sobrinho Benjamin por me dar alegria nos difíceis dias de estudo em casa. Também agradeço ao meu cachorro Bob, pelo amor incondicional que só um cãozinho pode dar.

Deixo meus agradecimentos para meu orientador Guilherme, por ter me dado a oportunidade de vir estudar em São Carlos com bolsa FAPESP. Isso facilita muito a vida de qualquer estudante, visto que atualmente vivemos em um momento delicado perante ao financiamento à ciência brasileira. Ele me deu essa chance mesmo depois de eu ter demorado 18 dias para responder seu *email*, o qual caiu em minha caixa de *spam* (uma das conclusões dessa dissertação é: não usar o *hotmail* nunca mais!). Obrigado por compreender a minha difícil decisão de não continuar para o doutorado em Matemática Pura, inclusive me auxiliando na transição para outra área (algo que nem é sua obrigação). Além de me ensinar Matemática, ele me ensinou muito sobre o profissionalismo no meio acadêmico. Agradeço por ter me colocado em um grupo de futebol logo na minha primeira semana em São Carlos e por aguentar eu reclamando do Grêmio na Série B. Enfim, posso dizer que nesses 2 anos o Guilherme foi praticamente um irmão mais velho.

Agradeço ao pessoal do meu grupo de pesquisa pelas divertidas partidas de *Gartic* nas reuniões a distância e pelos incríveis *happy hours* no S** G*** (aqui é sem anúncio grátis). Obrigado a todos os meninos do futebol da Matemática (até os da Álgebra) pelo futebol de altíssimo nível jogado nas quadras de São Carlos. Agradeço também a todos do Laboratório de Análise, que sempre contribuíram para um ambiente legal de estudos no ICMC.

Agradeço em especial a todos meus amigos de minha cidade natal (Sobradinho-RS). São eles: Felipe, Shosha, Gi, Lou, Media, Luci, Cauê, Arthur, Dudi, Oberdan,

Roger, Jaque, Nadine, Japa, Gabi, Vivi, Zanca. Todos vocês tornaram o período de aulas a distância muito mais alegre e suportável. De fato, o único lado bom da pandemia foi poder ver vocês com mais frequência. Também não posso esquecer de agradecer ao grupo de futebol de Sobradinho, mais um escape dos dias duros de estudo. Agora gostaria de fazer alguns agradecimentos especiais:

- Felipe, obrigado por se juntar a mim na destruição total de um bar porto alegrense, na invasão de uma formatura aleatória e por me presentear com uma rosa, tudo isso no último dia antes de me mudar para São Carlos. Brincadeiras à parte, obrigado por ser um dos meus melhores amigos desde o ensino médio, gostaria de ser seu amigo há muito mais tempo. Obrigado por ter sido o melhor colega de apartamento que tive até hoje, infelizmente fomos separados por diferentes ambições pessoais e profissionais. Obrigado por me apresentar ao veganismo, o qual é uma das formas mais simples de fazer a diferença em busca de um mundo melhor, essa foi talvez uma das melhores coisas que aconteceram na minha vida.

- Gi, obrigado por abrir tua casa para dezenas de *rolés* durante o ano de mestrado a distância, esses dias contribuíram muito para minha sanidade mental. Muito obrigado por impedir que eu perdesse o ônibus para São Paulo, depois de eu esquecer de ligar o despertador após um *porre* na noite porto alegrense, já citado no item anterior (Guilherme, por favor, ignore esta parte). Obrigado por todos os dias de futebol durante 2021 e também pelas corridinhas na *Fejão*, só não agradeço por me obrigar a ir no *LAB*.

- Shosha, obrigado por ser o meu filósofo pós-moderno preferido e por ser o meu principal recomendador de filmes e músicas legais. Muito obrigado por estar sempre com a casa aberta com cervejinhas e *indie rock* durante 2021, me fazendo sair de casa em dias em que eu já estava preparado para dormir. Obrigado por todas as madrugadas em que ficávamos conversando sobre a vida mesmo depois que todos os outros iam embora do *rolé*. Se a gente tivesse se conhecido quando criança, a minha vida teria sido ainda melhor.

- Lou, obrigado por me acolher em sua casa em Mogi Mirim e fazer eu me sentir menos estrangeiro aqui em São Paulo. Essa foi a lenda de dois sobradinhenses perdidos em São Paulo, só quem viveu sabe. Muito obrigado por conversar comigo em dias difíceis, sobre as dificuldades de ser adulto e de estar longe das pessoas que a gente gosta.

- Paulo, obrigado por me avisar sobre a existência do programa de mestrado do ICMC, mesmo sabendo que eu seria um forte concorrente a uma bolsa. Obrigado pela parceria desde o início da graduação em Santa Maria até o fim do mestrado aqui em São Carlos.

- Alan, muito obrigado por ter me recebido em sua casa no meu primeiro dia em São Carlos, após eu ter chegado totalmente assustado, cansado das 27 horas de viagem e sem lugar para dormir.

- Victor Júlio, obrigado por ser uma das pessoas que melhor me acolheram em São Carlos, com todas as conversas aleatórias no *lab* e no RU. Obrigado por aturar minha militância vegana e minha chatice sulista inerente (os dois ao mesmo tempo).

Este projeto foi financiado pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), processo n° 2020/16783-8. Dessa forma, é claro que também agradeço à FAPESP, por proporcionar uma estrutura financeira estável e confortável para estudantes como eu continuarem desenvolvendo a ciência no país. Infelizmente esse privilégio é para poucos, por isso sou muito grato.

Agradeço a mim mesmo por ser forte o suficiente para completar os dois anos de mestrado: um deles durante a pandemia e outro longe da família e muitos amigos importantes. Deixar o conforto da minha casa na minha pequena cidade de 15 mil habitantes lá do Sul para me mudar para o estado de São Paulo foi absolutamente a coisa mais difícil que fiz na vida até então, eu amadureci muito com isso. Fico feliz por ter tido a coragem e honestidade comigo mesmo ao tomar a difícil decisão de não continuar para o doutorado em Matemática Pura. Após muitas caminhadas e reflexões sobre o que quero para a minha vida, cheguei a esta conclusão e ainda espero encontrar o que eu realmente gosto de fazer.

Por fim, agradeço a todos que participaram de alguma forma na minha vida durante esses anos. Sei que falta agradecer muitas pessoas aqui, pois vários momentos simples como um elogio ou uma conversa no almoço podem fazem muita diferença no nosso dia-a-dia. Seriam necessárias dezenas de páginas para descrever essa gratidão. Se você que está lendo participou da minha trajetória e não escrevi sobre você aqui, saiba que eu lembro desses momentos.

"From beasts we scorn as soulless,
in forest, field and den,
the cry goes up to witness
the soullessness of men."
(M. Frida Hartley)

RESUMO

VENDRUSCOLLO, V. G. **Sistemas de partículas aleatórias e processos de Schur**. 2023. 147 p. Dissertação (Mestrado em Ciências – Matemática) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2023.

Pode soar absurdo acreditar que um modelo de tráfego, autovalores de matrizes aleatórias, e o comprimento da maior subsequência crescente de permutações aleatórias possuem uma conexão. Matemáticos dizem o contrário. Resultados surpreendentes mostram que flutuações desses e outros modelos têm o mesmo comportamento assintótico. O desenvolvimento da probabilidade integrável, um ramo da teoria de representações aplicada à probabilidade, é crucial para tais descobertas. O TASEP (totally asymmetric simple exclusion process) é um sistema de partículas interagentes que simula o tráfego. Este é encontrado em uma parametrização particular de um objeto mais geral, o processo de Schur, o qual é o principal objeto deste estudo.

Palavras-chave: Processos de Schur, Probabilidade Integrável, Funções Simétricas, TASEP, Análise Assintótica.

ABSTRACT

VENDRUSCOLLO, V. G. **Random particle systems and Schur processes**. 2023. 147 p. Dissertação (Mestrado em Ciências – Matemática) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2023.

It may sound absurd to believe that a traffic jam model, eigenvalues of random matrices, and the length of the longest increasing subsequence of random permutations have a connection. Mathematicians say otherwise. Surprising results show that some fluctuations of these and many other models have the same asymptotic behavior. The development of integrable probability, a branch of representation theory applied to probability, is crucial for such findings. The totally asymmetric simple exclusion process (TASEP) is an interacting particle system that simulates traffic. We find the TASEP in a particular parametrization of a more general object, the Schur process, which is the main object of this study.

Keywords: Schur Processes, Integrable Probability, Symmetric Functions, TASEP, Asymptotic Analysis.

CONTENTS

1 INTRODUCTION . . . 19
1.1. TASEP, Integrable Models and the KPZ Universality Class . . . 20
1.1.1. Block-Push Process . . . *22*
1.2. About the content . . . 23

2 SYMMETRIC FUNCTIONS . . . 27
2.1. Summary of the Chapter . . . 27
2.2. Partitions . . . 29
2.2.1. Skew Young Diagrams . . . *32*
2.2.2. Young tableaux and the Robinson-Schensted correspondence . . . *33*
2.3. Symmetric Polynomials and Symmetric Functions . . . 35
2.3.1. The Ring of Symmetric Polynomials . . . *35*
2.3.2. The Ring of Formal Power Series . . . *37*
2.3.3. The Ring of Symmetric Functions . . . *39*
2.3.4. Other bases for the ring of symmetric functions . . . *41*
2.3.5. Skew-symmetric Polynomials . . . *47*
2.4. Schur Polynomials and Schur Functions . . . 49
2.4.1. Schur Polynomials . . . *49*
2.4.2. Schur Functions . . . *52*
2.4.3. A Scalar Product on the Ring of Symmetric Functions . . . *54*
2.4.4. Skew Schur Functions . . . *55*
2.5. Specializations . . . 58

3 SCHUR PROCESSES . . . 63
3.1. Summary of the Chapter . . . 63
3.2. Probabilities on Partitions . . . 65
3.2.1. Schur Measure . . . *65*
3.2.2. Markov Chains on Partitions . . . *67*
3.3. Probabilities on Sequences of Partitions . . . 70
3.3.1. The Schur Process . . . *70*
3.3.2. Markov Chains on Sequences of Partitions . . . *72*
3.4. Application: Block-Push Process and TASEP . . . 77
3.4.1. Block-Push Markov Chain . . . *77*

3.4.2. The discrete-time TASEP *83*

4 DETERMINANTAL POINT PROCESSES 85
4.1. Summary of the Chapter 85
4.2. Random Point Processes 87
4.2.1. Correlation Functions *88*
4.3. Determinantal Point Processes 94
4.3.1. The Airy Point Process *97*
4.3.2. Biorthogonal Ensembles *101*
4.4. The Schur Measure as a Determinantal Point Process 106
4.4.1. Shifted Schur Measure *106*
4.4.2. The Schur measure is a DPP *110*
4.4.3. TASEP and DPPs *111*

5 ASYMPTOTICS 115
5.1. Summary of the Chapter 115
5.2. Asymptotic Methods 116
5.2.1. Laplace Method *117*
5.2.2. Steepest Descent Method *118*
5.3. Large-time asymptotics in TASEP 122
5.3.1. Rescaled TASEP and the Airy Point Process *122*
5.3.2. Convergence of Fredholm Determinants *133*
5.3.3. Trace-Class Operators and Fredholm Determinants *134*
5.3.4. Next Step: The Convergence to the Tracy-Widom Distribution . . *137*

BIBLIOGRAPHY 143

CHAPTER

1

INTRODUCTION

The Central Limit Theorem states that the Gaussian distribution is the limiting distribution of the normalized sum of independent and identically distributed (i.i.d.) random variables. The value of this result relies on its generality: no matter what the distribution of the variables is, they only need to be i.i.d. (under mild assumptions). This is called the Gaussian universality.

The first glimpse of this result appeared in the XVIIIth century when De Moivre and Laplace considered the problem for i.i.d. Bernoulli random variables. Their proof has an algebraic part, where they computed explicit formulas for probabilities, and an analytic part, where they calculated the asymptotic distribution using Stirling's approximation for $n!$. Although there was a strong belief in the truth of the Gaussian universality, rigorous proofs appeared only at the beginning of the XXth century, with the works of Chebyshev, Markov and Lyapunov (FISCHER, 2011).

Nowadays, we still study similar problems. Take a sequence of random variables (not i.i.d.) and define some quantity related to them (usually non-linear). Then we try to conjecture a kind of central limit theorem (universality). In the field of Mathematical Physics, for example, these problems may appear in the study of interacting particle systems, since the interest of Statistical Mechanics is in finding the equilibrium state when the number of particles becomes large. Moreover, we are particularly interested in probabilistic systems that can be analyzed by algebraic methods. Such models are called *integrable models* (BORODIN; GORIN, 2012; BAIK, 2018; FERRARI, 2019).

To illustrate, Random Matrix Theory is part of this area and good progress has been made in the last decades. The references (ERDőS; SCHLEIN; YAU, 2011; BLEHER; ITS, 1999; DEIFT; GIOEV, 2007; PASTUR; SHCHERBINA, 1997; SOSHNIKOV, 1999; TAO; VU, 2010; TAO; VU, 2011; TRACY; WIDOM, 1994) are just some examples of several results found. In short, Dyson's conjecture says that eigenvalues statistics converge

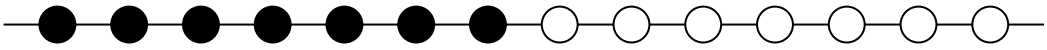
(when the matrix dimension becomes large) to a distribution that does not depend on the initial matrix distribution, but only on the symmetry class of the matrix ensemble (ERDőS; YAU, 2012). A limit distribution that appears in this theory is the well-known Tracy-Widom distribution (TRACY; WIDOM, 1994). The Semicircle Law and the Sine Process also appear in such problems.

The KPZ universality is related to statistics of random growth models (CORWIN, 2016), like the totally asymmetric simple exclusion process (TASEP), an integrable model which we present next. It is possible to find exact formulas for some statistics of the TASEP and then use them in the asymptotic analysis (BORODIN; GORIN, 2012). The limiting distribution for the fluctuations of the so-called *height function* of this model is the Tracy-Widom distribution (JOHANSSON, 2000). That is right, even though the TASEP does not seem related to random matrices, it is the same distribution found in random matrix theory! Many other models present this behavior such as the Corner Growth Model (JOHANSSON, 2000), the Last Passage Percolation (BORODIN; OKOUNKOV; OLSHANSKI, 2000; JOHANSSON, 2001a), the largest increasing subsequence in random permutations (BAIK; DEIFT; JOHANSSON, 1999) and the Poly-Nuclear Growth model (PRäHOFER; SPOHN, 2002). Therefore, mathematicians have enough evidence to suspect about universality here.

1.1. TASEP, Integrable Models and the KPZ Universality Class

The totally asymmetric simple exclusion process (TASEP) was first introduced by Spitzer (1970) and it is a simple but very rich interacting particle system. This model can be interpreted as a traffic model, for example. Suppose we have N particles positioned in $\mathbb{Z}$ denoted by $y_1(t) > \cdots > y_N(t)$, where t denotes the time. At time $t = 0$, they are positioned on the first N non-positive integers, that is, $y_1(0) = 0, y_2(0) = -1, \ldots, y_N(0) = 1 - N$ (see Figure 1).

Figure 1 – Initial condition in TASEP. The black circles are particles and the white ones are empty spaces.

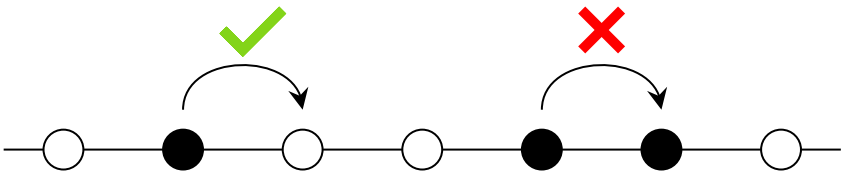


Source: Elaborated by the author.

The TASEP has the following dynamics: the particles want to jump from left to right, but they only do it when they are not blocked by another one (Figure 2). If they are free to move, we flip a coin to decide if the particle jumps (the coin does not need to be fair). For now, we may assume discrete time and we update the particles one by one,

going from the first one to the last one. When the last particle is updated, we say that one unit of time is finished. Check Patrik Ferrari's website[1] to see simulations of this model.

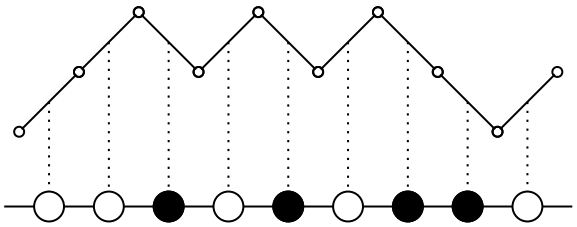
Figure 2 – Particles can only jump if the adjacent space on the right is empty.



Source: Elaborated by the author.

The main object studied in the literature about the TASEP is the *height function* $h(x,t)$. This function associates the current configuration of the model, at each time t, to a real function, in such a way that the difference $h(x+\frac{1}{2},t)-h(x+\frac{1}{2},0)$ counts the exact number of particles that have passed the position x until the time t. The graphic of $h(x,t)$ at time t is obtained connecting segments of slope -1 and 1, depending on the existence or not of a particle at position x, respectively (see Figure 3).

Figure 3 – The height function



Source: Elaborated by the author.

It was proven that fluctuations of this height function converge to a specific distribution in Random Matrix Theory (JOHANSSON, 2000). More precisely, Johansson obtained the convergence

$$\lim_{t\to\infty} \mathbb{P}\left(\frac{h(t,tx)-c_1(x)t}{c_2(x)t^{1/3}} \geq -s\right) = \mathrm{F}_2(s), \tag{1.1}$$

where c_1 and c_2 are explicit functions and F_2 is the Tracy-Widom distribution, which is the limiting distribution for fluctuations of the largest eigenvalue in the Gaussian Unitary Ensemble of random matrices (TRACY; WIDOM, 1994).

The route used to obtain this result is in direct analogy to the one used to prove particular cases for the Central Limit Theorem. First, there is an algebraic part, computing explicit formulas for the distribution of the height function. Next, we use these formulas to obtain the asymptotic distribution for the fluctuations of the height function.

[1] <https://wt.iam.uni-bonn.de/ferrari/research/jsanimationtasep>

The algebraic techniques cited above usually come from the Representation Theory. The subarea of representation theory devoted to finding algebraic methods to solve probabilistic problems is called Integrable Probability (BORODIN; GORIN, 2012). The theory of Symmetric Functions and the Robinson-Schensted-Knuth (RSK) correspondence are some examples of that. On the asymptotic analysis for integrable models, we often rely on Laplace method, Steepest Descent Method, Fredholm determinants and Riemann-Hilbert problems (CORWIN, 2016).

Surprisingly, many two-dimensional models (one spatial and one temporal) not related to the TASEP present the same asymptotic behavior (ADLER; MOERBEKE, 2005; BAIK; DEIFT; JOHANSSON, 1999; BORODIN, 2003; BORODIN *et al.*, 2007; BUFETOV; GORIN, 2018; GORIN; SHKOLNIKOV, 2015; OKOUNKOV; RESHETIKHIN, 2007). Such systems belong to the *Kardar-Parisi-Zhang (KPZ) universality class*. It is conjectured that all models in this class converge under the KPZ 1:2:3 scaling to a universal fluctuation field $\mathfrak{h}(x,t)$. This field does not depend on the model but only on the initial condition class. The invariant limit process is called *KPZ fixed point*, and it was recently constructed by Matetski, Quastel and Remenik (2021). In (CORWIN, 2016), we see several models, as random growth processes and stochastic PDEs, supporting this conjecture and revealing properties of the KPZ fixed point. However, at the moment we are arguably in a "De Moivre-Laplace stage" for this class of stochastic systems, say Borodin and Gorin (2012), in analogy to the Gaussian universality.

We hope the relevance of integrable models is clear. In this Master's Thesis, we only study the TASEP, but the reader should have in mind that this simple model is part of a wide class of models connected with each other by an object not very well understood yet: the KPZ fixed point.

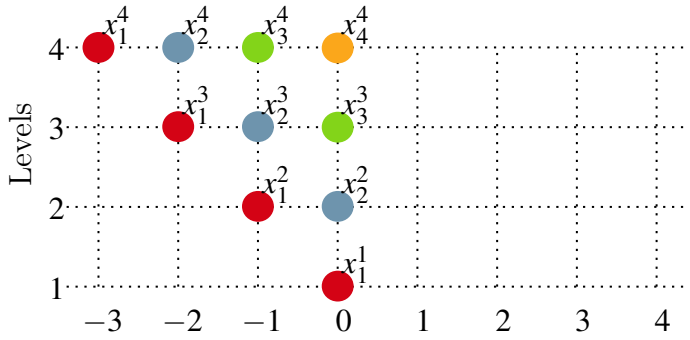
1.1.1. Block-Push Process

The Block-Push process is a model similar to the TASEP, but more general, in the sense that we can generate the TASEP from this model. To define it, let $\frac{N(N+1)}{2}$ be distinct particles positioned on N levels, labeled by $j = 1,\dots,N$, such that each level j has exactly j particles. The position of the i-th particle of the j-th level at time t is denoted by $x_i^j(t)$. For the initial conditions, we require that all particles of each level j lie on the first j non-positive integers (Figure 4), that is,

$$x_i^j(0) = i - j, \qquad 1 \le i \le j \le N.$$

As in TASEP, particles want to jump from left to right, staying in the same level. The movement is again decided by the flip of a coin performed particle by particle, now accordingly to the following rules:

Figure 4 – Block-Push: initial conditions (N=4). Particles with the same color have the same i-index.



Source: Elaborated by the author.

- (Block) A particle is "blocked" by the ones in the level below, that is, the ones with same i-index can never be at same position. To be more clear, see Figure 4, the particles with the same color can never be on the same vertical line;
- (Push) When a particle moves, the adjacent particle above also moves. Imagine that you have a pile of books, if you slowly push the lower one, the ones above also move;
- We update each particle individually, starting from the bottom to the top and from the right to the left, applying the rules above. A particle cannot move twice in the same update.

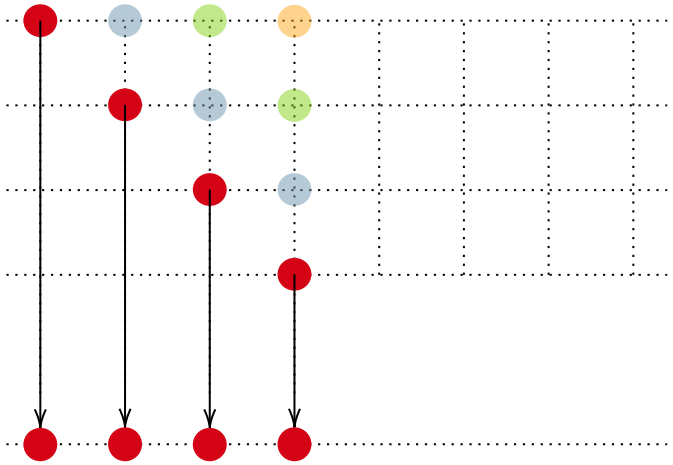
See Patrik Ferrari's website[2] for computer simulations. One could ask "why are we defining such complicated model if we just want to study the TASEP?". For now, we justify this by noticing that the TASEP is a restriction of the Block-Push process to the particles $y_i(t) = x_1^i(t)$, as if we were projecting the last particle of each level on a line (see Figure 5). Therefore, if we understand the Block-Push process, we may obtain results about the TASEP. In this work, we will see that the Block-Push process is just a parametrization of a large class of models.

1.2. About the content

Now we present the content of this Master's Thesis, which is a study of the existent literature. We do not present any new results. The goal is to produce an accessible and useful text for beginners in the area, since most of the literature is in form of journal articles and advanced lecture notes. As for prerequisites, the reader is expected to be familiar to the basics of *Measure Theory*, *Probability* and *Complex Analysis*, so that an undergraduate student in final years can follow the text. The end of the last chapter is the only part that requires some background in *Functional Analysis*. To facilitate access

[2] <https://wt.iam.uni-bonn.de/ferrari/research/jsanimationakpz>

Figure 5 – Looking only at the red particles, we obtain the TASEP.



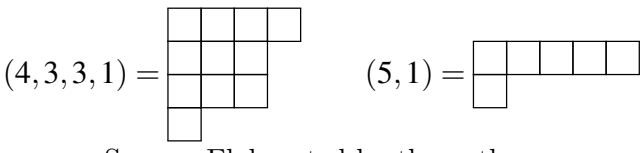
Source: Elaborated by the author.

to the individual topics, the chapters are rendered as self-contained as possible. Let us present a summary of the content of each chapter.

In Chapter 2, we present the necessary theory of symmetric functions. They can be understood as "polynomials" in infinitely many variables that are invariant under permutations of its variables. A fundamental class of symmetric functions, called *Schur functions*, is used to compute probabilities over a particular type of objects: the *partitions*.

The partitions are finite decreasing sequences of non-negative integers. They can be visualized as diagrams of boxes, called Young diagrams (Figure 6). We will show later that the TASEP can be described by an evolution of these diagrams. The main results of this chapter are the *Cauchy* identities, because they allow the computation of probabilities on the set of partitions.

Figure 6 – Young diagrams. The number of boxes in each row corresponds to each number of the sequence.



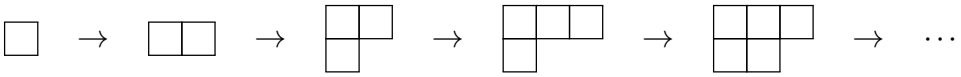
Source: Elaborated by the author.

After this, we proceed to the main topic of the text, the *Schur processes* (Chapter 3). In this chapter, some of the magic happens: all the nice identities obtained in Chapter 2 are explored together to define the Schur processes and to calculate probabilities on partitions in a systematic way.

We start defining the *Schur measure*, which is a probability measure on the set of partitions. In other words, the Schur measure is an "infinite die" with partitions on its faces. Then we define a *Markov chain* on partitions, which is a random evolution of

Young diagrams, as in Figure 7. We will show that this Markov chain has the important property of preserving Schur measures.

Figure 7 – A pictorial example of a Markov chain on partitions, which produces a "growth" of a diagram.



Source: Elaborated by the author.

The Schur processes are probability measures on vectors of partitions, generalizing the idea above. Again we define Markov chains for these objects, such that they preserve Schur processes. To finish the chapter, we show how one can study the TASEP as a particular parametrization of a Schur process (which is the Block-Push process discussed above). This is not the only way to study the TASEP, we can study it by solving Kolmogorov equations via Bethe ansatz method (BETHE; METALLE, 1931; GWA; SPOHN, 1992; SCHüTZ, 1997; TRACY; WIDOM, 2008; BAIK; LIU, 2018), however, this is beyond the scope of this work. Moreover, Schur processes are far more accessible for beginners since it requires just basic knowledge in Algebra.

At each time t, the particles in TASEP are positioned over a random subset of integers. So in Chapter 4, we study a general theory used to understand probabilities about random particles in a given space, called *determinantal point processes* (DPP). The important property of DPPs is that we can compress a lot of information about the particles in just a two variable function $K(x,y)$, called *correlation kernel*. We show how to apply this theory in TASEP and how the probabilities computed in Chapter 3 are expressed in this language.

In Chapter 5, the last one, we discuss the asymptotics for the TASEP, that is, its behavior in large time scale. We show how a rescaled version of the model is related to the *Airy point process*, which gives rise to the celebrated Tracy-Widom distribution (TRACY; WIDOM, 1994). We tried to reproduce this well-established result in the literature, using a more accessible approach.

As you will see, the content of this Master's thesis is quite extensive. So before going through the full text, we strongly recommend to the reader to follow the first section of each chapter, where we present a short summary for the respective chapter. In this way, it will be easier to understand the goals of this work. More precisely, start reading in the following order:

Section 2.1 $\rightarrow$ Section 3.1 $\rightarrow$ Section 4.1 $\rightarrow$ Section 5.1.

After that, you can read the full chapters to explore the details.

CHAPTER

2

SYMMETRIC FUNCTIONS

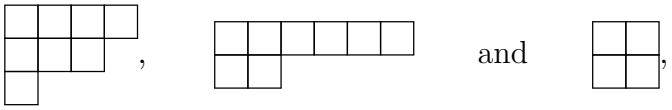
In this chapter, we define the Schur functions and obtain the Cauchy identities. They are the basic objects necessary for the later definition of Schur processes, in Chapter 3. The text is accessible to undergraduate students. The experienced reader is invited to see (SAGAN, 2001, Ch. 4) or (BAIK; DEIFT; SUIDAN, 2016, Ch. 10) for a more direct (albeit not detailed) approach, or simply read the summary provided here. For more details and results, see (MACDONALD, 1998). All the text of this chapter is based on the three references above.

2.1. Summary of the Chapter

The **symmetric functions** are symmetric polynomial expressions in infinitely many variables, as

$$x_1 + x_2 + x_3 + \cdots \qquad \text{or} \qquad x_1x_2 + x_1x_3 + x_2x_3 + \cdots .$$

To understand the space of symmetric functions, denoted by Λ, we need the so-called **partitions**, which are finite sequences of non-increasing non-negative integers. Each partition is represented by a **Young diagram**. For example, the partitions $(4,3,1)$, $(6,2)$ and $(2,2)$ are represented by the diagrams of boxes



respectively. The diagrams are self-explanatory: the number of boxes on the i-th row represents the i-th term in the partition. Let $\mathbb{Y}$ denote the set of all partitions.

A **semistandard Young tableau** (SSYT) of shape $\lambda \in \mathbb{Y}$ is a filling of boxes of the diagram of λ, with natural numbers, such that they weakly increase along the rows

and strictly increase down the columns. For example, let $\lambda = (4,3,1)$, then

$$\begin{array}{|c|c|c|c|}\hline 1&1&3&4\\\hline 3&3&7\\\cline{1-3} 4\\\cline{1-1}\end{array} \quad \text{is a SSYT and} \quad \begin{array}{|c|c|c|c|}\hline 1&1&3&5\\\hline 1&2&6\\\cline{1-3} 4\\\cline{1-1}\end{array} \quad \text{is not.} \tag{2.1}$$

Let T be a SSYT and let n_i be the number of times that $i \in \mathbb{N}$ appears on T. We can associate T to the monomial

$$\mathrm{x}^T = x_1^{n_1} x_2^{n_2} \cdots x_k^{n_k}.$$

For example, in the SSYT in 2.1, we have $x^T = x_1^2 x_3^3 x_4^2 x_7$. Given $\lambda \in \mathbb{Y}$, the associated **Schur function** is defined by the formal series

$$\mathrm{s}_\lambda(x_1, x_2, \ldots) = \sum_T \mathrm{x}^T,$$

where the sum is over all SSYTs with shape λ. For example, for $\lambda = (2,1)$, some of the possible SSYTs are

$$\begin{array}{|c|c|}\hline 1&1\\\hline 2\\\cline{1-1}\end{array}, \quad \begin{array}{|c|c|}\hline 1&2\\\hline 2\\\cline{1-1}\end{array}, \quad \begin{array}{|c|c|}\hline 1&1\\\hline 3\\\cline{1-1}\end{array}, \quad \begin{array}{|c|c|}\hline 1&3\\\hline 3\\\cline{1-1}\end{array}, \quad \begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}, \quad \begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array}, \quad \begin{array}{|c|c|}\hline 1&2\\\hline 4\\\cline{1-1}\end{array}, \quad \begin{array}{|c|c|}\hline 1&4\\\hline 2\\\cline{1-1}\end{array}, \ldots,$$

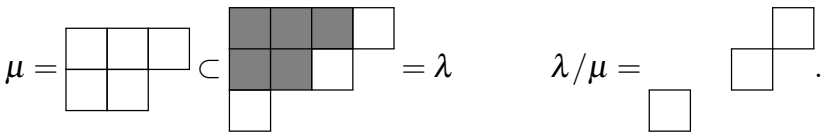
so we have

$$\mathrm{s}_\lambda(\mathrm{x}) = x_1^2 x_2 + x_1 x_2^2 + x_1^2 x_3 + x_1 x_3^2 + 2x_1 x_2 x_3 + 2x_1 x_2 x_4 + \cdots.$$

One can prove that the Schur functions are symmetric, i.e., invariant under permutations of variables.

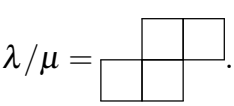
Let $\mu, \lambda \in \mathbb{Y}$, we write $\mu \subset \lambda$ when the diagram of μ is contained in the diagram of λ. If $\mu \subset \lambda$, the diagram obtained by "subtracting" μ from λ is called **skew Young diagram** and it is denoted by λ/μ. For instance, given $\lambda = (4,3,1)$ and $\mu = (3,2)$, we have $\mu \subset \lambda$, as Figure 8 shows.

Figure 8 – Skew Young diagram (μ is highlighted with darker boxes inside λ).



Source: Elaborated by the author.

Following the same process we used to define Schur functions, we define the **skew Schur functions**. For example, for partitions $\lambda = (3,2)$ and $\mu = (1)$, we have $\mu \subset \lambda$ and



We fill the boxes with weakly increasing numbers along the rows and strictly increasing numbers down the columns, in all possible ways. Some of the possibilities are

$$\begin{array}{|c|c|}\hline \multicolumn{1}{|c}{} & 1 & 1 \\ \hline 1 & 2 \\ \cline{1-2}\end{array}\,, \quad \begin{array}{|c|c|}\hline \multicolumn{1}{|c}{} & 1 & 2 \\ \hline 1 & 2 \\ \cline{1-2}\end{array}\,, \quad \begin{array}{|c|c|}\hline \multicolumn{1}{|c}{} & 1 & 2 \\ \hline 2 & 2 \\ \cline{1-2}\end{array}\,, \quad \begin{array}{|c|c|}\hline \multicolumn{1}{|c}{} & 2 & 2 \\ \hline 1 & 3 \\ \cline{1-2}\end{array}\,, \quad \dots .$$

Again we sum the corresponding monomials to define the skew Schur function

$$s_{\lambda/\mu}(x) = x_1^3x_2 + x_1^2x_2^2 + x_1x_2^3 + x_1x_2^2x_3 + \cdots .$$

These are the intuitive definitions for Schur functions and skew Schur functions. If they seem too confusing, do not worry: along the chapter we present them formally.

The main results of this chapter are the interesting Cauchy identities

$$\sum_{\lambda\in\mathbb{Y}} s_\lambda(x)s_\lambda(y) = \prod_{i,j=1}^{\infty} \frac{1}{1-x_iy_j} \tag{2.2}$$

and

$$\sum_{\nu\in\mathbb{Y}} s_{\nu/\lambda}(x)s_{\nu/\mu}(y) = \prod_{i,j=1}^{\infty} \frac{1}{1-x_ix_j} \sum_{\kappa\in\mathbb{Y}} s_{\lambda/\kappa}(y)s_{\mu/\kappa}(x). \tag{2.3}$$

They compress information about the Schur functions into a simple product $\prod_{i,j=1}^{\infty} \frac{1}{1-x_iy_j}$ depending only on the variables. Notice that the sums above run over all partitions, and as we will see, these formulas allow us to define probability measures on $\mathbb{Y}$.

2.2. Partitions

We start discussing partitions in more details.

Definition 2.1. A **partition** is a non-increasing sequence $\lambda = (\lambda_1, \lambda_2, \lambda_3, \dots)$ of non-negative integers with a finite number of non-zero terms. More precisely, $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \cdots \geq 0$ and there exists $n_0 \in \mathbb{N}$ such that $\lambda_n = 0$ for all $n \geq n_0$.

The compact notation $\lambda = (\lambda_1^{n_1}, \lambda_2^{n_2}, \dots)$ can be used if some terms appear more than once. Another common notation is $\lambda = (1^{m_1}, 2^{m_2}, 3^{m_3}, \dots)$, where for all $i \in \mathbb{N}$,

$$m_i = \mathrm{Card}\{j : \lambda_j = i\},$$

that is, m_i is the number of times that i repeats itself in λ (when we are working with more than one partition, we write $m_i = m_i(\lambda)$). For example, we have

$$(4,4,2,2,2,1,0,0,\dots) = (4,4,2,2,2,1) = (4^2,2^3,1) = (1,2^3,3^0,4^2).$$

Notice that the null terms are usually omitted.

Definition 2.2. The sum $|\lambda| = \sum_{i\in\mathbb{N}} \lambda_i$ is called the **size** of λ. If $\lambda = n$, we write $\lambda \vdash n$.

The sum $|\lambda|$ is always finite, since the number of non-zero terms in λ is finite. We denote by $\mathbb{Y}_n$ the set of all partitions of size n. If $|\lambda| = n$, we also say that λ is a **partition of** n, since we are literally breaking n in smaller parts. For example, $\mathbb{Y}_3 = \{(3), (2,1), (1^3)\}$, which corresponds to all possible ways of writing the number 3 as a sum of positive integers.

Example 2.1. These are all the fifteen elements of $\mathbb{Y}_7$: (7), $(6,1)$, $(5,2)$, $(5,1^2)$, $(4,3)$, $(4,2,1)$, $(4,1^2)$, $(3^2,1)$, $(3,2^2)$, $(3,2,1^2)$, $(3,1^4)$, $(2^3,1)$, $(2^2,1^3)$, $(2,1^5)$, (1^6).

Definition 2.3. The number of non-zero terms of λ is called the **length** of λ and it is denoted by $\ell(\lambda)$.

Example 2.2. For $\lambda = (5,4,1,0,0,\ldots)$, we have $|\lambda| = 10$ and $\ell(\lambda) = 3$.

The partition $\varnothing := (0,0,\ldots)$ is called the **empty partition**. The set of all partitions of any size is denoted by $\mathbb{Y}$. It is straightforward to note that $\mathbb{Y} = \dot{\bigcup}_{n=1}^{\infty} \mathbb{Y}_n$, where the symbol $\dot{\cup}$ means that the union is disjoint.

For easy visualization, we often represent partitions as a set of boxes, called **Young diagrams**, where each row, from the top to the bottom, represents a term of the partition (see Figure 9).

Figure 9 – Young diagrams

(a) $\lambda = (4,3,1)$ (b) $\lambda' = (3,2,2,1)$

If we look from the left to the right on Figure 9a, the number of boxes in each column of $\lambda = (4,3,1)$ defines a new partition $\lambda' = (3,2,2,1)$ (Figure 9b). This leads us to the following definition.

Definition 2.4. Let $\lambda \in \mathbb{Y}$. The partition λ' defined by $\lambda'_i := \mathrm{Card}\{j : \lambda_j \geq i\}$ is called the **transpose** of λ.

From definition above and from the geometric representation of partitions, we obtain $\lambda_1 = \ell(\lambda')$ and $\lambda'_1 = \ell(\lambda)$. Note that λ' is well-defined, since $\{j : \lambda_j \geq i\} \supset \{j : \lambda_j \geq i+1\}$, so $\lambda'_i \geq \lambda'_{i+1}$.

Proposition 2.1. An equivalent definition for the transpose λ' is $\lambda'_i = \max\{j : \lambda_j \geq i\}$.

Proof. Let $j_0 = \max\{j : \lambda_j \geq i\}$, then

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{j_0-1} \geq \lambda_{j_0} \geq i > \lambda_{j_0+1} \geq \lambda_{j_0+2} \geq \cdots,$$

so $\text{Card}\{j : \lambda_j \geq i\} = j_0$. □

Given two partitions λ and μ, we would like to be able to compare them. We have several ways to do this and each one has some importance.

Definition 2.5. (Natural ordering) Given $\lambda, \mu \vdash n$, we write $\lambda \trianglelefteq \mu$ if

$$\forall i \geq 1, \quad \lambda_1 + \lambda_2 + \cdots + \lambda_i \leq \mu_1 + \mu_2 + \cdots + \mu_i.$$

In addition, if λ and μ satisfy the definition above and $\lambda \neq \mu$, we simply write $\lambda \triangleleft \mu$.

Definition 2.6. (Lexicographic Ordering) Given $\lambda, \mu \vdash n$, we write $\lambda < \mu$ if, for some $i \in \mathbb{N}$, we have $\lambda_j = \mu_j$ for all $j < i$ and $\lambda_i < \mu_i$.

Notice that both natural and lexicographic ordering compare partitions with the same size.

Example 2.3. The natural order is not total as we can see that $(3,3), (4,1,1) \in \mathbb{Y}_6$ are not comparable, since $3 \leq 4$ and $3+3 > 4+1$. On the other hand, the lexicographic order is total. To illustrate, in $\mathbb{Y}_6$ we have

$$(6) > (5,1) > (4,2) > \left(4,1^2\right) > \left(3^2\right) > (3,2,1) > \left(3,1^3\right) > \left(2^3\right) > \left(2^2,1^2\right) > \left(2,1^4\right) > (1^6).$$

The natural ordering dominates the lexicographic one, in the sense of the following proposition.

Proposition 2.2. Let $\lambda, \mu \vdash n$. If $\lambda \triangleleft \mu$, then $\lambda < \mu$.

Proof. If $\lambda \triangleleft \mu$, then $\lambda \neq \mu$, so there exists $i = \min\{j : \lambda_j \neq \mu_j\}$. If we had $\lambda_i > \mu_i$, then

$$\lambda_1 + \cdots + \lambda_{i-1} + \lambda_i = \mu_1 + \cdots + \mu_{i-1} + \lambda_i > \mu_i + \cdots + \mu_i,$$

thus $\lambda \not\trianglelefteq \mu$, which is a contradiction. So we must have $\lambda_i < \mu_i$. □

The next proposition provides a characterization for partitions, with respect to the natural ordering.

Proposition 2.3. A sequence λ of n non-negative integers defines a partition if, and only if, $\lambda \trianglerighteq \sigma(\lambda)$ for all permutations $\sigma \in \mathscr{S}_n$.

Proof. ($\Rightarrow$) Since $\lambda_1 \geq \cdots \geq \lambda_n$, we have $\lambda_1 + \cdots + \lambda_i \geq \lambda_{\sigma(1)} + \cdots + \lambda_{\sigma(i)}$ for any i, so $\lambda \trianglerighteq \sigma(\lambda)$.

($\Leftarrow$) For all $i = 1, \ldots, n-1$, take the permutation σ_i that switches λ_i by λ_{i+1}. Since $\lambda \trianglerighteq \sigma_i(\lambda)$ for each i, we have

$$(\lambda_1, \ldots, \lambda_{i-1}, \lambda_i, \lambda_{i+1}, \ldots, \lambda_n) \trianglerighteq (\lambda_1, \ldots, \lambda_{i-1}, \lambda_{i+1}, \lambda_i, \ldots, \lambda_n).$$

Particularly, we obtain

$$\lambda_1 + \cdots + \lambda_{i-1} + \lambda_i \geq \lambda_1 + \cdots + \lambda_{i-1} + \lambda_{i+1},$$

so $\lambda_i \geq \lambda_{i+1}$. □

2.2.1. Skew Young Diagrams

We have more three ways to compare partitions. Each one is necessary for this work.

Definition 2.7. We write $\mu \subset \lambda$ if the diagram of λ contains the diagram of μ, that is, $\mu_i \leq \lambda_i$ for all i (or equivalently, $\mu'_i \leq \lambda'_i$). The diagram obtained by subtracting μ from λ is called **skew Young diagram** and it is denoted by λ/μ.

Definition 2.8. (Interlacing Property) We write $\mu \preccurlyeq \lambda$, and say that λ and μ **interlace**, when $\lambda_{i+1} \leq \mu_i \leq \lambda_i$ for all i. In this case, the skew Young diagram λ/μ is called a **horizontal strip**.

Proposition 2.4. We have $\mu \preccurlyeq \lambda$ if, and only if, $0 \leq \lambda'_i - \mu'_i \leq 1$ for every i.

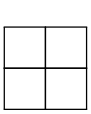
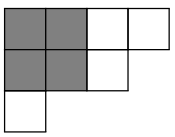
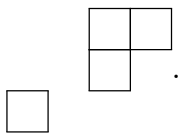
Proof. Assume $\mu \preccurlyeq \lambda$, then $0 \leq \lambda_i - \mu_i$, so $\mu \subset \lambda$ and then $0 \leq \lambda'_i - \mu'_i$. By Proposition 2.1, we have $\lambda'_i = j_0 = \max\{j : \lambda_j \geq i\}$, so by the interlacing property,

$$\mu_1 \geq \cdots \geq \mu_{j_0 - 1} \geq \lambda_{j_0} \geq i,$$

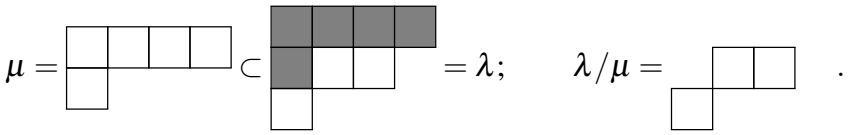
and this means that $\mu'_i \geq j_0 - 1 = \lambda'_i - 1$, proving the second inequality. For the reciprocal, we use the same idea. □

Definition 2.9. We write $\mu \preccurlyeq_{\mathrm{v}} \lambda$, when $0 \leq \lambda_i - \mu_i \leq 1$ for all i (in particular, $\mu \subset \lambda$). In this case, the skew Young diagram λ/μ is called a **vertical strip**.

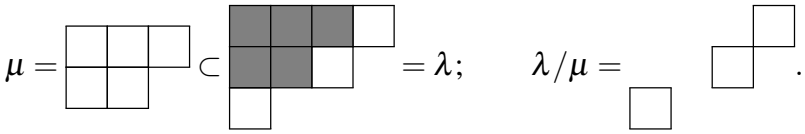
Example 2.4. Given $\lambda = (4,3,1)$ and $\mu = (2,2)$, we have $\mu \subset \lambda$ as the figure below shows (μ is highlighted with darker boxes). On the right, we show the skew Young diagram λ/μ, which is neither a horizontal strip nor a vertical strip, since $\lambda'_3 - \mu'_3 = 2$ and $\lambda_1 - \mu_1 = 2$, respectively.

$\mu =$  $\subset$  $= \lambda;$ $\qquad \lambda/\mu =$  .

Example 2.5. Given $\lambda = (4,3,1)$ and $\mu = (4,1)$, we see that λ/μ is an horizontal strip, but not a vertical strip.



Example 2.6. Given $\lambda = (4,3,1)$ and $\mu = (3,2)$, we see that λ/μ is both a horizontal and a vertical strip.



2.2.2. Young tableaux and the Robinson-Schensted correspondence

Let us go a little further on the study of Young diagrams. In this section, we show a simple way to compute probabilities on $\mathbb{Y}$, using a well-known result on representations of the symmetric group $\mathscr{S}_N$. We start defining a new object by filling the Young diagrams with natural numbers.

Definition 2.10. A **Young tableau** of shape $\lambda \in \mathbb{Y}_N$ is a filling of boxes on the diagram of λ with the naturals $1,2,\ldots,N$. If T is a Young tableau of shape λ, we denote $\mathrm{sh}(T) = \lambda$.

If the numbers strictly increase along the rows and down the columns of a Young tableau T, we say that T is a **standard Young tableau**.

Figure 10 – Young tableaux of shape $\lambda = (4,3,1)$

1	3	4	6
2	6	8	
7			

(a) Standard.

2	6	4	8
3	7	1	
5			

(b) Not standard.

We denote by $\mathbf{SYT}_N$ the set of all standard Young tableaux of size N, that is, $\mathrm{sh}(T) \vdash N$. The interesting result below was firstly proven by Robinson (1938) and then found in another form by Schensted (1961).

Theorem 2.1. (Robinson-Schensted correspondence) There is a bijection between the symmetric group $\mathscr{S}_N$ and the set

$$\{(T,U) \in \mathbf{SYT}_N \times \mathbf{SYT}_N : \mathrm{sh}(T) = \mathrm{sh}(U)\},$$

i.e., the set consisted of pairs of standard Young tableaux of size N with same shape.

The bijection established on Theorem 2.1 is obtained by an algorithm, but we do not enter in such details. We are only interested in the information that the set defined in the theorem has exactly $N!$ elements.

Example 2.7. The set SYT_3 is composed by the tableaux

$$T_1 = \begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline 3 \\ \hline \end{array}, \quad T_2 = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1} \end{array}, \quad T_3 = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1} \end{array} \quad \text{and} \quad T_4 = \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline \end{array},$$

so there are 3! pairs with the same shape. They are (T_1,T_1), (T_2,T_2), (T_3,T_3), (T_4,T_4), (T_2,T_3) and (T_3,T_2), agreeing to Theorem 2.1.

Given $\lambda \vdash N$, define

$$\dim(\lambda) := \mathrm{Card}\{T \in \mathrm{SYT}_N : \mathrm{sh}(T) = \lambda\},$$

in other words, $\dim(\lambda)$ is the number of possible standard Young tableaux with shape λ. So the number of pairs $(T,U) \in \mathrm{SYT}_N \times \mathrm{SYT}_N$ such that $\mathrm{sh}(T) = \mathrm{sh}(U) = \lambda$ is exactly $\dim(\lambda)^2$, hence Theorem 2.1 implies that

$$\sum_{\lambda \vdash N} \dim(\lambda)^2 = N!. \tag{2.4}$$

Definition 2.11. The **Plancherel measure** is the function defined by

$$\mathbb{P}_N(\lambda) = \frac{\dim(\lambda)^2}{N!}, \quad \lambda \in \mathbb{Y}_N.$$

By equation (2.4), $\mathbb{P}_N$ is a probability measure on $\mathbb{Y}_N$, that is, we have $\sum_{\lambda \in \mathbb{Y}_N} \mathbb{P}_N(\lambda) = 1$. In other words, $\mathbb{P}_N$ is a biased die with the partitions of $\mathbb{Y}_N$ on its faces.

Example 2.8. For $N = 3$, we have by Example 2.7,

$$\dim\left(\ydiagram{3}\right)^2 = \dim\left(\ydiagram{1,1,1}\right)^2 = 1 \quad \text{and} \quad \dim\left(\ydiagram{2,1}\right)^2 = 4,$$

so $\mathbb{P}_3$ is given by

$$\mathbb{P}_3\left(\ydiagram{3}\right) = \mathbb{P}_3\left(\ydiagram{1,1,1}\right) = \frac{1}{6} \quad \text{and} \quad \mathbb{P}_3\left(\ydiagram{2,1}\right) = \frac{4}{6}.$$

So we can say that $\mathbb{P}_3$ is an ordinary die with six faces, but the partitions (3) and $(1,1,1)$ appear on one face, respectively, and $(2,1)$ appears on four faces (Figure 11).

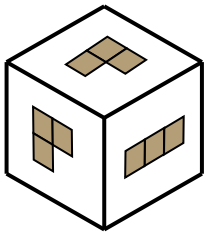
Using the Plancherel measures, we can define a measure on $\mathbb{Y}$ in the following way.

Definition 2.12. Let $t > 0$. The **Poissonized**[1] **Plancherel measure** parametrized by t is the function defined by

$$\mathbb{P}_t^{\mathrm{Poisson}}(\lambda) = \mathrm{e}^{-t} \frac{t^{|\lambda|}}{(|\lambda|!)^2} \dim(\lambda)^2, \quad \lambda \in \mathbb{Y}.$$

[1] The name is related to the expression $\mathrm{e}^{-t}t^n/n!$, which is the probability mass function for the Poisson distribution.

Figure 11 – Die for $\mathbb{P}_3$



Source: Elaborated by the author.

The function $\mathbb{P}_t$ is indeed a probability measure on $\mathbb{Y}$, since

$$\sum_{\lambda\in\mathbb{Y}} \mathbb{P}_t(\lambda) = \sum_{N=0}^{\infty} \sum_{\lambda\vdash N} \mathrm{e}^{-t} \frac{t^N}{(N!)^2} \dim(\lambda)^2 = \mathrm{e}^{-t} \underbrace{\sum_{N=0}^{\infty} \frac{t^N}{N!} \underbrace{\sum_{\lambda\vdash N} \mathbb{P}_N(\lambda)}_{=1}}_{=\mathrm{e}^t} = 1.$$

In Chapter 3, we show that the Poissonized Plancherel measure is a particular parametrization of a more general probability measure. But before this, we need to dive into the theory of symmetric functions.

2.3. Symmetric Polynomials and Symmetric Functions

2.3.1. The Ring of Symmetric Polynomials

Given a permutation $\sigma \in \mathscr{S}_N$ and a polynomial $p(x_1,\ldots,x_N)$, the action of σ in p is given by

$$\sigma(p) = \sigma \cdot p(x_1,x_2,\ldots,x_N) \coloneqq p\left(x_{\sigma(1)},x_{\sigma(2)},\ldots,x_{\sigma(N)}\right).$$

In other words, σ acts in p by switching the variables according to σ.

Example 2.9. For $p(x_1,x_2,x_3) = x_1^{1997} + x_2^7 x_3$ and $\sigma = (1\ 3\ 2)$, we have

$$\sigma \cdot p(x_1,x_2,x_3) = x_3^{1997} + x_1^7 x_2.$$

Definition 2.13. A polynomial p in N variables is called **symmetric** if, for any permutation $\sigma \in \mathscr{S}_N$, we have $\sigma \cdot p = p$.

The set of all symmetric polynomials in N variables is denoted by Λ_N.

Example 2.10. The polynomial of Example 2.9 is not symmetric, but

$$q(x_1,x_2,x_3) = x_1^{2022} + x_2^{2022} + x_3^{2022} + x_1x_2 + x_1x_3 + x_2x_3$$

is symmetric.

Let $\mathbb{C}[x_1,\ldots,x_N]$ be the vector space (and ring) of polynomials in N variables with complex coefficients. It is straightforward to verify that Λ_N is closed under the operations of sum and multiplication of polynomials, so the next proposition holds.

Proposition 2.5. The set Λ_N is a vector subspace and a subring of $\mathbb{C}[x_1,\ldots,x_N]$.

It is our interest to find a linear basis for Λ_N. For this purpose, we need to define the "simplest" symmetric polynomials. We are now adapting the ideas of (SAGAN, 2001, Ch. 4) for finite variables.

Definition 2.14. Let λ be any partition with length $\ell = \ell(\lambda) \leq N$, the polynomial

$$\mathrm{m}_{\lambda,N} = \mathrm{m}_\lambda(x_1,\ldots,x_N) \coloneqq \sum x_{i_1}^{\lambda_1}\cdots x_{i_\ell}^{\lambda_\ell},$$

summed over all ℓ-tuples $(x_{i_1},\ldots,x_{i_\ell})$ of distinct variables, is called **monomial symmetric polynomial** in N variables. If $\ell(\lambda) > N$, we set $\mathrm{m}_{\lambda,N} \coloneqq 0$.

Example 2.11. For $\lambda = (3,1)$, we have

$$\begin{aligned}
\mathrm{m}_{\lambda,2} &= x_1^3x_2 + x_1x_2^3;\\
\mathrm{m}_{\lambda,3} &= x_1^3x_2 + x_1^3x_3 + x_1x_2^3 + x_1x_3^3 + x_2^3x_3 + x_2x_3^3;\\
\mathrm{m}_{\lambda,4} &= x_1^3x_2 + x_1^3x_3 + x_1^3x_4 + x_1x_2^3 + x_1x_3^3 + x_1x_4^3 + x_2^3x_3 + x_2^3x_4 + x_2x_3^3 + x_2x_4^3 + x_3^3x_4 + x_3x_4^3.
\end{aligned}$$

From Example 2.11, we see that $\mathrm{m}_\lambda(x_1,\ldots,x_{N-1},0) = \mathrm{m}_\lambda(x_1,\ldots,x_{N-1})$ for every N and λ. Since $\mathrm{m}_{\lambda,N}$ is obtained by summing over all permutations of variables, it must be symmetric ($\mathrm{m}_{\lambda,N} \in \Lambda_N$). We can interpret $\mathrm{m}_{\lambda,N}$ as the "smallest" symmetric polynomial containing the monomial $x_1^{\lambda_1}\cdots x_\ell^{\lambda_\ell}$, that is, the polynomial with less monomials satisfying this property.

We say that a polynomial is **homogeneous of degree** n if all monomials in its expression have degree n.

Example 2.12. The polynomial

$$p(x_1,x_2,x_3) = x_1^{21} + x_1^{14}x_2^7 + x_1^7x_2^{14} + x_2^{21} + x_1^{14}x_3^7 + x_2^{14}x_3^7 + x_1^7x_3^{14} + x_2^7x_3^{14} + x_3^{21}$$

is homogeneous of degree 21.

Let

$$\Lambda_N^n \coloneqq \{p \in \Lambda_N : p \text{ is homogeneous of degree } n\} \cup \{0\},$$

be the space of symmetric homogeneous polynomials of degree n, which is a vector subspace of Λ_N. Notice that if $f \in \Lambda_N^n$ and $g \in \Lambda_N^m$, then $fg \in \Lambda_N^{n+m}$. We also have the direct sum

$$\Lambda_N = \bigoplus_{n=0}^{\infty} \Lambda_N^n, \tag{2.5}$$

so to prove that the $m_{\lambda,N}$'s provide a basis for Λ_N, it suffices to find a basis for each Λ_N^n.

Proposition 2.6. The set $\mathscr{B} = \{\mathrm{m}_{\lambda,N} : \lambda \vdash n, \ell(\lambda) \leq N\}$ is a linear basis for Λ_N^n.

Proof. If $\lambda \vdash n$, then $\mathrm{m}_{\lambda,N}$ is homogeneous of degree n, so $\mathbf{Span}(\mathscr{B}) \subset \Lambda_N^n$. On the other hand, let $p \in \Lambda_N^n$. It can be written as a sum of monomials of degree n, i.e.

$$p(x_1,\ldots,x_N) = \sum_i c_i x_{i_1} \cdots x_{i_n}.$$

Fixed some index i, the monomial $c_i x_{\sigma(i_1)} \cdots x_{\sigma(i_n)}$ is also in the expression of p, for any permutation $\sigma \in \mathscr{S}_N$. So there exists λ such that $c_i \mathrm{m}_{\lambda,N}$ appears in the expression of p. Since i is arbitrary, we obtain that p is a linear combination of $\mathrm{m}_{\lambda,N}$'s. Finally, we shall prove that $\mathscr{B}$ is linearly independent. Take any linear combination

$$a_1 \mathrm{m}_{\lambda_1} + \cdots + a_k \mathrm{m}_{\lambda_k} = 0,$$

assuming all λ_i's distinct. This is a sum of distinct monomials, with repeated coefficients and equal to zero, so we have $a_1 = \cdots = a_k = 0$. □

By Proposition 2.6 and (2.5), we obtain the corollary below.

Corollary 2.1. The set $\{\mathrm{m}_{\lambda,N} : \lambda \in \mathbb{Y}, \ell(\lambda) \leq N\}$ is a basis for Λ_N.

2.3.2. The Ring of Formal Power Series

In this section, we present the ring of formal power series. Basically, they are polynomials in infinitely many variables. This object is necessary to understand the ring of symmetric functions.

Consider for each $N \in \mathbb{N}$ the projection

$$\begin{aligned} \pi_N : \mathbb{C}[x_1,\ldots,x_N] &\to \mathbb{C}[x_1,\ldots,x_{N-1}] \\ p(x_1,\ldots,x_N) &\mapsto p(x_1,\ldots,x_{N-1},0). \end{aligned}$$

In other words, π_N acts in a polynomial by setting $x_N = 0$.

Example 2.13. For the polynomial

$$p(x,y,z) = xy + xz + yz \in \mathbb{C}[x,y,z],$$

we have

$$\pi_3 p(x,y,z) = p(x,y,0) = xy \in \mathbb{C}[x,y].$$

Definition 2.15. Let $\mathrm{x} = \{x_1, x_2, x_3, \ldots\}$ be a countable set of infinitely many variables. The set of **formal power series** is defined by

$$\mathbb{C}[\mathrm{x}] = \mathbb{C}[x_1, x_2, \ldots] \coloneqq \{(p_1, p_2, \ldots) : \forall N \in \mathbb{N}, p_N \in \mathbb{C}[x_1,\ldots,x_N], \pi_N p_N = p_{N-1}\}.$$

The elements of $\mathbb{C}[\mathrm{x}]$ are sequences of polynomials such that when we set $x_N = 0$ in p_N, we obtain p_{N-1}.

Example 2.14. For each $N \in \mathbb{N}$, let $p_N(x_1,\ldots,x_N) = x_1 + \cdots + x_N$, so

$$p_N(x_1,\ldots,x_{N-1},0) = x_1 + \cdots + x_{N-1} + 0 = p_{N-1}(x_1,\ldots,x_{N-1}),$$

therefore, $p = (p_N)_{N\in\mathbb{N}} \in \mathbb{C}[\mathrm{x}]$.

Note 2.1. We often use some abuse of notation to represent formal power series. To illustrate, we would simply write p, from Example 2.14, as $p = x_1 + x_2 + x_3 + \cdots$. That is why we use the word *formal*: there are no worries about the convergence of such series (yet), because we are not evaluating the series in any values.

Example 2.15. The reader should be cautious about the abuse of notation described in Note 2.1. For instance, the infinite product

$$p(x_1,x_2,\ldots) = x_1x_2x_3\cdots$$

is not a formal power series. Indeed, writing p accordingly to Definition 2.16, we would have $p = (x_1x_2\cdots x_N)_{N\in\mathbb{N}}$ and we see that $\pi_N(x_1\cdots x_N) = 0 \neq x_1\cdots x_{N-1}$.

We equip $\mathbb{C}[\mathrm{x}]$ with the usual operations of sum and multiplication of polynomials, applied elementwise in the sequences. The good definition of these operations follows directly by the analogous operations for polynomials. Furthermore, it follows that $\mathbb{C}[\mathrm{x}]$ is a vector space and a ring.

For any monomial $p = x_1^{\lambda_1}\cdots x_N^{\lambda_N}$, the sum $\sum \lambda_i$ is the degree of p. The degree of a polynomial $p \neq 0$, denoted by $\deg(p)$, is the maximum of the degrees of monomials in its expression. We now extend this definition to formal power series.

Definition 2.16. Let $p = (p_N)_{N\in\mathbb{N}} \in \mathbb{C}[\mathrm{x}]\backslash\{0\}$, the **degree** of p is given by

$$\deg(p) := \sup_{N\in\mathbb{N}}\{\deg(p_N)\}.$$

If $\{\deg(p_N) : N \in \mathbb{N}\}$ is not bounded, we define $\deg(p) := \infty$.

The notion of degree allows the following remark, which is important for future manipulation of formal power series.

Example 2.16. The multiplication in $\mathbb{C}[\mathrm{x}]$ must be handled carefully. We have already seen that the monomial of infinite degree $\prod_{i=1}^{\infty} x_i$ is not a formal power series (Example 2.15). On the other hand, $p = \prod_{i=1}^{\infty}(1+x_i)$ is a formal power series since we can do this multiplication taking only monomials with finite degree, in the following way:

$$\begin{aligned}\prod_{i=1}^{\infty}(1+x_i) &= (1+x_1)(1+x_2)(1+x_3)\ldots \\ &= 1 + \sum x_i + \sum x_{i_1}x_{i_2} + \sum x_{i_1}x_{i_2}x_{i_3} + \cdots,\end{aligned}$$

where the sums above run through all permutations of $1,2,3,\ldots$ variables, respectively. Also, we have $\deg(p) = \infty$.

Example 2.17. The formal power series $x_1^2 + x_2 + x_3^2 + x_4 + \cdots$ has degree 2.

As we expect, the usual properties for degree of polynomials are still valid here: we have $\deg(p+q) = \max\{\deg(p), \deg(q)\}$ and $\deg(pq) = \deg(p) + \deg(q)$.

2.3.3. The Ring of Symmetric Functions

In analogy to the construction of the ring of symmetric polynomials, we now repeat the process with infinitely many variables.

Definition 2.17. Given $\lambda \in \mathbb{Y}$, the **monomial symmetric function** corresponding to λ is given by $\mathrm{m}_\lambda := (\mathrm{m}_{\lambda,N})_{N\in\mathbb{N}}$.

Note that $\mathrm{m}_\lambda \in \mathbb{C}[\mathrm{x}]$, since $\pi_N \mathrm{m}_{\lambda,N} = \mathrm{m}_{\lambda,N-1}$ for each N. We use the m_λ's as generators for the space of symmetric functions.

Definition 2.18. The vector space $\Lambda := \mathrm{Span}\{\mathrm{m}_\lambda : \lambda \in \mathbb{Y}\}$ is called the **ring of symmetric functions**.

We have defined Λ as a vector space, but it is also a ring since it is closed under the usual product of $\mathbb{C}[\mathrm{x}]$. Remember that every m_λ is symmetric, so are all $f \in \Lambda$. Notice that each $f \in \Lambda$ is written as a finite linear combination of m_λ's.

Example 2.18. We have $f = 2x_1 + 2x_2 + 2x_3 + \cdots \in \Lambda$, since f is exactly $2\mathrm{m}_{(1)}$.

Example 2.19. The product $f = \prod_{i=1}^{\infty}(1+x_i) \in \mathbb{C}[\mathrm{x}]$ is not a symmetric function, even though it is invariant under permutations of its variables. If f could be written as a finite linear combination of m_λ's, then f would have bounded degree, which does not happen, therefore, $f \notin \Lambda$.

Let us continue to reproduce the results proved in Section 2.3.1. Again, it is useful to decompose Λ as a direct sum of finite dimension subspaces. So we define the space of **symmetric homogeneous functions** of degree n, given by

$$\Lambda^n := \mathrm{Span}\{\mathrm{m}_\lambda : \lambda \vdash n\}.$$

Since the number of partitions of n is finite, Λ^n has finite dimension. Moreover, every $f \in \Lambda^n$ has degree n, because every m_λ for $\lambda \vdash n$ has degree n, so we have the natural decomposition

$$\Lambda = \bigoplus_{n=0}^{\infty} \Lambda^n. \tag{2.6}$$

The m_λ's, for $\lambda \vdash n$, not only generate Λ^n, but they are indeed a basis.

Proposition 2.7. The set $\{m_\lambda : \lambda \vdash n\}$ is linearly independent, so it is a basis for Λ^n.

Proof. Write

$$a_1 \mathrm{m}_{\lambda_1} + \cdots + a_k \mathrm{m}_{\lambda_k} = 0$$

where $k = p(n)$ (the number of partitions of n). By definition of m_λ, we have for every $N \in \mathbb{N}$,

$$a_1 \mathrm{m}_{\lambda_1,N} + \cdots + a_k \mathrm{m}_{\lambda_k,N} = 0,$$

so if we take N sufficiently large, we have $a_1 = \cdots = a_k = 0$, since the $\mathrm{m}_{\lambda,N}$'s provide a basis for Λ^n_N (Proposition 2.6). □

Example 2.20. We have $\mathbb{Y}_3 = \{(3),(2,1),(1,1,1)\}$, so $\{\mathrm{m}_{(3)}, \mathrm{m}_{(2,1)}, \mathrm{m}_{(1,1,1)}\}$ is a basis for Λ^3, which has dimension 3.

The next proposition sets a natural characterization for Λ^n, in terms of homogeneous polynomials in Λ^n_N.

Proposition 2.8. For each $n \in \mathbb{N}$, we have

$$\Lambda^n = \{(f_N)_{N\in\mathbb{N}} \in \mathbb{C}[\mathrm{x}] : f_N \in \Lambda^n_N\}.$$

Proof. If $f = (f_N) \in \Lambda^n$, we have $f = \sum_{\lambda\vdash n} c_\lambda \mathrm{m}_\lambda$, so for each N, $f_N = \sum_{\lambda\vdash n} c_\lambda \mathrm{m}_{\lambda,N} \in \Lambda^n_N$. Conversely, let f be such that $f_N \in \Lambda^n_N = \mathrm{Span}\{\mathrm{m}_{\lambda,N} : \lambda \vdash n\}$ for each N (using Proposition 2.6), then $f_N = \sum_{\lambda\vdash n} c_\lambda \mathrm{m}_{\lambda,N}$, therefore,

$$f = (f_N) = \left(\sum_{\lambda\vdash n} c_\lambda \mathrm{m}_{\lambda,N}\right) = \sum_{\lambda\vdash n} c_\lambda \mathrm{m}_\lambda \in \Lambda^n.$$

□

A similar characterization is obtained for Λ.

Proposition 2.9. We have

$$\Lambda = \{f = (f_N)_{N\in\mathbb{N}} \in \mathbb{C}[\mathrm{x}] : f_N \in \Lambda_N, \deg(f) < \infty\}.$$

Proof. By (2.6), if $f \in \Lambda$, then there exists $M \in \mathbb{N}$ and homogeneous symmetric functions $g^{(k)} \in \Lambda^k$ for $k = 1,\ldots,M$, such that

$$f = g^{(1)} + \cdots + g^{(M)} \Rightarrow f_N = g^{(1)}_N + \cdots + g^{(M)}_N \in \bigoplus_{k=0}^{\infty} \Lambda^k_N = \Lambda_N.$$

Also, we have $\deg(f) = \max\{\deg(g^{(k)})\} \leq M < \infty$. Reciprocally, if $\deg(f) < \infty$, there exists $M = \sup\{\deg(f_N)\}$, so each f_N can be written as

$$f_N = g_N^{(1)} + \cdots + g_N^{(M)},$$

where $g_N^{(k)} \in \Lambda_N^k$ for each k. Defining $g^{(k)} = \left(g_1^{(k)}, g_2^{(k)}, \ldots\right)$, we have $g^{(k)} \in \Lambda^k$ and $f = g^{(1)} + \cdots + g^{(M)} \in \oplus\Lambda^k = \Lambda$. □

Proposition 2.9 says that every symmetric function is a formal power series composed by a sequence of symmetric polynomials such that the set of their degrees is bounded.

Example 2.21. The hypothesis $\deg(f) < \infty$ is necessary in the last proposition. To illustrate, take again the formal power series $f = (f_N)$ where $f_N = \prod_{i=1}^{N}(1 + x_i)$. It is clear that $f_N \in \Lambda_N$ for each N, but as we saw in Example 2.16, $\deg(f) = \infty$ and $f \notin \Lambda$.

The following property shows that Λ is closed under the usual multiplication of formal power series. So it follows that Λ is a subring.

Proposition 2.10. If $f \in \Lambda^n$ and $g \in \Lambda^m$, then $fg \in \Lambda^{n+m}$.

Proof. Write $f = (f_N)$ and $g = (g_N)$, so by Proposition 2.8, $f_N \in \Lambda_N^n$ and $g_N \in \Lambda_N^m$ for each $N \in \mathbb{N}$, so $f_N g_N \in \Lambda_N^{n+m}$. Applying Proposition 2.8 again, we have $fg = (f_N g_N) \in \Lambda^{n+m}$. □

Corollary 2.2. The space Λ is a subring of $\mathbb{C}[\mathrm{x}]$ (therefore, it is an algebra).

2.3.4. Other bases for the ring of symmetric functions

We saw that $\{\mathrm{m}_\lambda : \lambda \vdash n\}$ is a basis for Λ^n, but there are other important bases for Λ^n. To present them, we need to define the following symmetric functions.

Definition 2.19. The n-th **elementary symmetric function** e_n is the sum of all products of n distinct variables. More precisely, for any $n \geq 1$,

$$\mathrm{e}_n := \mathrm{m}_{(1^n)} = \sum_{i_1 < \cdots < i_n} x_{i_1} \cdots x_{i_n}$$

and we set $\mathrm{e}_0 := 1$.

The n-th **complete homogeneous function** h_n is the sum of all monomials of degree n, that is,

$$\mathrm{h}_n := \sum_{\lambda \vdash n} \mathrm{m}_\lambda = \sum_{i_1 \leq \cdots \leq i_n} x_{i_1} \cdots x_{i_n}$$

for $n \geq 1$. Again we set $\mathrm{h}_0 := 1$.

The n-th **power sum** is defined by

$$\mathrm{p}_n := \mathrm{m}_{(n)} = \sum_{i=1}^{\infty} x_i^n$$

for all $n \geq 1$.

Example 2.22. We have

$$\mathrm{e}_1 = x_1 + x_2 + x_3 + \cdots,$$

$$\mathrm{e}_2 = x_1x_2 + x_1x_3 + x_2x_3 + \cdots,$$

and

$$\mathrm{e}_3 = x_1x_2x_3 + x_1x_3x_4 + x_2x_3x_4 + \cdots.$$

Example 2.23. We have $\mathrm{h}_1 = \mathrm{e}_1$ and

$$\mathrm{h}_2 = \mathrm{e}_2 + x_1^2 + x_2^2 + x_3^2 + \cdots.$$

Note 2.2. Sometimes it is useful to work with finite variables. In this case, Definition 2.19 remains the same.

In order to prove some results about these functions, we define their generating functions. We will see that it is actually easier to manipulate the generating functions than the functions themselves.

Definition 2.20. The **generating functions** for e_n, h_n and p_n are, respectively,

$$\mathrm{E}(z) := \sum_{n=0}^{\infty} \mathrm{e}_n z^n, \quad \mathrm{H}(z) := \sum_{n=0}^{\infty} \mathrm{h}_n z^n \quad \text{and} \quad \mathrm{P}(z) := \sum_{n=1}^{\infty} \mathrm{p}_n z^{n-1}.$$

We can see the generating functions as formal power series in the ring $\mathbb{C}[z, x_1, x_2, \ldots]$. The next proposition shows nice expressions for them.

Proposition 2.11. We have $\mathrm{E}(z) = \prod_{i=1}^{\infty}(1 + x_i z)$, $\mathrm{H}(z) = \prod_{i=1}^{\infty}(1 - x_i z)^{-1}$ and $\mathrm{P}(z) = \frac{\mathrm{d}}{\mathrm{d}z}\sum_{i=1}^{\infty} \ln(1 - x_i z)^{-1}$.

Proof. For the first identity, we expand the product, obtaining

$$\begin{aligned}\prod_{i=1}^{\infty}(1 + x_i z) &= 1 + (x_1 + x_2 + \cdots)z + (x_1x_2 + x_1x_3 + x_2x_3 + \cdots)z^2 \\ &\quad + (x_1x_2x_3 + x_1x_2x_4 + x_1x_3x_4 + x_2x_3x_4 + \cdots)z^3 + \cdots \\ &= \mathrm{e}_0 + \mathrm{e}_1 z + \mathrm{e}_2 z^2 + \mathrm{e}_3 z^3 + \cdots = \mathrm{E}(z).\end{aligned}$$

To prove the second one, we are going to use the geometric series $(1-x_iz)^{-1} = \sum_{j=0}^{\infty}(x_iz)^j$. We can write

$$\begin{aligned}\prod_{i=1}^{\infty}(1-x_iz)^{-1} &= \prod_{i=1}^{\infty}\sum_{j=0}^{\infty}(x_iz)^j \\ &= (1+x_1z+x_1^2z^2+\cdots)(1+x_2z+x_2^2z^2+\cdots)(1+x_3z+x_3^2z^2+\cdots)\cdots \\ &= 1+(x_1+x_2+\cdots)z+\left[\left(x_1^2+x_2^2+\cdots\right)+(x_1x_2+x_1x_3+x_2x_3+\cdots)\right]z^2+\cdots \\ &= \mathrm{h}_0+\mathrm{h}_1z+\mathrm{h}_2z^2+\cdots = \mathrm{H}(z).\end{aligned}$$

For the last one, the logarithm should be interpreted as its power series, obtaining

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}z}\sum_{i=1}^{\infty}\ln(1-x_iz)^{-1} &= \frac{\mathrm{d}}{\mathrm{d}z}\sum_{i,j=1}^{\infty}\frac{x_i^jz^j}{j} = \sum_{i,j=1}^{\infty}x_i^jz^{j-1} \\ &= (x_1+x_2+\cdots)+\left(x_1^2+x_2^2+\cdots\right)z+\cdots \\ &= \sum_{n=1}^{\infty}\mathrm{p}_nz^{n-1} = \mathrm{P}(z).\end{aligned}$$

□

Note 2.3. The calculations in the last proof are just formal, that is, we are not worried about the convergence of the series, since we are not evaluating them. The notation $\frac{\mathrm{d}}{\mathrm{d}z}$ is not the derivative defined in Real Analysis, but just the formal derivative used in Algebra. Whenever we write some expression that does not look like a formal power series, we should interpret them as the power series that we know from its Taylor expansion, for example,

$$\frac{1}{1-x} = 1+x+x^2+x^3+\cdots$$

and

$$\mathrm{e}^x = 1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\cdots.$$

The expressions obtained in Proposition 2.11 are used to get results that would be harder to prove using just the definitions.

Corollary 2.3. We have $\mathrm{H}(z)\mathrm{E}(-z) = 1$.

Proof. From Proposition 2.11,

$$\mathrm{H}(z)\mathrm{E}(-z) = \prod_{i=1}^{\infty}(1-x_iz)^{-1}\prod_{i=1}^{\infty}(1+x_i(-z)) = \prod_{i=1}^{\infty}\frac{1-x_iz}{1-x_iz} = 1.$$

□

Corollary 2.4. We have

$$\mathrm{H}(z) = \frac{1}{\mathrm{E}(-z)} = \exp\left(\sum_{n=1}^{\infty}\frac{p_nz^n}{n}\right)$$

Proof. The first equality follows from Corollary 2.3. From Proposition 2.11 and by the product property of logarithms,

$$\sum_{n=1}^{\infty} \frac{\mathrm{p}_n z^n}{n} = \int \mathrm{P}(z)\mathrm{d}z = \sum_{i=1}^{\infty} \ln \frac{1}{1-x_i z} = \ln\left(\prod_{i=1}^{\infty} \frac{1}{1-x_i z}\right) = \ln \mathrm{H}(z). \tag{2.7}$$

Now we get the result by taking the exponential on both sides. □

The next corollary sets a relation between the generating functions $\mathrm{E}(z)$, $\mathrm{H}(z)$ and $\mathrm{P}(z)$.

Corollary 2.5. We have

$$\mathrm{P}(z) = \frac{\mathrm{H}'(z)}{\mathrm{H}(z)} \quad \text{and} \quad \mathrm{P}(-z) = \frac{\mathrm{E}'(z)}{\mathrm{E}(z)}.$$

Proof. From equation (2.7),

$$\mathrm{P}(z) = \frac{\mathrm{d}}{\mathrm{d}z} \ln \mathrm{H}(z) = \frac{\mathrm{H}'(z)}{\mathrm{H}(z)}.$$

For the second identity, we apply Corollary (2.3). □

Corollary 2.6. For any natural $n \geq 1$,

$$\sum_{i=0}^{n} (-1)^i \mathrm{e}_i \mathrm{h}_{n-i} = 0.$$

Proof. Expanding the product $\mathrm{H}(z)\mathrm{E}(-z)$, we obtain

$$\mathrm{H}(z)\mathrm{E}(-z) = \sum_{i,j=0}^{\infty} (-1)^i \mathrm{e}_i \mathrm{h}_j z^{i+j}.$$

Fixing a natural $n \geq 1$, we see that the coefficient of z^n is the sum

$$\sum_{\substack{i,j \in \mathbb{N} \\ i+j=n}} (-1)^i \mathrm{e}_i \mathrm{h}_j = \sum_{i=0}^{n} (-1)^i \mathrm{e}_i \mathrm{h}_{n-i},$$

which must be zero due to Corollary (2.3). □

Let $\lambda = (\lambda_1, \ldots, \lambda_k)$, if $\mathrm{f} = \mathrm{e}$, p or h, then we define the symmetric function $\mathrm{f}_\lambda = \mathrm{f}_{\lambda_1}\mathrm{f}_{\lambda_2}\cdots\mathrm{f}_{\lambda_n}$. By Definition 2.19, $\mathrm{f}_{\lambda_i} \in \Lambda^{\lambda_i}$ for each i. Also, applying Proposition 2.10 successively, we have $\mathrm{f}_\lambda \in \Lambda^{|\lambda|}$, therefore, the formal power series e_λ, h_λ and p_λ are well-defined symmetric functions.

Proposition 2.12. We have $\mathrm{H}(z) = \sum_{\lambda \in \mathbb{Y}} \frac{\mathrm{p}_\lambda z^{|\lambda|}}{\xi_\lambda}$, where $\xi_\lambda := \prod_{i=1}^{\infty} i^{m_i(\lambda)} m_i(\lambda)!$ and $m_i(\lambda)$ is the number of times that the number i repeats itself in λ.

Proof. By Equation (2.7), we have $\ln \mathrm{H}(z) = \sum_{k=1}^{\infty} \frac{\mathrm{p}_k z^k}{k}$. Taking the exponential on both sides, we obtain

$$\mathrm{H}(z) = \exp\left(\sum_{k=1}^{\infty} \frac{\mathrm{p}_k z^k}{k}\right) = \prod_{k=1}^{\infty} \exp\left(\frac{\mathrm{p}_k z^k}{k}\right) = \prod_{k=1}^{\infty}\left(\sum_{n=0}^{\infty} \frac{\mathrm{p}_k^n z^{kn}}{k^n n!}\right).$$

So

$$\mathrm{H}(z) = \left(1 + \mathrm{p}_1 z + \frac{\mathrm{p}_1^2 z^2}{2!} + \frac{\mathrm{p}_1^3 z^3}{3!} + \cdots\right)\left(1 + \frac{\mathrm{p}_2 z^2}{2} + \frac{\mathrm{p}_2^2 z^4}{2^2 2!} + \frac{\mathrm{p}_2^3 z^6}{2^3 3!} + \cdots\right)\cdots$$

and we see that $\mathrm{H}(z)$ is an infinite sum such that each term is the product

$$\prod_{i=1}^{\infty} \frac{(\mathrm{p}_i z^i)^{m_i}}{i^{m_i} m_i!} = z^{\sum i m_i} \prod_{i=1}^{\infty} \frac{1}{i^{m_i} m_i!} \prod_{i=1}^{\infty} \mathrm{p}_i^{m_i} \tag{2.8}$$

and the sum is over all the sequences $(m_1, m_2, m_3, \ldots)$ with a finite number of non-zero terms. Since each sequence $(m_1, m_2, \ldots)$ defines a partition $\lambda = (1^{m_1}, 2^{m_2}, \ldots)$ the sum is over all partitions. Moreover, each term in (2.8) assumes the form $\frac{\mathrm{p}_\lambda z^{|\lambda|}}{\xi_\lambda}$, as we wanted to prove. □

Example 2.24. We know that $\{\mathrm{m}_{(3)}, \mathrm{m}_{(2,1)}, \mathrm{m}_{(1,1,1)}\}$ is a basis for Λ^3. Ordering this basis accordingly to the lexicographic order $((3) > (2,1) > (1,1,1))$, one obtain the coordinates of each p_λ for $\lambda \vdash 3$:

$$\begin{aligned}
\mathrm{p}_{(3)} &= 1\mathrm{m}_{(3)} + 0\mathrm{m}_{(2,1)} + 0\mathrm{m}_{(1,1,1)};\\
\mathrm{p}_{(2,1)} &= 1\mathrm{m}_{(3)} + 1\mathrm{m}_{(2,1)} + 0\mathrm{m}_{(1,1,1)};\\
\mathrm{p}_{(1,1,1)} &= 1\mathrm{m}_{(3)} + 3\mathrm{m}_{(2,1)} + 6\mathrm{m}_{(1,1,1)}.
\end{aligned}$$

The matrix formed by the coordinates is triangular with no zeros on the main diagonal, so it is invertible. Consequently, the set $\{\mathrm{p}_{(3)}, \mathrm{p}_{(2,1)}, \mathrm{p}_{(1,1,1)}\}$ is linearly independent, and since $\dim \Lambda^3 = 3$, it must be a basis for Λ^3.

Example 2.25. Let us write all the e_λ's, for $\lambda \vdash 4$, in the basis $\{\mathrm{m}_\lambda : \lambda \vdash 4\}$. We have

$$\begin{aligned}
\mathrm{e}_{(4)} &= \mathrm{m}_{(1,1,1,1)};\\
\mathrm{e}_{(3,1)} &= 4\mathrm{m}_{(1,1,1,1)} + \mathrm{m}_{(2,1,1)};\\
\mathrm{e}_{(2,2)} &= 6\mathrm{m}_{(1,1,1,1)} + 2\mathrm{m}_{(2,1,1)} + \mathrm{m}_{(2,2)};\\
\mathrm{e}_{(2,1,1)} &= 12\mathrm{m}_{(1,1,1,1)} + 5\mathrm{m}_{(2,1,1)} + 2\mathrm{m}_{(2,2)} + \mathrm{m}_{(3,1)};\\
\mathrm{e}_{(1,1,1,1)} &= 24\mathrm{m}_{(1,1,1,1)} + 12\mathrm{m}_{(2,1,1)} + 6\mathrm{m}_{(2,2)} + 4\mathrm{m}_{(3,1)} + \mathrm{m}_{(4)}
\end{aligned}$$

and we observe the same pattern from the previous example, thus, $\{\mathrm{e}_\lambda : \lambda \vdash 4\}$ is a basis for Λ^4.

These examples provide intuition to enunciate the next theorem.

Theorem 2.2. The sets $\{\mathrm{p}_\lambda : |\lambda| = n\}$, $\{\mathrm{e}_\lambda : |\lambda| = n\}$ and $\{\mathrm{h}_\lambda : |\lambda| = n\}$ are linear bases for Λ^n.

Proof. Since all the sets above have $\dim \Lambda^n$ elements, it suffices to prove that they are linearly independent or generate Λ^n. The proofs presented here are found in (MACDONALD, 1998) and (SAGAN, 2001). Given $\lambda \vdash n$, we have

$$\mathrm{p}_\lambda = \mathrm{p}_{\lambda_1} \cdots \mathrm{p}_{\lambda_\ell} = (x_1^{\lambda_1} + x_2^{\lambda_1} + \cdots)(x_1^{\lambda_2} + x_2^{\lambda_2} + \cdots) \cdots (x_1^{\lambda_\ell} + x_2^{\lambda_\ell} + \cdots),$$

so p_λ is a sum of monomials in the form $x_{i_1}^{\lambda_1} \cdots x_{i_\ell}^{\lambda_\ell}$. When we have $i_j = i_k$ for some j and k, the exponents are summed, then the monomial above is associated with a partition with higher natural order. Therefore, when we write p_λ as a linear combination of m_μ's, we always have $\lambda \trianglelefteq \mu$, and since the natural order implies the lexicographic order, we can write

$$\mathrm{p}_\lambda = \sum_{\mu \geq \lambda} c_\mu \mathrm{m}_\mu.$$

Furthermore, we have $c_\lambda \neq 0$, so the matrix of coordinates of p_λ's is triangular with no zeros on the main diagonal, that is, the set $\{\mathrm{p}_\lambda : |\lambda| = n\}$ is linearly independent. For the e_λ's, we use the same strategy. Given $\lambda \vdash n$ and its transpose λ', the product $\mathrm{e}_{\lambda'} = \mathrm{e}_{\lambda'_1} \cdots \mathrm{e}_{\lambda'_\ell}$ is a sum of monomials of the form

$$(x_{i_1} \cdots x_{i_{\lambda'_1}})(x_{j_1} \cdots x_{j_{\lambda'_2}}) \cdots (x_{k_1} \cdots x_{k_{\lambda'_\ell}}) = x^{\alpha_1} \cdots x^{\alpha_r}.$$

By definition of e_k, we can suppose $i_1 < i_2 < \cdots < i_{\lambda'_1}$, $j_1 < j_2 < \cdots < j_{\lambda'_2}$ ans so on. Now imagine that we are entering the numbers $i_1, i_2, \ldots, i_{\lambda'_1}$ in the first column of the diagram of λ, $j_1, j_2, \ldots, j_{\lambda'_2}$ in the second and so on. Since these indexes are strictly increasing, all the numbers $\leq t$ must occur in the top t rows of λ, so we have $\alpha_1 + \cdots + \alpha_t \leq \lambda_1 + \cdots + \lambda_t$ for every t, that is, $\alpha \trianglelefteq \lambda$. By Propositions 2.2 and 2.3, it follows that

$$\mathrm{e}_{\lambda'} = \sum_{\mu \leq \lambda} d_\mu \mathrm{m}_\mu.$$

This argument also proves that $d_\lambda = 1$, so again we have a triangular matrix of coordinates with no zeros on the main diagonal, thus $\{\mathrm{e}_\lambda : |\lambda| = n\}$ is a basis. Finally, we shall prove that the h_λ's, for $\lambda \vdash n$, generate Λ^n. For this, it suffices to prove that the e_k's are polynomials in h_l's. We know that $\mathrm{e}_1 = \mathrm{h}_1$, and by Corollary 2.6, we have for every natural k,

$$\mathrm{e}_k = \mathrm{h}_1 \mathrm{e}_{k-1} - \mathrm{h}_2 \mathrm{e}_{k-2} + \cdots + (-1)^{k+1} \mathrm{h}_k,$$

hence we get the result by induction over k. □

Corollary 2.7. The sets $\{\mathrm{p}_\lambda : \lambda \in \mathbb{Y}\}$, $\{\mathrm{e}_\lambda : \lambda \in \mathbb{Y}\}$ and $\{\mathrm{h}_\lambda : \lambda \in \mathbb{Y}\}$ are linear bases for Λ.

Corollary 2.8. Every $f \in \Lambda$ can be written as a polynomial in each of the following sets of variables: $\{\mathrm{p}_1, \mathrm{p}_2, \mathrm{p}_3, \ldots\}$, $\{\mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3, \ldots\}$, $\{\mathrm{h}_1, \mathrm{h}_2, \mathrm{h}_3, \ldots\}$.

2.3.5. Skew-symmetric Polynomials

Ironically, to prove more interesting properties about symmetric functions, we should look at the skew-symmetric polynomials (or anti-symmetric polynomials).

Definition 2.21. A polynomial p is called **skew-symmetric** if $\sigma \cdot p = \operatorname{sgn}(\sigma)p$, for every permutation $\sigma \in \mathscr{S}_N$. The set of all skew symmetric polynomials in N variables is denoted by A_N.

Example 2.26. For every transposition $(i\ j)$, we have $\operatorname{sgn}(i\ j) = -1$. So if p is skew-symmetric, we have $(i\ j)\cdot p = -p$.

The next properties are straightforward.

Proposition 2.13. Let $p, q \in \mathbb{C}[x_1, \ldots, x_N]$, then:

(a) If p and q are symmetric, then pq is symmetric;

(b) If p and q are skew-symmetric, then pq is symmetric;

(c) If p is symmetric and q is skew-symmetric, then pq is skew-symmetric.

If q divides p, then the same is valid for the quotient p/q.

Consider a set of N variables $\mathrm{x} = \{x_1, \ldots, x_N\}$ and a sequence $\alpha = (\alpha_1, \ldots, \alpha_N)$ of N non-negative integers. Let $\mathrm{x}^\alpha := x_1^{\alpha_1} \cdots x_N^{\alpha_N}$. The **anti-symmetrization** of x^α is the polynomial

$$\mathrm{a}_\alpha(x_1, \ldots, x_N) := \sum_{\sigma \in \mathscr{S}_N} \operatorname{sgn}(\sigma)\sigma(\mathrm{x}^\alpha).$$

Example 2.27. Let $\alpha = (2,1)$. The anti-symmetrization of $\mathrm{x}^\alpha = x_1^2 x_2$ is $\mathrm{a}_\alpha(x_1, x_2) = x_1^2 x_2 - x_2^2 x_1$.

Proposition 2.14. The polynomial a_α is skew-symmetric.

Proof. Take any transposition $(i\ j) \in \mathscr{S}_N$ $(i \neq j)$, then

$$\begin{aligned}
(i\ j)\cdot \mathrm{a}_\alpha(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_N) &= \mathrm{a}_\alpha(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_N) \\
&= \sum_{\sigma \in \mathscr{S}_N} \operatorname{sgn}(\sigma)\sigma\left(x_1^{\alpha_1} \cdots x_j^{\alpha_i} \cdots x_i^{\alpha_j} \cdots x_N^{\alpha_N}\right) \\
&= \sum_{\sigma \in \mathscr{S}_N} \operatorname{sgn}(\sigma)\,(\sigma \circ (i\ j))\,(\mathrm{x}^\alpha) \\
&= \sum_{\tau \in \mathscr{S}_N} \operatorname{sgn}(\tau \circ (i\ j))\tau(\mathrm{x}^\alpha) \\
&= -\sum_{\tau \in \mathscr{S}_N} \operatorname{sgn}(\tau)\tau(\mathrm{x}^\alpha) = -\mathrm{a}_\alpha.
\end{aligned}$$

Since every $\sigma \in \mathscr{S}_N$ is a composition of transpositions, we obtain the result. □

Corollary 2.9. If $\alpha_i = \alpha_j$ for $i \neq j$, then $\mathrm{a}_\alpha = 0$.

Proof. By Proposition 2.14, $-\mathrm{a}_\alpha = (i\ j)\cdot \mathrm{a}_\alpha = \mathrm{a}_\alpha$. □

Let $\delta := (N-1, N-2, \ldots, 1, 0)$. By Corollary 2.9, we may assume $\alpha_1 > \alpha_2 > \cdots > \alpha_N \geq 0$, so that we have $\alpha = \lambda + \delta$ for some partition λ satisfying $\ell(\lambda) \leq N$. By definition of determinant, we have

$$\mathrm{a}_{\lambda+\delta} = \det\left[x_i^{\lambda_j+N-j}\right]_{i,j=1}^{N}.$$

In particular, taking $\lambda = \varnothing$, we have

$$\mathrm{a}_\delta = \det\left[x_i^{N-j}\right]_{i,j=1}^{N} = \prod_{1\leq i<j\leq N}(x_i - x_j),$$

which is the well-known **Vandermonde determinant**.

Proposition 2.15. For every partition λ such that $\ell(\lambda) \leq N$, the polynomial $\mathrm{a}_{\lambda+\delta}$ is divisible by a_δ.

Proof. Let i and j be distinct integers such that $x_i = x_j$. Since $\mathrm{a}_{\lambda+\delta}$ is skew-symmetric, we have

$$\begin{aligned}\mathrm{a}_{\lambda+\delta}(x_1,\ldots,x_i,\ldots,x_j,\ldots,x_N) &= (i\ j)\cdot \mathrm{a}_{\lambda+\delta}(x_1,\ldots,x_i,\ldots,x_j,\ldots,x_N)\\ &= -\mathrm{a}_{\lambda+\delta}(x_1,\ldots,x_i,\ldots,x_j,\ldots,x_N),\end{aligned}$$

hence $\mathrm{a}_{\lambda+\delta}(x_1,\ldots,x_i,\ldots,x_j,\ldots,x_N) = 0$. So for every pair (i,j) such that $1 \leq i < j \leq N$, the difference $x_i - x_j$ divides $\mathrm{a}_{\lambda+\delta}$, thus $\mathrm{a}_{\lambda+\delta}$ is divisible by a_δ. □

Proposition 2.16. We have $\mathrm{A}_N = \{\mathrm{a}_\delta s : s \in \Lambda_N\}$.

Proof. By Proposition 2.13 and the fact that a_δ is skew-symmetric, we have $\mathrm{a}_\delta s \in \mathrm{A}_N$ for every $s \in \Lambda_N$. Conversely, if p is skew-symmetric, we can repeat the proof of Proposition 2.15 to show that a_δ divides p, so $s = p/\mathrm{a}_\delta$ is a well-defined symmetric polynomial, by Proposition 2.13, hence $p = \mathrm{a}_\delta s$. □

Corollary 2.10. A_N is a vector subspace.

Proof. By Proposition 2.16, A_N is closed under sum and scalar multiplication, since Λ_N is a vector subspace. □

Corollary 2.11. Λ_N and A_N are linearly isomorphic.

Proof. The linear application $\varphi : \Lambda_N \to \mathrm{A}_N$ defined by $\varphi(s) = \mathrm{a}_\delta s$ is the desired isomorphism. □

As the m_λ's are the "smallest" symmetric polynomials, the $\mathrm{a}_{\lambda+\delta}$'s are the "smallest" skew-symmetric polynomials, so the next proposition is analogous to the one for symmetric polynomials (Proposition 2.1).

Proposition 2.17. The set $\{\mathrm{a}_{\lambda+\delta}(x_1,\ldots,x_N) : \lambda \in \mathbb{Y}, \ell(\lambda) \leq N\}$ is a basis for A_N.

2.4. Schur Polynomials and Schur Functions

2.4.1. Schur Polynomials

The polynomial $\mathrm{a}_{\lambda+\delta}/\mathrm{a}_\delta$, well-defined by Proposition 2.15, has a special name.

Definition 2.22. For every partition λ with $\ell(\lambda) \leq N$, the **Schur polynomial** in N variables corresponding to λ is given by

$$\mathrm{s}_{\lambda,N} = \mathrm{s}_\lambda(x_1,\ldots,x_N) \coloneqq \frac{\mathrm{a}_{\lambda+\delta}(x_1,\ldots,x_N)}{\mathrm{a}_\delta(x_1,\ldots,x_N)} = \frac{\det\left[x_i^{\lambda_j+N-j}\right]_{i,j=1}^N}{\prod_{1\leq i<j\leq N}(x_i-x_j)}.$$

If $\ell(\lambda) > N$, we define $\mathrm{s}_{\lambda,N} \coloneqq 0$.

Besides the name, this definition is due to Jacobi (1841). However, it was Schur (1901) who discovered important results about such polynomials, as their relevance to the representation theory of symmetric groups.

Example 2.28. For the empty partition $\varnothing = (0,0,0,\ldots)$, we have $\mathrm{s}_\varnothing = 1$.

Example 2.29. We have

$$\begin{aligned}
\mathrm{s}_{(1)}(x_1) &= x_1;\\
\mathrm{s}_{(1)}(x_1,x_2) &= x_1+x_2;\\
&\vdots\\
\mathrm{s}_{(1)}(x_1,\ldots,x_N) &= x_1+x_2+\cdots+x_N.
\end{aligned}$$

Example 2.30. For any $n \geq 1$,

$$\begin{aligned}
\mathrm{s}_{(n)}(x) &= x^n\\
\mathrm{s}_{(n,0)}(x,y) &= x^n + x^{n-1}y + \cdots + xy^{n-2} + y^n.
\end{aligned}$$

Example 2.31. For any pair $n_1 \geq n_2$,

$$\mathrm{s}_{(n_1,n_2)}(x_1,x_2) = \mathrm{s}_{(n_1,n_2)}(x_1,x_2) = \frac{x_1^{n_1+1}x_2^{n_2} - x_1^{n_1}x_2^{n_2+1}}{x_1-x_2} = x_1^{n_1}x_2^{n_2} + x_1^{n_1-1}x_2^{n_2+1} + \cdots + x_1^{n_2}x_2^{n_1}.$$

Corollary 2.12. Every Schur polynomial is symmetric.

Proof. Since $\mathrm{a}_{\lambda+\delta}$ and a_{δ} are skew-symmetric, the quotient must be symmetric. □

Proposition 2.18. The degree of the Schur polynomial s_{λ} is $|\lambda|$.

Proof. We have $\deg(\mathrm{s}_{\lambda}) = \deg(\mathrm{a}_{\lambda+\delta}) - \deg(\mathrm{a}_{\delta})$. Notice that

$$\deg(\mathrm{a}_{\delta}) = \binom{N}{2} = \frac{N(N-1)}{2}$$

and

$$\deg(\mathrm{a}_{\lambda+\delta}) = \sum_{i=1}^{N}(\lambda_i + N - i) = |\lambda| + \frac{N(N-1)}{2},$$

hence it follows the result. □

The next proposition is essential to define the Schur symmetric functions in the next section. We show that when we make $x_N = 0$ in a Schur polynomial in N variables, we obtain the Schur polynomial in $N-1$ variables (both with respect to the same partition λ).

Proposition 2.19. For every $\lambda \in \mathbb{Y}$ and $N \in \mathbb{N}$, we have $\pi_N \mathrm{s}_{\lambda,N} = \mathrm{s}_{\lambda,N-1}$.

Proof. For $N \leq \ell(\lambda)$, the result is straightforward. Suppose $N > \ell(\lambda)$. When $x_N = 0$, we have

$$\mathrm{a}_{\delta}(x_1,\ldots,x_{N-1},0) = x_1\cdots x_{N-1}\prod_{1\leq i<j\leq N-1}(x_i - x_j) = x_1\cdots x_{N-1}\mathrm{a}_{\delta}(x_1,\ldots,x_{N-1})$$

Moreover, $x_N^{\lambda_j+N-j} = 1$ when $j = N$ and $x_N^{\lambda_j+N-j} = 0$ for $j = 1,\ldots,N-1$. So by Laplace expansion in the last row of the matrix that defines $\mathrm{a}_{\lambda+\delta}$, we have

$$\begin{aligned}\mathrm{a}_{\lambda+\delta}(x_1,\ldots,x_{N-1},0) &= \det\left[x_i^{\lambda_j+N-j}\right]_{i,j=1}^{N-1}\\ &= x_1\cdots x_{N-1}\det\left[x_i^{\lambda_j+N-1-j}\right]_{i,j=1}^{N-1}\\ &= x_1\cdots x_{N-1}\mathrm{a}_{\lambda+\delta}(x_1,\ldots,x_{N-1}).\end{aligned}$$

Now we divide $\mathrm{a}_{\lambda+\delta}$ by a_{δ} to obtain the result. □

The Schur polynomials provide a new and important basis for Λ_N.

Corollary 2.13. The set $\{\mathrm{s}_{\lambda,N} : \ell(\lambda) \leq N\}$ is a basis for Λ_N.

Proof. Take the isomorphism $\varphi : \Lambda_N \to \mathrm{A}_N$ from Corollary 2.11. Since the $\mathrm{a}_{\lambda+\delta}$'s form a basis for A_N (Proposition 2.17), the elements

$$\varphi^{-1}(\mathrm{a}_{\lambda+\delta}) = \varphi^{-1}(\mathrm{a}_{\delta}\mathrm{s}_{\lambda,N}) = \mathrm{s}_{\lambda,N}$$

form a basis for Λ_N. □

The following theorem is essential for this study, since we need it to define an important probability measure on $\mathbb{Y}$. To prove it, we need the two lemmas below. The first one is a classic result from linear algebra and the second is in (BORODIN; GORIN, 2012).

Lemma 2.1. (Cauchy-Binet identity) Let $n \leq k$. For two matrices $A = [a_{i,j}]_{n\times k}$ and $B = [b_{i,j}]_{k\times n}$,

$$\det(AB) = \sum_{1\leq \ell_1 < \cdots < \ell_n \leq k} \det[a_{i,\ell_j}] \det[b_{\ell_i,j}].$$

Particularly, for $n = k$, we obtain the usual identity $\det(AB) = \det(A)\det(B)$. The proposition still works for $k = \infty$, in this case, each entry of AB is a series, so we must verify that each one converges.

Lemma 2.2. (Cauchy determinant formula) The following identity holds:

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

Theorem 2.3. (Cauchy identity) If $a = (a_1, \ldots, a_N)$ and $b = (b_1, \ldots, b_N)$ are such that $|a_i b_j| < 1$ for all $i, j = 1, \ldots, N$, then

$$\sum_{\lambda\in\mathbb{Y}} s_\lambda(a) s_\lambda(b) = \prod_{i,j=1}^{N} \frac{1}{1-a_i b_j}. \tag{2.9}$$

Proof. (BAIK, 2018) The sum on the left side of (2.9) is finite since $s_\lambda = 0$ when $\ell(\lambda) > N$. By definition,

$$\sum_{\lambda\in\mathbb{Y}} s_\lambda(a) s_\lambda(b) = \sum_{\lambda_1\geq\cdots\geq\lambda_N\geq 0} \frac{\det\left[a_i^{\lambda_j+N-j}\right]\det\left[b_i^{\lambda_j+N-j}\right]}{\prod_{i<j}(a_i-a_j)(b_i-b_j)}.$$

Let $\mu_j := \lambda_j + N - j$, so that $\mu_1 > \cdots > \mu_N$. By Cauchy-Binet identity (Lemma 2.1), the numerator is equal to

$$\sum_{\mu_1>\cdots>\mu_N\geq 0} \det\left[a_i^{\mu_j}\right]\det\left[b_i^{\mu_j}\right] = \det\left[\sum_{n=0}^{\infty} a_i^n b_j^n\right] = \det\left[\frac{1}{1-a_i b_j}\right],$$

since $\sum_{n=0}^{\infty} a_i^n b_j^n < \infty$. So it follows from Lemma 2.2, that

$$\sum_{\lambda\in\mathbb{Y}} s_\lambda(a) s_\lambda(b) = \frac{\det\left[\frac{1}{1-a_i b_j}\right]}{\prod_{i<j}(a_i-a_j)(b_i-b_j)} = \prod_{i,j}\frac{1}{1-a_i b_j},$$

as we wanted to prove. □

2.4.2. Schur Functions

Consider an infinite set of variables $\mathrm{x} = \{x_1, x_2, \ldots\}$. Let $\mathrm{s}_{\lambda,N} = \mathrm{s}_\lambda(x_1, \ldots, x_N)$ be the Schur polynomial in N variables with respect to λ. From the properties discussed in the previous sections, the sequence $(\mathrm{s}_{\lambda,N})_{N\in\mathbb{N}}$ is a symmetric function.

Definition 2.23. The symmetric function given by $\mathrm{s}_\lambda(\mathrm{x}) \coloneqq (\mathrm{s}_{\lambda,N}(x_1, \ldots, x_N))_{N\in\mathbb{N}}$ is called the **Schur function** with respect to λ.

This is the classical definition of Schur functions, but there is a definition based on combinatorics, briefly presented in the introduction of this chapter. For this modern approach, see (SAGAN, 2001, Ch. 4), there it is shown the equivalence of these two definitions.

Proposition 2.20. The set of Schur functions $\{\mathrm{s}_\lambda : \lambda \in \mathbb{Y}\}$ is a basis for Λ.

Proof. By Corollary 2.13, for every $N \in \mathbb{N}$, $\{\mathrm{s}_{\lambda,N} : \ell(\lambda) \leq N\}$ is a basis for Λ_N. Let $f = (f_N)_{N\in\mathbb{N}} \in \Lambda$, then $f_N \in \Lambda_N$ for each N, so

$$f_N = \sum_{\lambda\in\mathbb{Y}} c_\lambda \mathrm{s}_{\lambda,N}$$

(remember we defined $\mathrm{s}_{\lambda,N} = 0$ when $N < \ell(\lambda)$). Since $\pi_N f_N = f_{N-1}$, we have

$$f_{N-1} = \sum_{\lambda\in\mathbb{Y}} c_\lambda \mathrm{s}_{\lambda,N-1},$$

hence the constants in the linear combination remain the same for every natural N. Therefore, we conclude that

$$f = (f_N)_{N\in\mathbb{N}} = \left(\sum_{\lambda\in\mathbb{Y}} c_\lambda \mathrm{s}_{\lambda,N}\right)_{N\in\mathbb{N}} = \sum_{\lambda\in\mathbb{Y}} c_\lambda \left(\mathrm{s}_{\lambda,N}\right)_{N\in\mathbb{N}} = \sum_{\lambda\in\mathbb{Y}} c_\lambda \mathrm{s}_\lambda,$$

that is, $\{\mathrm{s}_\lambda : \lambda \in \mathbb{Y}\}$ generates Λ. Now, to prove that the set of Schur functions is linearly independent, take any finite linear combination such that

$$c_1 \mathrm{s}_{\lambda_1} + \cdots + c_k \mathrm{s}_{\lambda_k} = 0$$

and choose $N \geq \max \ell(\lambda_i)$. So we have

$$c_1 \mathrm{s}_{\lambda_1,N} + \cdots + c_k \mathrm{s}_{\lambda_k,N} = 0,$$

which implies that $c_1 = \cdots = c_k = 0$ since $\{\mathrm{s}_{\lambda,N} : \ell(\lambda) \leq N\}$ is linearly independent in Λ_N. □

Example 2.32. The Schur function parametrized by the empty partition $\varnothing = (0, 0, \ldots)$ is $\mathrm{s}_\varnothing = 1$.

Example 2.33. For $\lambda = (1)$, we have the Schur polynomials

$$\begin{aligned} s_\lambda(x_1) &= x_1 \\ s_\lambda(x_1,x_2) &= x_1+x_2 \\ s_\lambda(x_1,x_2,x_3) &= x_1+x_2+x_3 \\ &\vdots \\ s_\lambda(x_1,\ldots,x_N) &= x_1+x_2+\cdots+x_N, \end{aligned}$$

so the Schur function with respect to λ is the formal series

$$s_\lambda(x_1,x_2,x_3,\ldots) = x_1+x_2+x_3+\cdots.$$

By Proposition 2.8, each Schur function is expressed as a polynomial in the generators h_k or e_k of Λ. The formulas are given below. The proof is technical and it is not in the scope of this work. One proof is found in (MACDONALD, 1998) and a second one, by Lindström (1973), is presented in (SAGAN, 2001).

Theorem 2.4. (Jacobi-Trudi formulas) Given a partition λ, the Schur function s_λ can be expressed as

$$s_\lambda = \det\left[h_{\lambda_i-i+j}\right]_{i,j=1}^{\ell(\lambda)} = \det\left[e_{\lambda'_i-i+j}\right]_{i,j=1}^{\ell(\lambda')},$$

where we agree that $h_k = e_k = 0$ for $k < 0$.

We end this section presenting the Cauchy identity for Schur functions (in Theorem 2.3, we obtained it for finite variables). Consider two sets of variables $\mathrm{x} = \{x_1,x_2,x_3,\ldots\}$ and $\mathrm{y} = \{y_1,y_2,y_3,\ldots\}$. Now we have to specify which set of variables is considered for some symmetric function. For example, the power sums in the variables x and y are denoted by $p_k(\mathrm{x})$ and $p_k(\mathrm{y})$. For any partition $\lambda \in \mathbb{Y}$, remember the number ξ_λ defined in Proposition 2.12.

Theorem 2.5 (Cauchy identity)**.** Let x and y be two sets of infinite variables, then

$$\prod_{i,j=1}^{\infty}\frac{1}{1-x_iy_j} = \sum_{\lambda\in\mathbb{Y}} s_\lambda(\mathrm{x})s_\lambda(\mathrm{y}) = \sum_{\lambda\in\mathbb{Y}}\frac{p_\lambda(\mathrm{x})p_\lambda(\mathrm{y})}{\xi_\lambda}. \tag{2.10}$$

Proof. Since the set $\{x_iy_j\}_{i,j=1}^{\infty}$ is countable, we can take an enumeration $\mathrm{w} = \{w_1,w_2,\ldots\}$ of this set. For each $n \in \mathbb{N}$, we have

$$p_n(\mathrm{w}) = \sum_{k=1}^{\infty} w_k^n = \sum_{i,j=1}^{\infty} x_i^n y_j^n = p_n(\mathrm{x})p_n(\mathrm{y}),$$

so for all $\lambda \in \mathbb{Y}$, we obtain $p_\lambda(\mathrm{w}) = p_\lambda(\mathrm{x})p_\lambda(\mathrm{y})$. By Propositions 2.11 and 2.12, we have

$$H(z;\mathrm{w}) = \prod_{k=1}^{\infty}\frac{1}{1-w_kz} = \sum_{\lambda\in\mathbb{Y}}\frac{p_\lambda(\mathrm{w})z^{|\lambda|}}{\xi_\lambda} = \sum_{\lambda\in\mathbb{Y}}\frac{p_\lambda(\mathrm{x})p_\lambda(\mathrm{y})z^{|\lambda|}}{\xi_\lambda}.$$

Replacing each w_k by $x_i y_j$, we obtain

$$\prod_{i,j=1}^{\infty} \frac{1}{1-x_i x_j z} = \sum_{\lambda \in \mathbb{Y}} \frac{\mathrm{p}_\lambda(\mathrm{x})\mathrm{p}_\lambda(\mathrm{y}) z^{|\lambda|}}{\xi_\lambda}.$$

Finally, we evaluate this expression on $z = 1$ to obtain one of the identities.

Remember that we have already proved the Cauchy identity for Schur polynomials in Theorem 2.3. So for any $N \in \mathbb{N}$, and for the variables $\mathrm{x} = \{x_1, \ldots, x_N, 0, 0, \ldots\}$ and $\mathrm{y} = \{y_1, \ldots, y_N, 0, 0, \ldots\}$, we have

$$\prod_{i,j=1}^{\infty} \frac{1}{1-x_i x_j} = \prod_{i,j=1}^{N} \frac{1}{1-x_i x_j} = \sum_{\lambda \in \mathbb{Y}} \mathrm{s}_\lambda(x_1, \ldots, x_N)\mathrm{s}_\lambda(y_1, \ldots, y_N) = \sum_{\lambda \in \mathbb{Y}} \mathrm{s}_\lambda(\mathrm{x})\mathrm{s}_\lambda(\mathrm{y}).$$

Since this identity holds for every $N \in \mathbb{N}$, we can send N to infinity to obtain the formal power series identity. □

2.4.3. A Scalar Product on the Ring of Symmetric Functions

Since $\{\mathrm{h}_\lambda : \lambda \in \mathbb{Y}\}$ and $\{\mathrm{m}_\lambda : \lambda \in \mathbb{Y}\}$ are bases for Λ, we can define a scalar product $\langle \cdot, \cdot \rangle$ in Λ by requiring that such bases are **biorthogonal**, that is, for all $\lambda, \mu \in \mathbb{Y}$,

$$\langle \mathrm{h}_\lambda, \mathrm{m}_\mu \rangle := \delta_{\lambda\mu} = \begin{cases} 1, & \text{if } \lambda = \mu; \\ 0, & \text{if } \lambda \neq \mu. \end{cases}$$

We claim that the Schur functions provide an orthonormal basis to Λ with respect to the scalar product defined above. In order to prove this result, we need the technical lemma below (proof in (MACDONALD, 1998)).

Lemma 2.3. Let $\{u_\lambda : \lambda \in \mathbb{Y}\}$ and $\{v_\lambda : \lambda \in \mathbb{Y}\}$ be bases of Λ such that u_λ, $v_\lambda \in \Lambda^n$ for all $n \geq 0$. The following conditions are equivalent:

(a) $\langle u_\lambda, v_\mu \rangle = \delta_{\lambda\mu}$ for all λ, $\mu \in \mathbb{Y}$;

(b) $\sum_{\lambda \in \mathbb{Y}} u_\lambda(x) v_\lambda(y) = \prod_{i,j} \frac{1}{1-x_i y_j}$.

Corollary 2.14. The basis $\{\mathrm{s}_\lambda : \lambda \in \mathbb{Y}\}$ of Λ is orthonormal.

Proof. By Cauchy identity (Theorem 2.5),

$$\prod_{i,j} \frac{1}{1-x_i y_j} = \sum_{\lambda \in \mathbb{Y}} \mathrm{s}_\lambda(\mathrm{x})\mathrm{s}_\lambda(\mathrm{y}),$$

so it follows from Lemma 2.3 that $\langle \mathrm{s}_\lambda, \mathrm{s}_\mu \rangle = \delta_{\lambda\mu}$ for every $\lambda, \mu \in \mathbb{Y}$. □

Corollary 2.15. The scalar product $\langle \cdot, \cdot \rangle$ is symmetric, that is, $\langle f, g \rangle = \langle g, f \rangle$ for any symmetric functions f and g.

Proof. Write f and g as a linear combination of s_λ's and apply Corollary 2.14. □

2.4.4. Skew Schur Functions

Given partitions μ and ν, we have $s_\mu s_\nu \in \Lambda$. Since $\{s_\lambda : \lambda \in \mathbb{Y}\}$ is an orthonormal basis, we can write

$$s_\mu s_\nu = \sum_{\lambda \in \mathbb{Y}} c^\lambda_{\mu\nu} s_\lambda, \quad c^\lambda_{\mu\nu} := \langle s_\lambda, s_\mu s_\nu \rangle .$$

The coefficients $c^\lambda_{\mu\nu}$ are known as **Littlewood-Richardson coefficients** (MACDONALD, 1998). They are non-negative integers that can be computed by a combinatorial rule (LITTLEWOOD; RICHARDSON, 1934).

Definition 2.24. For λ, $\mu \in \mathbb{Y}$, the symmetric function given by

$$s_{\lambda/\mu} := \sum_{\nu \in \mathbb{Y}} c^\lambda_{\mu\nu} s_\nu$$

is called **skew Schur function** parametrized by λ and μ.

Directly from the definitions above, we deduce several properties.

Proposition 2.21. We have the following properties:

(a) If $|\mu| + |\nu| \neq |\lambda|$, then $c^\lambda_{\mu\nu} = 0$ (this property guarantees that the sum in the definition of $s_{\lambda/\mu}$ is finite);

(b) If $|\lambda| < |\mu|$, then $s_{\lambda/\mu} = 0$;

(c) If $|\lambda| \geq |\mu|$, then $s_{\lambda/\mu}$ is a homogeneous symmetric function of degree $|\lambda| - |\mu|$;

(d) For all λ, μ, $\nu \in \mathbb{Y}$, we have $\langle s_{\lambda/\mu}, s_\nu \rangle = \langle s_\lambda, s_\mu s_\nu \rangle$;

(e) For all $\lambda \in \mathbb{Y}$, $s_{\lambda/\varnothing} = s_\lambda$;

(f) For all $\lambda \in \mathbb{Y}$, $s_{\lambda/\lambda} = 1$;

(g) For all $\lambda, \mu \in \mathbb{Y}$, $s_{\lambda/\mu}(0,0,0,\ldots) = \delta_{\lambda\mu}$.

Proof. (a) We have $\deg(s_\mu s_\nu) = |\mu| + |\nu|$ and $\deg(s_\lambda) = |\lambda|$, moreover, $s_\mu s_\nu$ and s_λ are homogeneous. So by definition of $s_\mu s_\nu$, each $c^\lambda_{\mu\nu}$ with $|\mu| + |\nu| \neq |\lambda|$ must be zero.

(b) If $|\lambda| \leq |\mu|$, then $|\lambda| < |\mu| + |\nu|$ for any ν. So by (a), $c^\lambda_{\mu\nu} = 0$ for all ν, thus $s_{\lambda/\mu} = 0$.

(c) By (a), $c^\lambda_{\mu\nu} = 0$ for $|\mu| + |\nu| \neq |\lambda|$. So in the definition of $s_{\lambda/\mu}$, the sum is over partitions ν such that $|\nu| = |\lambda| - |\mu|$, hence $\deg(s_{\lambda/\mu}) = |\lambda| - |\mu|$.

(d) By definition of $s_{\lambda/\mu}$, we have $c^\lambda_{\mu\nu} = \langle s_{\lambda/\mu}, s_\nu \rangle$.

(e) By item (d), for all $\nu \in \mathbb{Y}$,

$$\langle s_{\lambda/\varnothing}, s_\nu \rangle = \langle s_\lambda, s_\varnothing s_\nu \rangle = \langle s_\lambda, s_\nu \rangle,$$

so we must have $s_{\lambda/\varnothing} = s_\lambda$.

(f) Remember that $s_\varnothing = 1$, so applying (d) for $\nu = \varnothing$, we have

$$\langle s_{\lambda/\lambda}, 1 \rangle = \langle s_\lambda, s_\lambda s_\varnothing \rangle = \langle s_\lambda, s_\lambda \rangle = 1,$$

by Corollary 2.14. On the other hand, by (c), $\deg(s_{\lambda/\lambda}) = 0$, so $s_{\lambda/\lambda} = c = cs_\varnothing$, where $c \in \mathbb{C}$, thus

$$1 = \langle c, 1 \rangle = \langle cs_\varnothing, s_\varnothing \rangle = c \langle s_\varnothing, s_\varnothing \rangle = c.$$

(g) If $\lambda = \mu$, we have $s_{\lambda/\mu}(0) = 1$ from (f). If $\lambda \neq \mu$, notice that $s_\nu(0) = 0$ for every ν such that $|\nu| > 0$. Therefore, by definition of $s_{\lambda/\mu}$,

$$s_{\lambda/\mu}(0) = c^\lambda_{\mu\varnothing} s_\varnothing(0) = c^\lambda_{\mu\varnothing},$$

but

$$c^\lambda_{\mu\varnothing} = \langle s_\lambda, s_\mu s_\varnothing \rangle = \langle s_\lambda, s_\mu \rangle = 0.$$

□

We want to generalize the Cauchy identity (Theorem 2.5) for skew Schur functions. We start with the two lemmas below. Consider again two sets of variables $\mathrm{x} = \{x_1, x_2, \ldots\}$ and $\mathrm{y} = \{y_1, y_2, \ldots\}$. For convenience, let $\prod(\mathrm{x}, \mathrm{y}) := \prod_{i,j} \frac{1}{1 - x_i y_j}$.

Lemma 2.4. For any $\lambda, \mu \in \mathbb{Y}$, we have

$$\sum_{\lambda \in \mathbb{Y}} s_{\lambda/\mu}(\mathrm{x}) s_\lambda(\mathrm{y}) = s_\mu(\mathrm{y}) \prod(\mathrm{x}, \mathrm{y}).$$

Proof. (BAIK, 2018) From definition of skew Schur function and from Cauchy identity,

$$\begin{aligned}
\sum_{\lambda \in \mathbb{Y}} s_{\lambda/\mu}(\mathrm{x}) s_\lambda(\mathrm{y}) &= \sum_{\lambda \in \mathbb{Y}} \sum_{\nu \in \mathbb{Y}} c^\lambda_{\mu\nu} s_\nu(\mathrm{x}) s_\lambda(\mathrm{y}) \\
&= \sum_{\nu \in \mathbb{Y}} s_\nu(\mathrm{x}) \sum_{\lambda \in \mathbb{Y}} c^\lambda_{\mu\nu} s_\lambda(\mathrm{y}) \\
&= \sum_{\nu \in \mathbb{Y}} s_\nu(\mathrm{x}) s_\mu(\mathrm{y}) s_\nu(\mathrm{y}) \\
&= s_\mu(\mathrm{y}) \prod(\mathrm{x}, \mathrm{y}).
\end{aligned}$$

□

Note that we can define the Schur function in two variables $s_\lambda(\mathrm{x}, \mathrm{y})$, since the union of variables x and y is still countable.

Lemma 2.5. For any $\lambda, \mu \in \mathbb{Y}$, we have

$$s_\lambda(x,y) = \sum_{\mu \in \mathbb{Y}} s_{\lambda/\mu}(x) s_\mu(y).$$

Proof. (BAIK, 2018) Let $z = \{z_1, z_2, \ldots\}$ be another set of variables. By Lemma 2.4 and Cauchy identity,

$$\begin{aligned}\sum_{\lambda \in \mathbb{Y}} \sum_{\mu \in \mathbb{Y}} s_{\lambda/\mu}(x) s_\lambda(z) s_\mu(y) &= \sum_{\mu \in \mathbb{Y}} s_\mu(y) \sum_{\lambda \in \mathbb{Y}} s_{\lambda/\mu}(x) s_\lambda(z) \\ &= \sum_{\mu \in \mathbb{Y}} s_\mu(z) s_\mu(y) \prod(x,z) \\ &= \prod(x,z) \prod(y,z).\end{aligned}$$

On the other hand, again by Cauchy identity,

$$\sum_{\lambda \in \mathbb{Y}} s_\lambda(x,y) s_\lambda(z) = \prod(x,z) \prod(y,z),$$

and since $\{s_\lambda(z)\}$ is a basis we get the result by matching the coordinates. □

Theorem 2.6 (Skew summation formula)**.** For any $\lambda, \nu \in \mathbb{Y}$, we have

$$s_{\lambda/\nu}(x,y) = \sum_{\mu \in \mathbb{Y}} s_{\lambda/\mu}(x) s_{\mu/\nu}(y).$$

Proof. (BAIK, 2018) Let z be a new variable, by Lemma 2.5,

$$\begin{aligned}\sum_{\mu \in \mathbb{Y}} \sum_{\nu \in \mathbb{Y}} s_{\lambda/\mu}(x) s_{\mu/\nu}(y) s_\nu(z) &= \sum_{\mu \in \mathbb{Y}} s_{\lambda/\mu}(x) s_\mu(y,z) \\ &= s_\lambda(x,y,z) = \sum_{\nu \in \mathbb{Y}} s_{\lambda/\nu}(x,y) s_\nu(z)\end{aligned}$$

and since $\{s_\nu(z)\}$ is a basis, we have the result by matching the coordinates. □

Theorem 2.7 (Skew Cauchy identity)**.** For any $\lambda, \mu \in \mathbb{Y}$, we have

$$\sum_{\nu \in \mathbb{Y}} s_{\nu/\lambda}(x) s_{\nu/\mu}(y) = \prod(x,y) \sum_{\kappa \in \mathbb{Y}} s_{\lambda/\kappa}(y) s_{\mu/\kappa}(x).$$

Proof. (BAIK, 2018) Let z and w be two new sets of variables. By Lemmas 2.4 and 2.5,

$$\begin{aligned}\sum_{\nu,\lambda,\mu \in \mathbb{Y}} s_{\nu/\lambda}(x) s_{\nu/\mu}(y) s_\lambda(z) s_\mu(w) &= \sum_{\nu,\lambda \in \mathbb{Y}} s_{\nu/\lambda}(x) s_\nu(y,w) s_\lambda(z) \\ &= \sum_{\lambda \in \mathbb{Y}} s_\lambda(y,w) \prod(x,y) \prod(x,w) s_\lambda(z) \\ &= \prod(x,y) \sum_{\lambda \in \mathbb{Y}} s_\lambda(y,w) \prod(x,w) s_\lambda(z) \\ &= \prod(x,y) \sum_{\lambda,\kappa \in \mathbb{Y}} \prod(x,w) s_{\lambda/\kappa}(y) s_\kappa(w) s_\lambda(z) \\ &= \prod(x,y) \sum_{\lambda,\kappa,\mu \in \mathbb{Y}} s_{\lambda/\kappa}(y) s_{\mu/\kappa}(x) s_\mu(w) s_\lambda(z).\end{aligned}$$

Now we use that $\{s_\mu(w)\}$ and $\{s_\lambda(z)\}$ are bases to match the coordinates and get the result. □

We also have formulas for the skew Schur functions in terms of e_n's and h_n's, analogous to Jacobi-Trudi formulas (Theorem 2.4). For a proof, see (MACDONALD, 1998).

Proposition 2.22 (Skew Jacobi-Trudi formulas)**.** For any $\lambda, \mu \in \mathbb{Y}$, we have

$$s_{\lambda/\mu} = \det\left[h_{\lambda_i-\mu_j-i+j}\right]_{i,j=1}^{\max\{\ell(\lambda),\ell(\mu)\}} = \det\left[e_{\lambda'_i-\mu'_j-i+j}\right]_{i,j=1}^{\max\{\ell(\lambda'),\ell(\mu')\}},$$

where we agree that $h_k = e_k = 0$ for $k < 0$.

Proposition 2.23. If $\mu \not\subset \lambda$, then $s_{\lambda/\mu} = 0$.

Proof. Let $n = \max\{\ell(\lambda), \ell(\mu)\}$. Since $\mu \not\subset \lambda$, there exists k such that

$$\mu_1 \geq \cdots \geq \mu_{k-1} \geq \mu_k > \lambda_k \geq \lambda_{k+1} \geq \cdots \geq \lambda_n.$$

So for every pair (i, j) satisfying $i \in \{k, \ldots, n\}$ and $j \in \{1, \ldots, k\}$, we have $\lambda_i - \mu_j - i + j < 0$, thus $h_{\lambda_i-\mu_j-i+j} = 0$. By Proposition 2.22,

$$s_{\lambda/\mu} = \det\begin{pmatrix} f_{11} & \cdots & f_{1k} & f_{1(k+1)} & \cdots & f_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ f_{(k-1)1} & \cdots & f_{(k-1)k} & f_{(k-1)(k+1)} & \cdots & f_{(k-1)n} \\ 0 & \cdots & 0 & f_{k(k+1)} & \cdots & f_{kn} \\ 0 & \cdots & 0 & f_{(k+1)(k+1)} & \cdots & f_{(k+1)n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & f_{n(k+1)} & \cdots & f_{nn} \end{pmatrix}, \quad f_{ij} := h_{\lambda_i-\mu_j-i+j},$$

so applying the determinant formula $\det\left(\begin{smallmatrix} A & C \\ 0 & B \end{smallmatrix}\right) = \det A \det B$ for block matrices, we obtain $s_{\lambda/\mu} = 0$. □

2.5. Specializations

Until now, we have worked with symmetric functions interpreting them as formal power series, considering infinitely countable indeterminants $x_1, x_2, \ldots$. However, one could want to evaluate these functions in a sequence of complex numbers $\alpha_1, \alpha_2, \ldots$, so now we had better think about convergence of such series.

Example 2.34. Let $\alpha = (\alpha_1, \alpha_2, \ldots)$ be a sequence of complex numbers such that $S = \sum_i |\alpha_i| < \infty$ and fix $k \in \mathbb{N}$. For every $N \in \mathbb{N}$, we have

$$\sum_{i=1}^{N} \left|\alpha_i^k\right| = \sum_{i=1}^{N} |\alpha_i|^k \leq \left(\sum_{i=1}^{N} |\alpha_i|\right)^k \leq \left(\sum_{i=1}^{\infty} |\alpha_i|\right)^k = S^k,$$

so $\sum_i \alpha_i^k$ converges absolutely for any k and the evaluation of the power sum $\mathrm{p}_k(\mathrm{x})$ on α is well-defined. Since the power sums form a algebraic basis for Λ (Corollary 2.8), the evaluation of every symmetric function (as the Schur functions) on the variables α is well-defined. Moreover, it is immediate that the evaluation map

$$\begin{aligned} \alpha : \Lambda &\longrightarrow \mathbb{C} \\ f &\longmapsto f(\alpha) \end{aligned}$$

is an algebra homomorphism, that is, for any $f_1, f_2 \in \Lambda$ and $c \in \mathbb{C}$,

$$(f_1+f_2)(\alpha) = f_1(\alpha)+f_2(\alpha), \quad (f_1 f_2)(\alpha) = f_1(\alpha) f_2(\alpha) \quad \text{and} \quad (cf_1)(\alpha) = cf_1(\alpha).$$

The last example motivates the definition of a more general tool to evaluate symmetric functions.

Definition 2.25. A **specialization** is an algebra homomorphism $\rho : \Lambda \to \mathbb{C}$. If $\mathrm{s}_\lambda(\rho) \geq 0$ for every $\lambda \in \mathbb{Y}$, then ρ is called a **Schur-positive specialization**.

Note 2.4. We denote the application of ρ in $f \in \Lambda$ by $f(\rho)$ (instead of $\rho(f)$), as if we were evaluating f in a set of complex numbers.

Proposition 2.24. If ρ is a Schur-positive specialization, then $\mathrm{s}_{\lambda/\mu}(\rho) \geq 0$ for every $\lambda, \mu \in \mathbb{Y}$.

Proof. By the definition of skew Schur functions (Definition 2.24), we have

$$\mathrm{s}_{\lambda/\mu}(x) = \sum_{\nu \in \mathbb{Y}} c_{\mu\nu}^{\lambda} \mathrm{s}_\nu(x)$$

where the sum is finite (by Proposition 2.21(b)) and the Littlewood-Richardson coefficients $c_{\mu\nu}^{\lambda}$ are non-negative integers (MACDONALD, 1998; LITTLEWOOD; RICHARDSON, 1934). Since ρ is a Schur-positive specialization, we have

$$\mathrm{s}_{\lambda/\mu}(\rho) = \sum_{\nu \in \mathbb{Y}} \underbrace{c_{\mu\nu}^{\lambda}}_{\geq 0} \underbrace{\mathrm{s}_\nu(\rho)}_{\geq 0} \geq 0,$$

as we wanted to prove. □

The theorem below provides a characterization for Schur-positive specializations.

Theorem 2.8. Every Schur-positive specialization ρ is parametrized by two real sequences $\alpha = (\alpha_1 \geq \alpha_2 \geq \cdots \geq 0)$ and $\beta = (\beta_1 \geq \beta_2 \geq \cdots \geq 0)$ such that $\sum_i (\alpha_i + \beta_i) < \infty$, and an additional parameter $\gamma \geq 0$. We denote $\rho = (\alpha; \beta; \gamma)$ and ρ is defined by its evaluation on power sums. More precisely,

$$\mathrm{p}_1 \mapsto \mathrm{p}_1(\alpha;\beta;\gamma) = \gamma + \sum_{i=1}^{\infty} (\alpha_i + \beta_i); \tag{2.11}$$

$$\mathrm{p}_k \mapsto \mathrm{p}_k(\alpha;\beta;\gamma) = \sum_{i=1}^{\infty} \left(\alpha_i^k + (-1)^{k-1} \beta_i^k \right), \quad k \geq 2, \tag{2.12}$$

or equivalently,

$$\sum_{k=0}^{\infty} \mathrm{h}_k(\alpha;\beta;\gamma)z^k = \mathrm{e}^{\gamma z}\prod_{i=1}^{\infty}\frac{1+\beta_i z}{1-\alpha_i z}. \tag{2.13}$$

Proof. This proof is evolved and beyond the scope of this work. There exists several independently found proofs (EDREI, 1953; THOMA, 1964; VERSHIK; KEROV, 1981; KEROV; OKOUNKOV; OLSHANSKI, 1998; KEROV, 2003; OKOUNKOV, 1994). We only show the equivalence between conditions (2.11)-(2.12) and (2.13). By Equation (2.7),

$$\begin{aligned}
\sum_{k=0}^{\infty} \mathrm{h}_k(\rho)z^k = \mathrm{H}(z) &= \exp\left(\sum_{k=1}^{\infty}\frac{\mathrm{p}_k(\rho)}{k}z^k\right)\\
&= \mathrm{e}^{\gamma z}\exp\left(\sum_{k=1}^{\infty}\sum_{i=1}^{\infty}\frac{\alpha_i^k-(-\beta_i)^k}{k}z^k\right)\\
&= \mathrm{e}^{\gamma z}\exp\left(\sum_{k=1}^{\infty}\left(\frac{\mathrm{p}_k(\alpha)}{k}-\frac{\mathrm{p}_k(-\beta)}{k}\right)z^k\right)\\
&= \mathrm{e}^{\gamma z}\exp\left(\ln\mathrm{H}(z;\alpha)-\ln\mathrm{H}(z;-\beta)\right)\\
&= \mathrm{e}^{\gamma z}\frac{\mathrm{H}(z;\alpha)}{\mathrm{H}(z;-\beta)}\\
&= \mathrm{e}^{\gamma z}\prod_{i=1}^{\infty}\frac{1+\beta_i z}{1-\alpha_i z}.
\end{aligned}$$

□

Given two real sequences $\alpha = (\alpha_1 \geq \alpha_2 \geq \cdots \geq 0)$ and $\alpha' = (\alpha'_1 \geq \alpha'_2 \geq \cdots \geq 0)$, we can reorder the terms of the two sequences to define another non-increasing sequence, denoted by $\alpha \cup \alpha'$.

Proposition 2.25 (Union of Schur-positive specializations)**.** Given two Schur-positive specializations $\rho = (\alpha;\beta;\gamma)$ and $\rho' = (\alpha';\beta';\gamma')$, the specialization

$$\rho \cup \rho' := (\alpha \cup \alpha';\beta \cup \beta';\gamma+\gamma')$$

is Schur-positive.

Proof. The sequences $\alpha \cup \alpha' = (a_1 \geq a_2 \geq \cdots \geq 0)$ and $\beta \cup \beta' = (b_1 \geq b_2 \geq \cdots \geq 0)$ satisfy, by Theorem 2.8,

$$\sum_i (a_i + b_i) = \sum_j (\alpha_j + \beta_j) + \sum_j (\alpha'_j + \beta'_j) < \infty.$$

Moreover, we have $\gamma + \gamma' \geq 0$, so the specialization $\rho \cup \rho'$ is Schur-positive (again by Theorem 2.8). □

Example 2.35. Let $\alpha = (\alpha_1 \geq \alpha_2 \geq \cdots \geq 0)$ be such that $\sum_i \alpha_i < \infty$. By Theorem 2.8, the evaluation map $x_i \mapsto \alpha_i$ is exactly the Schur-positive specialization $\rho = (\alpha; 0; 0)$.

Consider two sets of variables x, y and a symmetric function f. We may want to work with the two variable function $\mathrm{f}(\mathrm{x},\mathrm{y})$. Let ρ_1 and ρ_2 be Schur-positive specializations. Since $\mathrm{f}(\mathrm{x},\mathrm{y})$ is symmetric in x, we can apply the specialization ρ_1 with respect to x, obtaining the symmetric function in y given by $\mathrm{f}(\rho_1,\mathrm{y})$. Then we can apply ρ_2 with respect to y, obtaining $\mathrm{f}(\rho_1,\rho_2)$. The technical proposition below shows that we could just apply $\rho_1 \cup \rho_2$ on $\mathrm{f}(\mathrm{x})$, now in one variable, obtaining the same result (BAIK, 2018).

Lemma 2.6. Let $\mathrm{f} \in \Lambda$ and consider two Schur-positive specializations ρ_1 and ρ_2. Then $\mathrm{f}(\rho_1,\rho_2) = \mathrm{f}(\rho_1 \cup \rho_2)$.

We finish the chapter with a series of useful propositions, found in (BORODIN; GORIN, 2012), about some specializations applied to Schur functions (proofs in (BAIK, 2018)). We denote $\mathbf{0} := (0,0,0,\ldots)$.

Proposition 2.26 (Single α specialization)**.** Let $c \geq 0$ and $\alpha = (c,0,0,\ldots)$. Consider the Schur-positive specialization $\rho = (\alpha; \mathbf{0}; 0)$. If λ/μ is a horizontal strip ($\mu \preccurlyeq \lambda$), then $\mathrm{s}_{\lambda/\mu}(\rho) = c^{|\lambda|-|\mu|}$. Otherwise, we have $\mathrm{s}_{\lambda/\mu}(\rho) = 0$.

Proposition 2.27 (Single β specialization)**.** Let $c \geq 0$ and $\beta = (c,0,0,\ldots)$. Consider the Schur-positive specialization $\rho = (\mathbf{0}; \beta; 0)$. If λ/μ is a vertical strip ($\mu \preccurlyeq_{\mathrm{v}} \lambda$), then $\mathrm{s}_{\lambda/\mu}(\rho) = c^{|\lambda|-|\mu|}$. Otherwise, we have $\mathrm{s}_{\lambda/\mu}(\rho) = 0$.

Note that Propositions 2.26 and 2.27 can be applied to Schur functions, since $\mathrm{s}_\lambda = \mathrm{s}_{\lambda/\varnothing}$ (2.21(e)). Given a $\lambda \in \mathbb{Y}$, remember that $\dim(\lambda)$ is defined as the number of possible standard Young tableaux with shape λ (see Section 2.2.2).

Proposition 2.28 (Single γ specialization)**.** Let $\gamma \geq 0$. For the Schur-positive specialization $\rho = (\mathbf{0}; \mathbf{0}; \gamma)$, we have

$$\mathrm{s}_\lambda(\rho) = \frac{\gamma^{|\lambda|}}{|\lambda|!} \dim(\lambda).$$

If you were brave enough to make this far in this text, we gift you with the next chapter, which will be definitely more enjoyable. In Chapter 3, we make use of most results and formulas presented until now. They are indispensable to define probabilities on the set of partitions.

CHAPTER

3

SCHUR PROCESSES

The main goal of this chapter is to define a stochastic process on $\mathbb{Y}$ and even on $\mathbb{Y}^N$. As we stated in the Introduction, our motivation is the fact that the TASEP can be obtained as a "projection" of an evolution of objects on $\mathbb{Y}^N$ (the Block-Push Process). First, we are going to define a probability measure on $\mathbb{Y}$, then we construct a Markov chain on $\mathbb{Y}$ that preserves such class of measures. After that, we repeat the construction for $\mathbb{Y}^N$. We will see why the formulas presented in Chapter 2 are so useful. We are mostly following the lecture notes (BORODIN; GORIN, 2012), (BAIK, 2018) and (FERRARI, 2019).

3.1. Summary of the Chapter

We define a probability measure on $\mathbb{Y}$ through the formula

$$\mathbb{S}_{\mathrm{x},\mathrm{y}}(\lambda) := \frac{\mathrm{s}_\lambda(\mathrm{x})\mathrm{s}_\lambda(\mathrm{y})}{\prod \frac{1}{1-x_i y_j}}, \qquad \lambda \in \mathbb{Y},$$

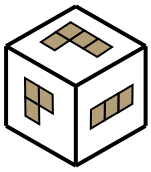
where we can assume for now that $\mathrm{x} = \{x_1, x_2, \ldots\}$ and $\mathrm{y} = \{y_1, y_2, \ldots\}$ are sets of real numbers. If we also ensure that $\prod_{i,j=1}^{\infty} \frac{1}{1-x_i y_j}$ is finite, then it is immediate by Cauchy identity (Equation (2.2)) that $\mathbb{S}_{\mathrm{x},\mathrm{y}}(\lambda)$ is a well-defined probability measure on $\mathbb{Y}$, that is, $\sum_{\lambda\in\mathbb{Y}} \mathbb{S}_{\mathrm{x},\mathrm{y}}(\lambda) = 1$. This measure is called **Schur measure**. In other words, it is an "infinite die" with partitions on its faces (Figure 12).

A random evolution of partitions is defined by the transitional probability

$$p_{\lambda\to\mu} := \frac{1}{\prod \frac{1}{1-x_i y_j}} \frac{\mathrm{s}_\mu(\mathrm{x})}{\mathrm{s}_\lambda(\mathrm{x})} \mathrm{s}_{\mu/\lambda}(\mathrm{y}),$$

which is the probability for going from λ to μ. These probabilities provide a way to obtain a random sequence of partitions, where the next partition depends only on the current

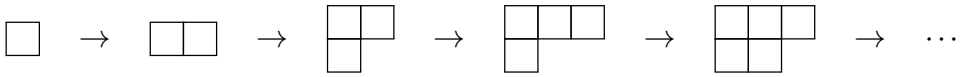
Figure 12 – The Schur measure is a way to select random partitions.



Source: Elaborated by the author.

partition (see Figure 13). This is called a Markov chain. The well-definition of this Markov chain relies again on the Cauchy identities (2.2) and (2.3).

Figure 13 – A pictorial example of a Markov chain on partitions, which produces a "growth" of a diagram.



Source: Elaborated by the author.

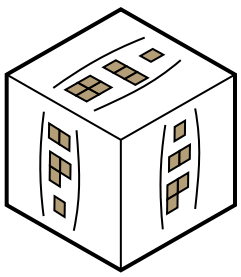
A nice property we obtain from these definitions is the preservation of the Schur measure in this Markov chain. More precisely, if a partition is randomly chosen according to a Schur measure (the die in Figure 12), the next partition of the chain is also chosen according to a Schur measure with slightly different parameters (a slightly different die).

After this we repeat a similar construction on $\mathbb{Y}^N$. First, we define a probability measure on $\mathbb{Y}^N$ given by

$$\mathbb{S}_{\rho_1,\ldots,\rho_N,\rho^-}\left(\lambda^{(1)},\ldots,\lambda^{(N)}\right) := \frac{s_{\lambda^{(1)}}(\rho_1)s_{\lambda^{(2)}/\lambda^{(1)}}(\rho_2)\cdots s_{\lambda^{(N)}/\lambda^{(N-1)}}(\rho_N)s_{\lambda^{(N)}}(\rho^-)}{\prod(\rho_1,\rho^-)\cdots\prod(\rho_N,\rho^-)}.$$

This measure is called **Schur process**. The objects ρ_1, ..., ρ_N and ρ^- are special parameters we will define later. The well-definition of this measure is again obtained by Cauchy identities. To continue our analogy, the Schur process is an infinite die with vectors of partitions on its faces (Figure 14).

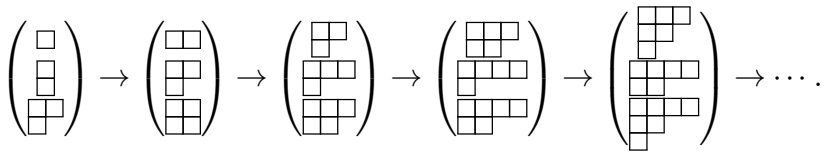
Figure 14 – The Schur process is a way to select random vectors of partitions.



Source: Elaborated by the author.

A Markov chain can be defined on $\mathbb{Y}^N$ to obtain sequences of random vectors of partitions (Figure 15). Analogously, we have a preservation property similar to the one obtained for the Schur measure.

Figure 15 – A pictorial example of a Markov chain on $\mathbb{Y}^N$.



Source: Elaborated by the author.

The importance of this framework is that we can recover (using a specific parametrization for the Markov chain) the Block-Push process (see Section 1.1.1). Then we use some general results (as the preservation property) to obtain a simple formula for a tricky problem on the TASEP model: what is the probability that all particles have passed a given point at time t? The answer is given simply by the Schur measure

$$\mathbb{P}_{\text{TASEP}}(y_N(t) - 1 + N \geq \ell) = \mathbb{S}_{\sigma_1,\sigma_2}(\nu_N \geq \ell),$$

for some parametrizations σ_1 and σ_2. This formula, obtained purely by algebraic methods, is crucial to advance in this work.

3.2. Probabilities on Partitions

3.2.1. Schur Measure

Let x and y be two sets of variables and consider the product set $\mathrm{xy} = \{x_i y_j\}_{i,j=1}^{\infty}$. From Equation (2.7), we have

$$\prod_{i,j=1}^{\infty} \frac{1}{1 - x_i y_j z} = \mathrm{H}(z; \mathrm{xy}) = \exp\left(\sum_{k=1}^{\infty} \frac{\mathrm{p}_k(\mathrm{xy}) z^k}{k}\right) = \exp\left(\sum_{k=1}^{\infty} \frac{\mathrm{p}_k(\mathrm{x})\mathrm{p}_k(\mathrm{y}) z^k}{k}\right).$$

Let $z = 1$ and recall the notation $\prod(\mathrm{x},\mathrm{y}) = \prod_{i,j=1}^{\infty} \frac{1}{1-x_i y_j}$. From Cauchy identity (Theorem 2.5) and from equation above we have

$$\sum_{\lambda} \mathrm{s}_\lambda(\mathrm{x})\mathrm{s}_\lambda(\mathrm{y}) = \prod(\mathrm{x},\mathrm{y}) = \exp\left(\sum_{k=1}^{\infty} \frac{\mathrm{p}_k(\mathrm{x})\mathrm{p}_k(\mathrm{y})}{k}\right).$$

The identity above is useful because the Schur-positive specializations are defined by their evaluation on power sums (Theorem 2.8). So given two specializations[1] ρ_1 and ρ_2, we have the specialized Cauchy identity

$$\prod(\rho_1, \rho_2) = \exp\left(\sum_{k=1}^{\infty} \frac{\mathrm{p}_k(\rho_1)\mathrm{p}_k(\rho_2)}{k}\right) = \sum_{\lambda} \mathrm{s}_\lambda(\rho_1)\mathrm{s}_\lambda(\rho_2). \tag{3.1}$$

Of course we need to guarantee that $\prod(\rho_1, \rho_2)$ is finite, since it depends on the convergence of a series.

[1] From now on, we assume that all specializations are Schur-positive.

Example 3.1. For the trivial specialization $\rho_0 = (\mathbf{0};\mathbf{0};0)$ and any other specialization ρ_1, we have $\prod(\rho_0,\rho_1) = 1$.

Example 3.2. Let $c \in [0,1]$. Consider the specializations

$$\rho_1 = (\underbrace{(1,\ldots,1}_{n},0,0,\ldots);\mathbf{0};0) \quad \text{and} \quad \rho_2 = (\mathbf{0};(\underbrace{c,\ldots,c}_{m},0,0,\ldots),0).$$

By Theorem 2.8, we have $p_k(\rho_1) = n$ and $p_k(\rho_2) = mc(-c)^{k-1}$ for every $k \geq 1$. So applying the alternating series test, we obtain

$$\prod(\rho_1,\rho_2) = \exp\left(\sum_{k=1}^{\infty} \frac{p_k(\rho_1)p_k(\rho_2)}{k}\right) = \exp\left(nmc\sum_{k=1}^{\infty} \frac{(-c)^{k-1}}{k}\right) < \infty.$$

Definition 3.1. (OKOUNKOV, 2001) Let ρ_1 and ρ_2 be two specializations satisfying $\prod(\rho_1,\rho_2) < \infty$. The measure defined by

$$\mathbb{S}_{\rho_1,\rho_2}(\lambda) := \frac{s_\lambda(\rho_1)s_\lambda(\rho_2)}{\prod(\rho_1,\rho_2)}, \qquad \lambda \in \mathbb{Y}, \tag{3.2}$$

is called the **Schur measure** parametrized by ρ_1 and ρ_2. Since $\mathbb{Y}$ is discrete, the underlying σ-algebra is the power set of $\mathbb{Y}$.

Note 3.1. The function defined in (3.2) is non-negative, since ρ_1 and ρ_2 are Schur-positive specializations (Definition 2.25). Moreover, it is a probability measure on $\mathbb{Y}$, since $\sum_\lambda \mathbb{S}_{\rho_1,\rho_2}(\lambda) = 1$, by Equation (3.1).

Example 3.3. Let $\rho_0 = (\mathbf{0};\mathbf{0};0)$, then $\prod(\rho,\rho_0) = 1$ for any other specialization ρ (Example 3.1), so $\mathbb{S}_{\rho,\rho_0}(\lambda) = s_\lambda(\rho)s_\lambda(\rho_0)$. Since $s_\varnothing = 1$, we have

$$\mathbb{S}_{\rho,\rho_0}(\lambda) = \begin{cases} 1, & \text{if } \lambda = \varnothing; \\ 0, & \text{otherwise,} \end{cases}$$

therefore, a random partition distributed with respect to $\mathbb{S}_{\rho,\rho_0}$ is equal to $\varnothing$ almost surely.

Example 3.4. Let $t \geq 0$. For the specialization $\rho_t = (\mathbf{0};\mathbf{0};t)$, we claim that the Schur measure $\mathbb{S}_{\rho_t,\rho_t}$ is exactly the Poissonized Plancherel measure with parameter t^2 (see Definition 2.12). First, we have $p_1(\rho_t) = t$ and $p_k(\rho_t) = 0$ for $k > 1$, by Theorem 2.8, so $\prod(\rho_t,\rho_t) = e^{t^2}$ by (3.1). For any $\lambda \in \mathbb{Y}$, we have by Proposition 2.28,

$$\mathbb{S}_{\rho_t,\rho_t}(\lambda) = \frac{s_\lambda(\rho_t)s_\lambda(\rho_t)}{\prod(\rho_t,\rho_t)} = e^{-t^2}\left(\frac{t^{|\lambda|}\dim(\lambda)}{|\lambda|!}\right)^2 = e^{-t^2}\left(t^2\right)^{|\lambda|}\frac{\dim(\lambda)^2}{(|\lambda|!)^2} = \mathbb{P}_{t^2}^{\text{Poisson}}(\lambda).$$

The Schur measure is a way to compute probabilities on $\mathbb{Y}$. In other words, it is an infinite die with partitions on its faces. Now we want to define an evolution of random partitions in time. We do it using Markov chains.

3.2.2. Markov Chains on Partitions

We start with a quick review on Markov chains. A **stochastic process** is a set of random variables $(X_t)_{t\geq 0}$ or $(X_t)_{t=0}^{\infty}$. We should think of t as time, and (X_t) as describing the time evolution of a certain quantity. Typically, the value X_t depends on previous values $X_{\bar{t}}$ $(\bar{t} < t)$, in some structural way. A case of particular interest are Markov chains.

Let S be a discrete set, as $\mathbb{N}$, $\{1,2,\ldots,N\}$ or $\mathbb{Y}$. A **Markov Chain** is a discrete-time stochastic process $(X_t)_{t=0}^{\infty}$, where the random variable X_t takes values on S (we call S the **state space**) and they have the following property: the state of X_{t+1} depends only on the state of X_t. More precisely,

$$\mathbb{P}(X_{t+1} = i_{t+1} \mid X_t = i_t, X_{t-1} = i_{t-1}, \ldots, X_0 = i_0) = \mathbb{P}(X_{t+1} = i_{t+1} \mid X_t = i_t).$$

Consider an enumeration $S = \{s_1, s_2, s_3, \ldots\}$. For each pair i and j, we define the number $p_{ij} := \mathbb{P}\left(X_{t+1} = s_j \mid X_t = s_i\right)$, called **transitional probability**, which is the probability for going from the state s_i to the state s_j. We join these probabilities in the **transition matrix**

$$P := \begin{pmatrix} p_{00} & p_{01} & p_{02} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots \\ p_{20} & p_{21} & p_{22} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

where the sum along each row must be 1 and the matrices may have infinite size. Such matrices are also called **stochastic matrices** and they are sufficient to define a Markov Chain. It is possible to verify that the transition matrix for the two-step evolution of the chain is P^2. For n steps, it is P^n. Also, we always need to define the initial distribution for X_0. Now we define a Markov chain in the state space $S = \mathbb{Y}$. Recall that specializations were defined in 2.25.

Definition 3.2. (Markov chains on $\mathbb{Y}$) Let ρ, ρ' be specializations such that $\prod(\rho, \rho') < \infty$. We define two Markov chains on $\mathbb{Y}$ by the two-dimensional transition matrices

$$p^{\uparrow}(\rho,\rho') = \left(p^{\uparrow}_{\lambda\to\mu}(\rho,\rho')\right)_{\lambda,\mu\in\mathbb{Y}} \qquad \text{and} \qquad p^{\downarrow}(\rho,\rho') = \left(p^{\downarrow}_{\lambda\to\mu}(\rho,\rho')\right)_{\lambda,\mu\in\mathbb{Y}},$$

indexed by partitions (you can take any enumeration of $\mathbb{Y}$), where

$$p^{\uparrow}_{\lambda\to\mu}(\rho,\rho') := \frac{s_\mu(\rho)}{\prod(\rho,\rho')s_\lambda(\rho)} s_{\mu/\lambda}(\rho') \quad \text{and} \quad p^{\downarrow}_{\lambda\to\mu}(\rho,\rho') := \frac{s_\mu(\rho)}{s_\lambda(\rho\cup\rho')} s_{\lambda/\mu}(\rho'). \quad (3.3)$$

Proposition 3.1. The matrices $p^{\uparrow}(\rho,\rho')$ and $p^{\downarrow}(\rho,\rho')$ are stochastic.

Proof. By definition (3.3), we have for every $\lambda \in \mathbb{Y}$ (each line of the matrix),

$$\sum_{\mu\in\mathbb{Y}} p^{\uparrow}_{\lambda\to\mu}(\rho,\rho') = \frac{1}{\prod(\rho,\rho')s_\lambda(\rho)} \sum_{\mu\in\mathbb{Y}} s_\mu(\rho) s_{\mu/\lambda}(\rho') = 1,$$

by Lemma 2.4. For the second matrix,

$$\sum_{\mu\in\mathbb{Y}} p^{\downarrow}_{\lambda\to\mu}(\rho,\rho') = \frac{1}{s_\lambda(\rho\cup\rho')}\sum_{\mu\in\mathbb{Y}} s_\mu(\rho)s_{\lambda/\mu}(\rho') = \frac{s_\lambda(\rho,\rho')}{s_\lambda(\rho\cup\rho')} = 1,$$

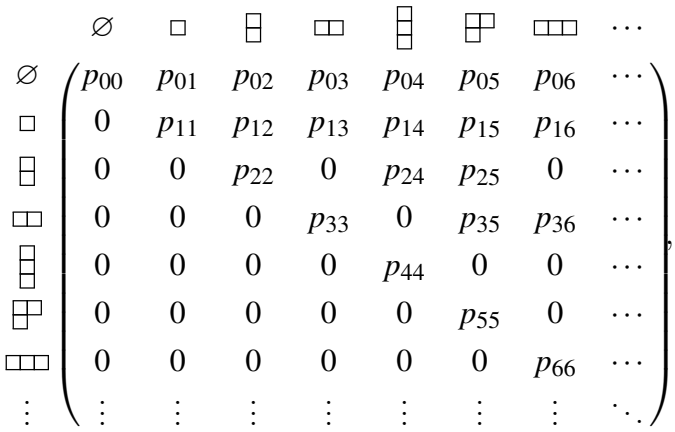
by Lemmas 2.5 and 2.6. □

Using the transition matrix $p^{\uparrow}(\rho,\rho')$ (or $p^{\downarrow}(\rho,\rho')$) we can define a Markov chain on $\mathbb{Y}$. For example, provided a initial partition $\lambda(0)=\mu$, the probability for $\lambda(1)=\nu$ is equal to $p^{\uparrow}_{\mu\to\nu}(\rho,\rho')$. Repeating this process, we obtain a partition $\lambda(t)$ (depending on $\lambda(t-1)$) for every $t=0,1,2,\ldots$.

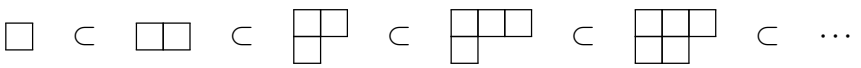
Recall from Proposition 2.23 that

$$\nu\not\subset\kappa \quad\Rightarrow\quad s_{\kappa/\nu}=0 \quad\Rightarrow\quad p^{\uparrow}_{\nu\to\kappa}(\rho,\rho')=0, \quad p^{\downarrow}_{\kappa\to\nu}(\rho,\rho')=0.$$

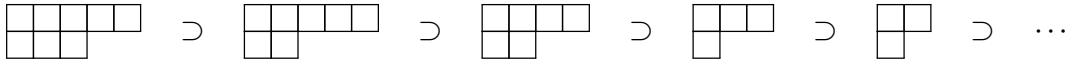
So the chain defined by $p^{\uparrow}(\rho,\rho')$ satisfies $\lambda(0)\subset\lambda(1)\subset\lambda(2)\subset\cdots$ and the one defined by $p^{\downarrow}(\rho,\rho')$ satisfies $\lambda(0)\supset\lambda(1)\supset\lambda(2)\supset\cdots$. To illustrate, the transition matrix $p^{\uparrow}(\rho,\rho')$ (indexed by Young diagrams) has the form



where each entry is a transitional probability. In terms of Young diagrams, the Markov chains defined above give us random sequences of diagrams as



for $p^{\uparrow}(\rho,\rho')$ (increasing the diagram), or



for $p^{\downarrow}(\rho,\rho')$ (decreasing the diagram). These Markov chains preserve Schur measures in the sense of the following proposition.

Proposition 3.2 (Schur measure preservation property)**.** Let $\mu\in\mathbb{Y}$. Consider the specializations ρ_1,ρ_2 and ρ_3, then

$$\sum_{\lambda\in\mathbb{Y}} \mathbb{S}_{\rho_1,\rho_2}(\lambda)p^{\uparrow}_{\lambda\to\mu}(\rho_2,\rho_3) = \mathbb{S}_{\rho_1\cup\rho_3,\rho_2}(\mu)$$

and

$$\sum_{\lambda\in\mathbb{Y}} \mathbb{S}_{\rho_1,\rho_2\cup\rho_3}(\lambda)p^{\downarrow}_{\lambda\to\mu}(\rho_2,\rho_3) = \mathbb{S}_{\rho_1,\rho_2}(\mu).$$

Proof. Applying Definitions 3.1 and 3.2, we have by Lemma 2.5,

$$\begin{aligned}\sum_{\lambda\in\mathbb{Y}} \mathbb{S}_{\rho_1,\rho_2}(\lambda)p^{\uparrow}_{\lambda\to\mu}(\rho_2,\rho_3) &= \sum_{\lambda\in\mathbb{Y}} \frac{s_\lambda(\rho_1)\cancel{s_\lambda(\rho_2)}}{\prod(\rho_1,\rho_2)}\frac{1}{\prod(\rho_2,\rho_3)}\frac{s_\mu(\rho_2)}{\cancel{s_\lambda(\rho_2)}}s_{\mu/\lambda}(\rho_3)\\ &= \frac{s_\mu(\rho_2)}{\prod(\rho_1,\rho_2)\prod(\rho_2,\rho_3)}\sum_{\lambda\in\mathbb{Y}} s_\lambda(\rho_1)s_{\mu/\lambda}(\rho_3)\\ &= \frac{s_\mu(\rho_2)s_\mu(\rho_1\cup\rho_3)}{\prod(\rho_2,\rho_1\cup\rho_3)} = \mathbb{S}_{\rho_1\cup\rho_3,\rho_2}(\mu).\end{aligned}$$

The second identity is similar: apply the definitions and one of Cauchy identities. □

The proposition above says that, if the distribution of the partition $\lambda(t)$ is a Schur measure, then the distribution of $\lambda(t+1)$ is still a Schur measure, with slightly different parameters. The sums in the proposition are from the matrix product between the starting distribution and the transition matrix (for more steps, we multiply by the transition matrix again). So if we start with $\lambda(0)$ distributed with respect to a Schur measure, we can apply Proposition 3.2 recursively, to obtain that $\lambda(t)$ is distributed by a Schur measure at any time t (in the language of Markov chains, we multiply the initial distribution by the transition matrix t times). We end the section with an useful technical lemma.

Lemma 3.1 (Commutation relation)**.** Let ρ_1,ρ_2 and ρ_3 be specializations. We have the matrix identity

$$p^{\uparrow}(\rho_1\cup\rho_2,\rho_3)p^{\downarrow}(\rho_1,\rho_2) = p^{\downarrow}(\rho_1,\rho_2)p^{\uparrow}(\rho_1,\rho_3).$$

Proof. Let $A = p^{\uparrow}(\rho_1\cup\rho_2,\rho_3)p^{\downarrow}(\rho_1,\rho_2)$ and $B = p^{\downarrow}(\rho_1,\rho_2)p^{\uparrow}(\rho_1,\rho_3)$, it suffices to prove that the entries $A_{\lambda\mu}$ and $B_{\lambda\mu}$ are equal for every pair $\lambda,\mu\in\mathbb{Y}$. By definition of product

of matrices and by skew Cauchy identity (Theorem 2.7),

$$\begin{aligned}
A_{\lambda\mu} &= \sum_{\nu\in\mathbb{Y}} p^{\uparrow}_{\lambda\to\nu}(\rho_1\cup\rho_2,\rho_3)\,p^{\downarrow}_{\nu\to\mu}(\rho_1,\rho_2)\\
&= \sum_{\nu\in\mathbb{Y}} \frac{1}{\prod(\rho_1\cup\rho_2,\rho_3)}\frac{s_\nu(\rho_1\cup\rho_2)}{s_\lambda(\rho_1\cup\rho_2)}s_{\nu/\lambda}(\rho_3)\frac{s_\mu(\rho_1)}{s_\nu(\rho_1\cup\rho_2)}s_{\nu/\mu}(\rho_2)\\
&= \frac{1}{\prod(\rho_1,\rho_3)\prod(\rho_2,\rho_3)}\frac{s_\mu(\rho_1)}{s_\lambda(\rho_1\cup\rho_2)}\sum_{\nu\in\mathbb{Y}} s_{\nu/\lambda}(\rho_3)s_{\nu/\mu}(\rho_2)\\
&= \frac{1}{\prod(\rho_1,\rho_3)}\frac{s_\mu(\rho_1)}{s_\lambda(\rho_1\cup\rho_2)}\sum_{\nu\in\mathbb{Y}} s_{\lambda/\nu}(\rho_2)s_{\mu/\nu}(\rho_3)\\
&= \sum_{\nu\in\mathbb{Y}} \frac{s_\nu(\rho_1)}{s_\lambda(\rho_1\cup\rho_2)}s_{\lambda/\nu}(\rho_2)\frac{1}{\prod(\rho_1,\rho_3)}\frac{s_\mu(\rho_1)}{s_\nu(\rho_1)}s_{\mu/\nu}(\rho_3)\\
&= \sum_{\nu\in\mathbb{Y}} p^{\downarrow}_{\lambda\to\nu}(\rho_1,\rho_2)\,p^{\uparrow}_{\nu\to\mu}(\rho_1,\rho_3) = B_{\lambda\mu},
\end{aligned}$$

as we wanted to show. □

At this point, the reader may have noticed that all the proofs are straightforward applications of the definitions, Theorems 2.5, 2.6 and 2.7 or Lemmas 2.4 and 2.5. The next section will not be different.

3.3. Probabilities on Sequences of Partitions

3.3.1. The Schur Process

We now generalize the study from previous section to sequences of partitions. We denote by

$$\mathbb{Y}^N = \mathbb{Y}\times\cdots\times\mathbb{Y} = \left\{ \left(\lambda^{(1)},\ldots,\lambda^{(N)}\right) \middle| \lambda^{(i)}\in\mathbb{Y}, i=1,\ldots,N \right\}$$

the set of all sequences of N partitions. We use the superscript notation

$$\lambda^{(i)} = \left(\lambda^{(i)}_1,\lambda^{(i)}_2,\ldots\right)$$

so that we have no confusion with the terms of partitions. For example, the j-th term of the i-th partition is given by $\lambda^{(i)}_j$. You can also think of the elements of $\mathbb{Y}^N$ as vectors consisting of N Young diagrams, such as

$$\left(\ydiagram{1}\right)\in\mathbb{Y}^2,\quad \left(\ydiagram{2}\right)\in\mathbb{Y}^3 \quad\text{or}\quad \left(\ydiagram{3}\right)\in\mathbb{Y}^4.$$

We first define a probability measure on $\mathbb{Y}^N$, generalizing the Schur measure (Definition 3.1).

Definition 3.3. Let $\rho_1,\dots,\rho_N$ and ρ^- be specializations. The probability measure on $\mathbb{Y}^N$ given by

$$\mathbb{S}_{\rho_1,\dots,\rho_N,\rho^-}\left(\lambda^{(1)},\dots,\lambda^{(N)}\right) := \frac{s_{\lambda^{(1)}}(\rho_1)s_{\lambda^{(2)}/\lambda^{(1)}}(\rho_2)\cdots s_{\lambda^{(N)}/\lambda^{(N-1)}}(\rho_N)s_{\lambda^{(N)}}(\rho^-)}{\prod(\rho_1,\rho^-)\cdots\prod(\rho_N,\rho^-)}$$

is called the **Schur process of rank** N parametrized by $\rho_1,\dots,\rho_N$ and ρ^-.

We have a few notes about Definition 3.3.

Note 3.2. Recall from Proposition 2.23 that if $\mu \not\subset \lambda$, then $s_{\lambda/\mu}=0$, so $\mathbb{S}_{\rho_1,\dots,\rho_N,\rho^-}$ is supported (i.e. it is non-zero) on sequences satisfying $\lambda^{(1)}\subset\cdots\subset\lambda^{(N)}$.

Note 3.3. When $N=1$, Definition 3.3 is exactly the Schur measure $\mathbb{S}_{\rho_1,\rho^-}$ from Definition 3.1.

Note 3.4. The Schur process $\mathbb{S}_{\rho_1,\dots,\rho_N,\rho^-}$ is indeed a probability measure. We can prove it by induction. By Note 3.3, the Schur process of rank 1 is a probability measure (base case). If we assume that any Schur process of rank $N-1$ is a probability measure (induction step), then we have by Lemma 2.4,

$$\begin{aligned}
&\sum_{\lambda^{(1)},\dots,\lambda^{(N)}} \mathbb{S}_{\rho_1,\dots,\rho_N,\rho^-}\left(\lambda^{(1)},\dots,\lambda^{(N)}\right)\\
&= \sum_{\lambda^{(1)},\dots,\lambda^{(N-1)}} \frac{s_{\lambda^{(1)}}(\rho_1)s_{\lambda^{(2)}/\lambda^{(1)}}(\rho_2)\cdots s_{\lambda^{(N-1)}/\lambda^{(N-2)}}(\rho_{N-1})}{\prod(\rho_1,\rho^-)\cdots\prod(\rho_{N-1},\rho^-)} \sum_{\lambda^{(N)}} \frac{s_{\lambda^{(N)}/\lambda^{(N-1)}}(\rho_N)s_{\lambda^{(N)}}(\rho^-)}{\prod(\rho_N,\rho^-)}\\
&= \sum_{\lambda^{(1)},\dots,\lambda^{(N-1)}} \frac{s_{\lambda^{(1)}}(\rho_1)s_{\lambda^{(2)}/\lambda^{(1)}}(\rho_2)\cdots s_{\lambda^{(N-1)}/\lambda^{(N-2)}}(\rho_{N-1})}{\prod(\rho_1,\rho^-)\cdots\prod(\rho_{N-1},\rho^-)} s_{\lambda^{(N-1)}}(\rho^-)\\
&= \sum_{\lambda^{(1)},\dots,\lambda^{(N-1)}} \mathbb{S}_{\rho_1,\dots,\rho_{N-1},\rho^-}\left(\lambda^{(1)},\dots,\lambda^{(N-1)}\right) = 1.
\end{aligned}$$

So the Schur process of rank N is a probability measure and the induction argument holds.

Example 3.5. Let $\rho^- = (\mathbf{0};\mathbf{0};0)$. For any specializations $\rho_1,\dots,\rho_N$, we have

$$\mathbb{S}_{\rho_1,\dots,\rho_N,\rho^-}(\varnothing,\dots,\varnothing)=1,$$

since $s_{\varnothing/\varnothing}=s_\varnothing=1$ and $\prod(\rho_i,\rho^-)=1$ for each $i=1,\dots,N$ (see Example 3.1). In other words, if $\left(\lambda^{(1)},\dots,\lambda^{(N)}\right)$ is distributed with respect to $\mathbb{S}_{\rho_1,\dots,\rho_N,\rho^-}$, where ρ^- is the trivial specialization, then $\left(\lambda^{(1)},\dots,\lambda^{(N)}\right)=(\varnothing,\dots,\varnothing)$ almost surely.

When we study probability in higher dimensions, we may want to obtain the **marginal probabilities**, that is, a probability with respect to just one of the variables. To illustrate, for a probability mass function $\mathbb{P}(x,y,z)$ with x,y,z in some discrete set, the marginal probability densities in x, y and z are respectively

$$\mathbb{P}_X(x) := \sum_{y,z}\mathbb{P}(x,y,z),\quad \mathbb{P}_Y(y) := \sum_{x,z}\mathbb{P}(x,y,z)\quad\text{and}\quad \mathbb{P}_Z(z) := \sum_{x,y}\mathbb{P}(x,y,z).$$

Proposition 3.3. The marginals of Schur processes are Schur measures. More precisely, for each $k = 1, \dots, N$, we have

$$\sum_{\lambda^{(1)},\dots,\lambda^{(k-1)},\lambda^{(k+1)},\dots,\lambda^{(N)}} \mathbb{S}_{\rho_1,\dots,\rho_N,\rho^-}\left(\lambda^{(1)},\dots,\lambda^{(k)}\dots,\lambda^{(N)}\right) = \mathbb{S}_{\rho,\rho^-}\left(\lambda^{(k)}\right)$$

where $\rho = \rho_1 \cup \dots \cup \rho_k$.

Proof. We proceed by induction over the rank of Schur processes. For rank $N = 1$, there is nothing to prove, since the Schur process is a Schur measure. Suppose that the result is true for any Schur process of rank $N-1$. If $1 \le k < N$, we have by a computation similar to Remark 3.4,

$$\begin{aligned}
&\sum_{\lambda^{(1)},\dots,\lambda^{(k-1)},\lambda^{(k+1)},\dots,\lambda^{(N)}} \mathbb{S}_{\rho_1,\dots,\rho_N,\rho^-}\left(\lambda^{(1)},\dots,\lambda^{(k)},\dots,\lambda^{(N-1)}\right) \\
&\quad = \sum_{\lambda^{(1)},\dots,\lambda^{(k-1)},\lambda^{(k+1)},\dots,\lambda^{(N-1)}} \mathbb{S}_{\rho_1,\dots,\rho_{N-1},\rho^-}\left(\lambda^{(1)},\dots,\lambda^{(k)},\dots,\lambda^{(N-1)}\right) = \mathbb{S}_{\rho,\rho^-}\left(\lambda^{(k)}\right),
\end{aligned}$$

with $\rho = \rho_1 \cup \dots \cup \rho_k$, using the induction hypothesis in the last equality. For $k = N$, we use the skew summation formula (Theorem 2.6) to obtain

$$\begin{aligned}
&\sum_{\lambda^{(1)},\dots,\lambda^{(N-1)}} \mathbb{S}_{\rho_1,\dots,\rho_N,\rho^-}\left(\lambda^{(1)},\dots,\lambda^{(N)}\right) \\
&= \sum_{\lambda^{(1)},\dots,\lambda^{(N-1)}} \frac{s_{\lambda^{(1)}}(\rho_1) s_{\frac{\lambda^{(2)}}{\lambda^{(1)}}}(\rho_2)\cdots s_{\frac{\lambda^{(N-2)}}{\lambda^{(N-3)}}}(\rho_{N-2}) s_{\frac{\lambda^{(N-1)}}{\lambda^{(N-2)}}}(\rho_{N-1}) s_{\frac{\lambda^{(N)}}{\lambda^{(N-1)}}}(\rho_N) s_{\lambda^{(N)}}(\rho^-)}{\prod(\rho_1,\rho^-)\cdots\prod(\rho_N,\rho^-)} \\
&= \sum_{\lambda^{(1)},\dots,\lambda^{(N-2)}} \frac{s_{\lambda^{(1)}}(\rho_1) s_{\frac{\lambda^{(2)}}{\lambda^{(1)}}}(\rho_2)\cdots s_{\frac{\lambda^{(N-2)}}{\lambda^{(N-3)}}}(\rho_{N-2}) s_{\lambda^{(N)}}(\rho^-)}{\prod(\rho_1,\rho^-)\cdots\prod(\rho_N,\rho^-)} \sum_{\lambda^{(N-1)}} s_{\frac{\lambda^{(N)}}{\lambda^{(N-1)}}}(\rho_N) s_{\frac{\lambda^{(N-1)}}{\lambda^{(N-2)}}}(\rho_{N-1}) \\
&= \sum_{\lambda^{(1)},\dots,\lambda^{(N-2)}} \frac{s_{\lambda^{(1)}}(\rho_1) s_{\frac{\lambda^{(2)}}{\lambda^{(1)}}}(\rho_2)\cdots s_{\frac{\lambda^{(N-2)}}{\lambda^{(N-3)}}}(\rho_{N-2}) s_{\lambda^{(N)}}(\rho^-)}{\prod(\rho_1,\rho^-)\cdots\prod(\rho_N,\rho^-)} s_{\lambda^{(N)}/\lambda^{(N-2)}}(\rho_{N-1}\cup\rho_N) \\
&= \sum_{\lambda^{(1)},\dots,\lambda^{(N-2)}} \mathbb{S}_{\rho_1,\dots,\rho_{N-2},\rho_{N-1}\cup\rho_N,\rho^-}\left(\lambda^{(1)},\dots,\lambda^{(N-2)},\lambda^{(N)}\right) = \mathbb{S}_{\rho,\rho^-}\left(\lambda^{(N)}\right)
\end{aligned}$$

with $\rho = \rho_1 \cup \dots \cup \rho_{N-2} \cup (\rho_{N-1} \cup \rho_N)$, using the induction hypothesis in the last equality. □

3.3.2. Markov Chains on Sequences of Partitions

We want to define a Markov chain on $\mathbb{Y}^N$ that preserves Schur processes, in analogy to Proposition 3.2. Consider the specializations $\rho_1, \dots, \rho_N, \rho'$. We define the transitional probabilities on $\mathbb{Y}^N$ by

$$\mathbb{P}\left(\begin{pmatrix}\nu^{(1)}\\ \vdots \\ \nu^{(N)}\end{pmatrix} \to \begin{pmatrix}\mu^{(1)}\\ \vdots \\ \mu^{(N)}\end{pmatrix}\right) := \prod_{\ell=1}^{N} \mathbb{P}_\ell\left(\nu^{(\ell)}, \mu^{(\ell-1)} \to \mu^{(\ell)}\right) \tag{3.4}$$

where

$$\mathbb{P}_\ell(\nu,\kappa\to\mu) := \frac{p^{\uparrow}_{\nu\to\mu}(\rho_1\cup\cdots\cup\rho_\ell,\rho')p^{\downarrow}_{\mu\to\kappa}(\rho_1\cup\cdots\cup\rho_{\ell-1},\rho_\ell)}{\sum_\lambda p^{\uparrow}_{\nu\to\lambda}(\rho_1\cup\cdots\cup\rho_\ell,\rho')p^{\downarrow}_{\lambda\to\kappa}(\rho_1\cup\cdots\cup\rho_{\ell-1},\rho_\ell)} \tag{3.5}$$

and we assume $\mu^{(0)}=\varnothing$ so that

$$\mathbb{P}_1\left(\nu^{(1)},\mu^{(0)}\to\mu^{(1)}\right)=p^{\uparrow}_{\nu^{(1)}\to\mu^{(1)}}(\rho_1,\rho').$$

If $N=1$, this reduces to the Markov chain on $\mathbb{Y}$ given by the matrix $p^{\uparrow}(\rho_1,\rho')$ (Definition 3.2). Notice that each $\mathbb{P}_\ell$ is a transitional probability from $\mathbb{Y}\times\mathbb{Y}$ to $\mathbb{Y}$, that is, for any ν and κ, we have

$$\sum_{\mu\in\mathbb{Y}}\mathbb{P}_\ell(\nu,\kappa\to\mu)=1. \tag{3.6}$$

The following proposition shows that the first k partitions (of the N partitions) depend just on the first k partitions on past time.

Proposition 3.4. For any $k=1,2,\ldots,N$,

$$\mathbb{P}\left(\mu^{(1)},\ldots,\mu^{(k)}\middle|\,\nu^{(1)},\ldots,\nu^{(N)}\right)=\mathbb{P}\left(\mu^{(1)},\ldots,\mu^{(k)}\middle|\,\nu^{(1)},\ldots,\nu^{(k)}\right).$$

Proof. We have

$$\begin{aligned}
&\mathbb{P}\left(\mu^{(1)},\ldots,\mu^{(k)}\middle|\,\nu^{(1)},\ldots,\nu^{(N)}\right)\\
&=\sum_{\mu^{(k+1)},\ldots,\mu^{(N)}}\mathbb{P}\left(\begin{pmatrix}\nu^{(1)}\\ \vdots\\ \nu^{(N)}\end{pmatrix}\to\begin{pmatrix}\mu^{(1)}\\ \vdots\\ \mu^{(N)}\end{pmatrix}\right)\\
&=\sum_{\mu^{(k+1)},\ldots,\mu^{(N)}}\prod_{\ell=1}^{N}\mathbb{P}_\ell\left(\nu^{(\ell)},\mu^{(\ell-1)}\to\mu^{(\ell)}\right)\\
&=\prod_{\ell=1}^{k}\mathbb{P}_\ell\left(\nu^{(\ell)},\mu^{(\ell-1)}\to\mu^{(\ell)}\right)\sum_{\mu^{(k+1)},\ldots,\mu^{(N)}}\prod_{\ell=k+1}^{N}\mathbb{P}_\ell\left(\nu^{(\ell)},\mu^{(\ell-1)}\to\mu^{(\ell)}\right)\\
&=\prod_{\ell=1}^{k}\mathbb{P}_\ell\left(\nu^{(\ell)},\mu^{(\ell-1)}\to\mu^{(\ell)}\right)\prod_{\ell=k+1}^{N}\underbrace{\sum_{\mu^{(\ell)}}\mathbb{P}_\ell\left(\nu^{(\ell)},\mu^{(\ell-1)}\to\mu^{(\ell)}\right)}_{=1,\text{ by Eq. (3.6)}}\\
&=\prod_{\ell=1}^{k}\mathbb{P}_\ell\left(\nu^{(\ell)},\mu^{(\ell-1)}\to\mu^{(\ell)}\right)\\
&=\mathbb{P}\left(\mu^{(1)},\ldots,\mu^{(k)}\middle|\,\nu^{(1)},\ldots,\nu^{(k)}\right)
\end{aligned}$$

□

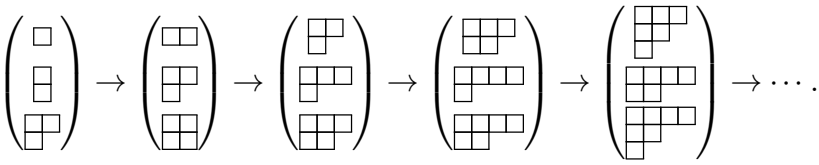
Corollary 3.1. The transitional probability defined in (3.4) defines a Markov chain $\left(\lambda^{(1)}(t),\ldots,\lambda^{(N)}(t)\right)$ on $\mathbb{Y}^N$.

Proof. We need to show that

$$\sum_{\mu^{(1)},\ldots,\mu^{(N)}} \mathbb{P}\left(\begin{pmatrix}\nu^{(1)}\\ \vdots \\ \nu^{(N)}\end{pmatrix} \to \begin{pmatrix}\mu^{(1)}\\ \vdots \\ \mu^{(N)}\end{pmatrix}\right) = 1,$$

that is, the transition matrix is stochastic (now the matrix is $\mathbb{Y}^N \times \mathbb{Y}^N$). To see this, we repeat the proof of Proposition 3.4, summing over all sequences $(\mu^{(1)},\ldots,\mu^{(N)})$. □

To keep in mind the objects we are working with, the Markov chain defined above is simply a sequence of random vectors consisting of Young diagrams evolving in time, as



Using the time notation for the Markov chain, we consider the update in one unit of time

$$\begin{pmatrix}\lambda^{(1)}(t)\\ \vdots \\ \lambda^{(N)}(t)\end{pmatrix} = \begin{pmatrix}\nu^{(1)}\\ \vdots \\ \nu^{(N)}\end{pmatrix} \to \begin{pmatrix}\mu^{(1)}\\ \vdots \\ \mu^{(N)}\end{pmatrix} = \begin{pmatrix}\lambda^{(1)}(t+1)\\ \vdots \\ \lambda^{(N)}(t+1)\end{pmatrix}.$$

By the following property for condition probability, given by

$$\mathbb{P}(A|BC) = \frac{\mathbb{P}(AB|C)}{\mathbb{P}(B|C)},$$

we have for any $k = 1,2,\ldots,N$,

$$\begin{aligned}\mathbb{P}\left(\mu^{(k)}\middle|\,\nu^{(1)},\ldots,\nu^{(k)},\mu^{(1)},\ldots,\mu^{(k-1)}\right) &= \frac{\mathbb{P}\left(\mu^{(1)},\ldots,\mu^{(k)}\middle|\,\nu^{(1)},\ldots,\nu^{(k)}\right)}{\mathbb{P}\left(\mu^{(1)},\ldots,\mu^{(k-1)}\middle|\,\nu^{(1)},\ldots,\nu^{(k)}\right)}\\ &= \frac{\mathbb{P}\left(\mu^{(1)},\ldots,\mu^{(k)}\middle|\,\nu^{(1)},\ldots,\nu^{(k)}\right)}{\mathbb{P}\left(\mu^{(1)},\ldots,\mu^{(k-1)}\middle|\,\nu^{(1)},\ldots,\nu^{(k-1)}\right)}\\ &= \mathbb{P}_k(\nu^{(k)},\mu^{(k-1)} \to \mu^{(k)}),\end{aligned}$$

where we used Proposition 3.4 and Equation (3.4). The calculation above means that the update of the chain from time t to time $t+1$ follows the N sequential steps below:

1. First, we update $\nu^{(1)} \to \mu^{(1)}$ with probability $\mathbb{P}_1\left(\nu^{(1)},\varnothing \to \mu^{(1)}\right)$;

2. Then given $\mu^{(1)}$, we update $\nu^{(2)} \to \mu^{(2)}$ with probability $\mathbb{P}_2\left(\nu^{(2)}, \mu^{(1)} \to \mu^{(2)}\right)$;

3. Given $\mu^{(2)}$, we update $\nu^{(3)} \to \mu^{(3)}$ with probability $\mathbb{P}_3\left(\nu^{(3)}, \mu^{(2)} \to \mu^{(3)}\right)$;

$\vdots$

N. Finally, given $\mu^{(N-1)}$, we update $\nu^{(N)} \to \mu^{(N)}$ with probability $\mathbb{P}_N\left(\nu^{(N)}, \mu^{(N-1)} \to \mu^{(N)}\right)$.

Lemma 3.2. The probability measure for the Schur process parametrized by $\rho_1, \dots, \rho_N$ and ρ^- can be written as

$$\mathbb{S}_{\rho_1,\dots,\rho_N,\rho^-}\left(\lambda^{(1)},\dots,\lambda^{(N)}\right) = \mathbb{S}_{\rho_1\cup\cdots\cup\rho_N,\rho^-}\left(\lambda^{(N)}\right)\prod_{\ell=2}^{N} p^{\downarrow}_{\lambda^{(\ell)}\to\lambda^{(\ell-1)}}\left(\rho_1\cup\cdots\cup\rho_{\ell-1},\rho_\ell\right).$$

Proof. Applying Definitions 3.1 and 3.2, the right-hand side is equal to

$$\begin{aligned}&\frac{s_{\lambda^{(N)}}(\rho_1\cup\cdots\cup\rho_N)s_{\lambda^{(N)}}(\rho^-)}{\prod(\rho_1\cup\cdots\cup\rho_N,\rho^-)}\prod_{\ell=2}^{N}\frac{s_{\lambda^{(\ell-1)}}(\rho_1\cup\cdots\cup\rho_{\ell-1})}{s_{\lambda^{(\ell)}}(\rho_1\cup\cdots\cup\rho_\ell)}s_{\lambda^{(\ell)}/\lambda^{(\ell-1)}}(\rho_\ell)\\ &=\frac{s_{\lambda^{(1)}}(\rho_1)s_{\lambda^{(2)}/\lambda^{(1)}}(\rho_2)\cdots s_{\lambda^{(N)}/\lambda^{(N-1)}}(\rho_N)s_{\lambda^{(N)}}(\rho^-)}{\prod(\rho_1,\rho^-)\cdots\prod(\rho_N,\rho^-)}\end{aligned}$$

after the cancellations. The last expression is exactly the Schur process $\mathbb{S}_{\rho_1,\dots,\rho_N,\rho^-}$ (Definition 3.3). □

Now we proceed to the main result of this section, which states that the Markov chain defined in (3.4) preserves Schur processes, in analogy to the one-dimensional case (Proposition 3.2).

Theorem 3.1. Consider the Markov chain on $\mathbb{Y}^N$ defined by (3.4), parametrized by the specializations $\rho_1, \dots, \rho_N$ and ρ'. It maps the Schur process $\mathbb{S}_{\rho_1,\dots,\rho_N,\rho^-}$ to the Schur process $\mathbb{S}_{\rho_1,\dots,\rho_N,\rho^-\cup\rho'}$, for any specialization ρ^-. More precisely, we have

$$\sum_{\lambda^{(1)},\dots,\lambda^{(N)}}\mathbb{S}_{\rho_1,\dots,\rho_N,\rho^-}\begin{pmatrix}\lambda^{(1)}\\ \vdots\\ \lambda^{(N)}\end{pmatrix}\mathbb{P}\left(\begin{pmatrix}\lambda^{(1)}\\ \vdots\\ \lambda^{(N)}\end{pmatrix}\to\begin{pmatrix}\mu^{(1)}\\ \vdots\\ \mu^{(N)}\end{pmatrix}\right)=\mathbb{S}_{\rho_1,\dots,\rho_N,\rho^-\cup\rho'}\begin{pmatrix}\mu^{(1)}\\ \vdots\\ \mu^{(N)}\end{pmatrix}. \quad (3.7)$$

The corollary below is immediate.

Corollary 3.2. Consider the Markov chain on $\mathbb{Y}^N$ defined by (3.4), parametrized by the specializations $\rho_1, \dots, \rho_N$ and ρ'. After t updates, it maps the Schur process $\mathbb{S}_{\rho_1,\dots,\rho_N,\rho^-}$ to the Schur process $\mathbb{S}_{\rho_1,\dots,\rho_N,\rho^-\cup\widehat{\rho}}$, for any specialization ρ^-, where $\widehat{\rho} = \rho'\cup\cdots\cup\rho'$ (the union is taken t times).

Proof. (of Theorem 3.1) Applying Lemma 3.2 and Definition 3.4, the left-hand side of (3.7) is equal to

$$\sum_{\lambda^{(1)},\dots,\lambda^{(N)}} \mathbb{S}_{\rho_1\cup\cdots\cup\rho_N,\rho^-}\left(\lambda^{(N)}\right) p^{\uparrow}_{\lambda^{(1)}\to\mu^{(1)}}(\rho_1,\rho')$$
$$\times\prod_{\ell=2}^{N} p^{\downarrow}_{\lambda^{(\ell)}\to\lambda^{(\ell-1)}}\left(\cup_{i=1}^{\ell-1}\rho_i,\rho_\ell\right)\frac{p^{\uparrow}_{\lambda^{(\ell)}\to\mu^{(\ell)}}(\rho_1\cup\cdots\cup\rho_\ell,\rho')p^{\downarrow}_{\mu^{(\ell)}\to\mu^{(\ell-1)}}(\rho_1\cup\cdots\cup\rho_{\ell-1},\rho_\ell)}{\sum_\nu p^{\uparrow}_{\lambda^{(\ell)}\to\nu}(\rho_1\cup\cdots\cup\rho_\ell,\rho')p^{\downarrow}_{\nu\to\mu^{(\ell-1)}}(\rho_1\cup\cdots\cup\rho_{\ell-1},\rho_\ell)}.$$

By the commutation relation (Lemma 3.1), we have the matrix identity

$$p^{\uparrow}(\rho_1\cup\cdots\cup\rho_\ell,\rho')p^{\downarrow}(\rho_1\cup\cdots\cup\rho_{\ell-1},\rho_\ell)=p^{\downarrow}(\rho_1\cup\cdots\cup\rho_{\ell-1},\rho_\ell)p^{\uparrow}(\rho_1\cup\cdots\cup\rho_{\ell-1},\rho'),$$

so we replace the sum on the denominator, obtaining

$$\sum_{\lambda^{(1)},\dots,\lambda^{(N)}} \mathbb{S}_{\rho_1\cup\cdots\cup\rho_N,\rho^-}\left(\lambda^{(N)}\right) p^{\uparrow}_{\lambda^{(1)}\to\mu^{(1)}}(\rho_1,\rho')$$
$$\times\prod_{\ell=2}^{N} p^{\downarrow}_{\lambda^{(\ell)}\to\lambda^{(\ell-1)}}\left(\cup_{i=1}^{\ell-1}\rho_i,\rho_\ell\right)\frac{p^{\uparrow}_{\lambda^{(\ell)}\to\mu^{(\ell)}}(\rho_1\cup\cdots\cup\rho_\ell,\rho')p^{\downarrow}_{\mu^{(\ell)}\to\mu^{(\ell-1)}}(\rho_1\cup\cdots\cup\rho_{\ell-1},\rho_\ell)}{\sum_\nu p^{\downarrow}_{\lambda^{(\ell)}\to\nu}(\rho_1\cup\cdots\cup\rho_{\ell-1},\rho_\ell)p^{\uparrow}_{\nu\to\mu^{(\ell-1)}}(\rho_1\cup\cdots\cup\rho_{\ell-1},\rho')}.$$

Ignoring all the terms without $\lambda^{(1)}$, we get

$$p^{\downarrow}_{\lambda^{(2)}\to\lambda^{(1)}}(\rho_1,\rho_2)p^{\uparrow}_{\lambda^{(1)}\to\mu^{(1)}}(\rho_1,\rho'),$$

so summing over $\lambda^{(1)}$, we cancel one of the denominators in the product. What remains is

$$\sum_{\lambda^{(2)},\dots,\lambda^{(N)}} \mathbb{S}_{\rho_1\cup\cdots\cup\rho_N,\rho^-}\left(\lambda^{(N)}\right) p^{\uparrow}_{\lambda^{(2)}\to\mu^{(2)}}(\rho_1\cup\rho_2,\rho')p^{\downarrow}_{\mu^{(2)}\to\mu^{(1)}}(\rho_1,\rho_2)$$
$$\times\prod_{\ell=3}^{N} p^{\downarrow}_{\lambda^{(\ell)}\to\lambda^{(\ell-1)}}\left(\cup_{i=1}^{\ell-1}\rho_i,\rho_\ell\right)\frac{p^{\uparrow}_{\lambda^{(\ell)}\to\mu^{(\ell)}}(\rho_1\cup\cdots\cup\rho_\ell,\rho')p^{\downarrow}_{\mu^{(\ell)}\to\mu^{(\ell-1)}}(\rho_1\cup\cdots\cup\rho_{\ell-1},\rho_\ell)}{\sum_\nu p^{\downarrow}_{\lambda^{(\ell)}\to\nu}(\rho_1\cup\cdots\cup\rho_{\ell-1},\rho_\ell)p^{\uparrow}_{\nu\to\mu^{(\ell-1)}}(\rho_1\cup\cdots\cup\rho_{\ell-1},\rho')}.$$

Ignoring all the terms without $\lambda^{(2)}$, we obtain

$$p^{\downarrow}_{\lambda^{(3)}\to\lambda^{(2)}}(\rho_1\cup\rho_2,\rho_3)p^{\uparrow}_{\lambda^{(2)}\to\mu^{(2)}}(\rho_1\cup\rho_2,\rho'),$$

so summing over $\lambda^{(2)}$, we again cancel one of the denominators in the product, obtaining

$$\sum_{\lambda^{(3)},\dots,\lambda^{(N)}} \mathbb{S}_{\rho_1\cup\cdots\cup\rho_N,\rho^-}\left(\lambda^{(N)}\right) p^{\uparrow}_{\lambda^{(3)}\to\mu^{(3)}}(\rho_1\cup\rho_2\cup\rho_3,\rho')p^{\downarrow}_{\mu^{(3)}\to\mu^{(2)}}(\rho_1\cup\rho_2,\rho_3)p^{\downarrow}_{\mu^{(2)}\to\mu^{(1)}}(\rho_1,\rho_2)$$
$$\times\prod_{\ell=4}^{N} p^{\downarrow}_{\lambda^{(\ell)}\to\lambda^{(\ell-1)}}\left(\cup_{i=1}^{\ell-1}\rho_i,\rho_\ell\right)\frac{p^{\uparrow}_{\lambda^{(\ell)}\to\mu^{(\ell)}}(\rho_1\cup\cdots\cup\rho_\ell,\rho')p^{\downarrow}_{\mu^{(\ell)}\to\mu^{(\ell-1)}}(\rho_1\cup\cdots\cup\rho_{\ell-1},\rho_\ell)}{\sum_\nu p^{\downarrow}_{\lambda^{(\ell)}\to\nu}(\rho_1\cup\cdots\cup\rho_{\ell-1},\rho_\ell)p^{\uparrow}_{\nu\to\mu^{(\ell-1)}}(\rho_1\cup\cdots\cup\rho_{\ell-1},\rho')}.$$

Repeating this process until the sum over $\lambda^{(N-1)}$ all the denominators are canceled and the remaining expression is

$$\sum_{\lambda^{(N)}} \mathbb{S}_{\rho_1\cup\cdots\cup\rho_N,\rho^-}\left(\lambda^{(N)}\right) p^{\uparrow}_{\lambda^{(N)}\to\mu^{(N)}}(\rho_1\cup\cdots\cup\rho_N,\rho')\prod_{\ell=2}^{N} p^{\downarrow}_{\mu^{(\ell)}\to\mu^{(\ell-1)}}(\rho_1\cup\cdots\cup\rho_{\ell-1},\rho_\ell).$$

By the Schur measure preservation property (Proposition 3.2), this is equal to

$$\mathbb{S}_{\rho_1\cup\cdots\cup\rho_N,\rho^-\cup\rho'}\left(\mu^{(N)}\right)\prod_{\ell=2}^{N} p^{\downarrow}_{\mu^{(\ell)}\to\mu^{(\ell-1)}}(\rho_1\cup\cdots\cup\rho_{\ell-1},\rho_\ell),$$

which is exactly $\mathbb{S}_{\rho_1,\ldots,\rho_N,\rho^-\cup\rho'}\left(\mu^{(1)},\ldots,\mu^{(N)}\right)$, by Lemma 3.2. □

This is all the theory we need for now. We show in the next section how to recover the TASEP model from this framework, justifying all the construction made.

3.4. Application: Block-Push Process and TASEP

3.4.1. Block-Push Markov Chain

In this section, we study a specific example of Markov chain defined in (3.4). Fix $N\in\mathbb{N}$, $b>0$ and consider the specializations

- $\rho_1=\cdots=\rho_N=((1,0,0,\ldots);\mathbf{0};0)$;
- $\rho'=(\mathbf{0};(b,0,0,\ldots);0)$;
- $\rho^-=(\mathbf{0};\mathbf{0};0)$.

By Proposition 2.26, we have for each $i=1,\ldots,N$,

$$s_{\lambda/\mu}(\rho_i)=\begin{cases}1, & \text{if } \lambda/\mu \text{ is a horizontal strip } (\mu\preccurlyeq\lambda);\\ 0, & \text{otherwise.}\end{cases}$$

Also, by Proposition 2.26,

$$s_{\lambda/\mu}(\rho')=\begin{cases}b^{|\lambda|-|\mu|}, & \text{if } \lambda/\mu \text{ is a vertical strip } (\mu\preccurlyeq_{\mathrm{v}}\lambda);\\ 0, & \text{otherwise.}\end{cases}$$

For the last, we have

$$s_{\lambda/\mu}(\rho^-)=\begin{cases}1, & \text{if } \lambda=\mu;\\ 0, & \text{otherwise,}\end{cases}$$

by Proposition 2.21(g). Consider the Markov chain $\left(\lambda^{(1)}(t),\dots,\lambda^{(N)}(t)\right)$ defined by the transition probability (3.4) parametrized by $\rho_1,\dots,\rho_N$ and ρ'. At time $t=0$, assume that the vector $\left(\lambda^{(1)}(0),\dots,\lambda^{(N)}(0)\right)$ is distributed by the Schur process parametrized by $\rho_1,\dots,\rho_N$ and ρ^-. By Example 3.5,

$$\left(\lambda^{(1)}(0),\dots,\lambda^{(N)}(0)\right)=(\varnothing,\dots,\varnothing)$$

almost surely.

Let us consider the transitional probabilities, given by

$$\mathbb{P}_\ell(\nu,\kappa\to\mu)=C_\ell(\nu,\kappa)p^{\uparrow}_{\nu\to\mu}(\rho_1\cup\cdots\cup\rho_\ell,\rho')p^{\downarrow}_{\mu\to\kappa}(\rho_1\cup\cdots\cup\rho_{\ell-1},\rho_\ell),$$

where the number $C_\ell(\nu,\kappa)$ does not depend on μ. So by definition,

$$\begin{aligned}\mathbb{P}_\ell(\nu,\kappa\to\mu)&=C_\ell(\nu,\kappa)\frac{s_\mu(\rho_1\cup\cdots\cup\rho_\ell)}{s_\nu(\rho_1\cup\cdots\cup\rho_\ell)}s_{\mu/\nu}(\rho')\frac{s_\kappa(\rho_1\cup\cdots\cup\rho_{\ell-1})}{s_\mu(\rho_1\cup\cdots\cup\rho_\ell)}s_{\mu/\kappa}(\rho_\ell)\\&=\widetilde{C}_\ell(\nu,\kappa)s_{\mu/\nu}(\rho')s_{\mu/\kappa}(\rho_k)\\&=\widetilde{C}_\ell(\nu,\kappa)b^{|\mu|-|\nu|}\mathbb{1}_{\nu\preccurlyeq_{\mathrm{v}}\mu}\mathbb{1}_{\kappa\preccurlyeq\mu},\end{aligned}$$

where the notations $\lambda\preccurlyeq_{\mathrm{v}}\mu$ and $\lambda\preccurlyeq\mu$ mean that μ/λ is a vertical-strip and a horizontal strip, respectively (see Chapter 1 if you do not remember). The indicator function $\mathbb{1}_{\nu\preccurlyeq\mu}$ is defined by 1 if $\nu\preccurlyeq\mu$ and 0 otherwise (similar for $\mathbb{1}_{\nu\preccurlyeq_{\mathrm{v}}\mu}$). Now for the Markov chain, each transition probability $\mathbb{P}_\ell(\nu^{(\ell)},\mu^{(\ell-1)}\to\mu^{(\ell)})$ is zero unless $\mu^{(\ell-1)}\preccurlyeq\mu^{(\ell)}$, by the second indicator function, so $\varnothing\preccurlyeq\mu^{(1)}\preccurlyeq\cdots\preccurlyeq\mu^{(N)}$. This means that for every t,

$$\varnothing\preccurlyeq\lambda^{(1)}(t)\preccurlyeq\cdots\preccurlyeq\lambda^{(N)}(t),$$

which implies that $\left|\lambda^{(k)}(t)\right|\le k$ for every $k=1,\dots,N$. Moreover, using the chain notation and the definition of vertical-strip, we obtain

$$\begin{aligned}&\mathbb{P}\left(\lambda^{(k)}(t+1)=\mu^{(k)}\middle|\,\lambda^{(k)}(t)=\nu^{(k)},\lambda^{(k-1)}(t+1)=\mu^{(k-1)}\right)\\&=\mathbb{P}_k(\nu^{(k)},\mu^{(k-1)}\to\mu^{(k)})=\widetilde{C}_k(\nu^{(k)},\mu^{(k-1)})\mathbb{1}_{\mu^{(k-1)}\preccurlyeq\mu^{(k)}}\prod_{i=1}^{k}b^{\mu_i^{(k)}-\nu_i^{(k)}}\mathbb{1}_{\mu_i^{(k)}-\nu_i^{(k)}\in\{0,1\}}.\end{aligned}\tag{3.8}$$

Now consider the update of the first partition, which we know that it has length 1, that is, $\lambda^{(1)}(t)=\left(\lambda_1^{(1)}(t)\right)$. Let $m,n\in\mathbb{Z}_+$, then by (3.8),

$$\mathbb{P}\left(\lambda^{(1)}(t+1)=(n)\middle|\,\lambda^{(1)}(t)=(m)\right)=\mathbb{P}_1(\lambda^{(1)}(t),\varnothing\to\lambda^{(1)}(t+1))=C_1(m)b^{n-m}\mathbb{1}_{n-m\in\{0,1\}}.$$

The probability is non-zero in just two cases: $n=m$ and $n=m+1$, so

$$\begin{aligned}\mathbb{P}\left(\lambda^{(1)}(t+1)=(m+1)\middle|\,\lambda^{(1)}(t)=(m)\right)&=bC_1(m);\\\mathbb{P}\left(\lambda^{(1)}(t+1)=(m)\middle|\,\lambda^{(1)}(t)=(m)\right)&=C_1(m),\end{aligned}$$

thus $C_1(m) = \frac{1}{1+b}$. We conclude that the first partition is increased by 1 with probability $\frac{b}{1+b}$ and stays the same with probability $\frac{1}{1+b}$.

Going further, we study now the update of the first two partitions $\left(\lambda^{(1)}(t), \lambda^{(2)}(t)\right)$. Assume $\lambda^{(1)}(t) = (m)$, $\lambda^{(2)}(t) = (m_1, m_2)$, $\lambda^{(1)}(t+1) = (n)$, $\lambda^{(2)}(t+1) = (n_1, n_2)$ and let us consider the possibilities for n, n_1 and n_2. By the interlacing property $\lambda^{(1)} \preccurlyeq \lambda^{(2)}$ for any t, we have $m_2 \leq m \leq m_1$ and $n_2 \leq n \leq n_1$. We have by the definition of the Markov chain,

$$\begin{aligned}\mathbb{P}\left(\lambda^{(1)}(t+1) = (n), \lambda^{(2)}(t+1) = (n_1, n_2) \middle| \lambda^{(1)}(t) = (m), \lambda^{(2)}(t) = (m_1, m_2)\right) \\ = \mathbb{P}\left(\lambda^{(2)}(t+1) = (n_1, n_2) \middle| \lambda^{(1)}(t+1) = (n), \lambda^{(2)}(t) = (m_1, m_2)\right) \\ \times \mathbb{P}\left(\lambda^{(1)}(t+1) = (n) \middle| \lambda^{(1)}(t) = (m)\right)\end{aligned}$$

and we know that the second factor is the update of the first partition $\lambda^{(1)}$, which we computed before. So by (3.8), the probability above is equal to

$$C_2(m, m_1, m_2) \mathbb{1}_{n_2 \leq n \leq n_1} \prod_{i=1}^{2} b^{n_i - m_i} \mathbb{1}_{n_i - m_i \in \{0,1\}} \left(\frac{b}{1+b} \mathbb{1}_{n=m+1} + \frac{1}{1+b} \mathbb{1}_{n=m} \right). \tag{3.9}$$

Assuming $n = m$ or $n = m+1$, the probability above is zero unless we have $n_1 - m_1 \in \{0,1\}$, $n_2 - m_2 \in \{0,1\}$ and $n_2 \leq n \leq n_1$, so we need to consider the cases below.

1. We first assume $n = m$.

   a) If $m_2 < m \leq m_1$, then the conditions $n_1 - m_1 \in \{0,1\}$, $n_2 - m_2 \in \{0,1\}$ imply the condition $n_2 \leq n \leq n_1$. The probabilities that $(n_1, n_2) - (m_1, m_2) = (0,0), (1,0), (0,1), (1,1)$ are equal to $C_2(m, m_1, m_2)$ times $1, b, b, b^2$, respectively, so we obtain $C_2(m, m_1, m_2) = \frac{1}{(1+b)^2}$. Now we can write (3.9) as

   $$\left(\frac{\mathbb{1}_{n_1 = m_1}}{1+b} + \frac{b \mathbb{1}_{n_1 = m_1 + 1}}{1+b} \right) \left(\frac{\mathbb{1}_{n_2 = m_2}}{1+b} + \frac{b \mathbb{1}_{n_2 = m_2 + 1}}{1+b} \right) \frac{1}{1+b},$$

   therefore, each row of $\lambda^{(2)}(t)$ may increase by 1, independently of each other, with probability $\frac{b}{1+b}$.

   b) If $m_2 = m \leq m_1$, then the condition $n_2 \leq n$ does not hold when $(n_1, n_2) - (m_1, m_2) \in \{(0,1), (1,1)\}$, since $n_2 = 1 + n$. So the probability (3.9) is zero unless $n_2 = m_2$. The probabilities that $(n_1, n_2) - (m_1, m_2) = (0,0), (1,0)$ are proportional to 1 and b, respectively, so we obtain $C_2(m, m_1, m_2) = \frac{1}{1+b}$. Then (3.9) equals

   $$\left(\frac{\mathbb{1}_{n_1 = m_1}}{1+b} + \frac{b \mathbb{1}_{n_1 = m_1 + 1}}{1+b} \right) \mathbb{1}_{n_2 = m_2} \frac{1}{1+b},$$

   and this means that only the first row of $\lambda^{(2)}(t)$ may grow by 1 with probability $\frac{b}{1+b}$. The second row stays the same, since it is **blocked** by $\lambda^{(1)}(t+1)$.

2. Now we consider $n = m+1$.

   a) If $m_2 \le m < m_1$, then this case is analogous to 1.(*a*): each row of $\lambda^{(2)}(t)$ may increase by 1, independently of each other, with probability $\frac{b}{1+b}$.

   b) If $m_2 \le m = m_1$, then for $(n_1,n_2)-(m_1,m_2) \in \{(0,0),(0,1)\}$, the condition $n \le n_1$ is violated, since $n_1 = n-1$. So (3.9) is zero unless $n_1 = m_1+1$. The probabilities that $(n_1,n_2)-(m_1,m_2) = (1,0),(1,1)$ are proportional to b and b^2, respectively, so we obtain $C_2(m,m_1,m_2) = \frac{1}{b(1+b)}$. Then (3.9) equals

   $$\mathbb{1}_{n_1=m_1+1}\left(\frac{\mathbb{1}_{n_2=m_2}}{1+b} + \frac{b\mathbb{1}_{n_2=m_2+1}}{1+b}\right)\frac{b}{1+b}.$$

   This computation shows that, when $\lambda^{(1)}(t)$ is increased by 1 (which occurs with probability $\frac{b}{1+b}$), the row $\lambda_1^{(2)}(t)$ is also increased by 1 automatically, that is, $\lambda_1^{(2)}$ is **pushed** by $\lambda_1^{(1)}$. Also, the second row $\lambda_2^{(2)}(t)$ may increase with probability $\frac{b}{1+b}$.

Assume we have a coin with probability $\frac{b}{1+b}$ of turning heads and $\frac{1}{1+b}$ for tails (we can assume $b = 1$ for a fair coin). All the computations above say that $\left(\lambda^{(1)}(t),\lambda^{(2)}(t)\right)$ evolves with the following **block-push** dynamics.

1. The first partition $\lambda^{(1)}(t) = \left(\lambda_1^{(1)}(t)\right)$ grows by 1, that is, $\lambda_1^{(1)}(t+1) = \lambda_1^{(1)}(t)+1$ if we obtain heads after flipping the coin;
2. If $\lambda_1^{(1)}(t+1)$ is larger that $\lambda_1^{(2)}(t)$, then $\lambda_1^{(2)}(t)$ is pushed by 1 (case 2.(b));
3. If $\lambda_1^{(1)}(t+1)$ is equal to $\lambda_2^{(2)}(t)$, then $\lambda_2^{(2)}(t)$ is blocked and stays the same (case 1.(b));
4. The rows that are neither pushed nor blocked grow independently by flipping the coin.

The evolution of $\left(\lambda^{(1)}(t),\ldots,\lambda^{(N)}(t)\right)$ follows the same dynamics recursively. This is called the **block-push process**. We now translate this process to the language of particle systems. For each pair i,j satisfying $1 \le i \le j \le N$, we define a particle x_i^j that is positioned at time t at

$$x_i^j(t) := \lambda_{j+1-i}^{(j)}(t) - j + i.$$

Since $\lambda^{(j)}(t)$ is a partition, we have for each $j = 1,\ldots,N$ and $i = 1,\ldots,j-1$,

$$x_i^j(t) = \lambda_{j+1-i}^{(j)}(t) - j + i < \lambda_{j+1-(i+1)}^{(j)}(t) - j + i + 1 = x_{i+1}^j(t),$$

so for every t, $x_1^j(t) < x_2^j(t) < \cdots < x_j^j(t)$. This property is desired, since we do not want to have two particles at the same position.

Moreover, for every t, we have the **interlacing property**

$$x_i^{j+1}(t) < x_i^j(t) \leq x_{i+1}^{j+1}(t) \tag{3.10}$$

for each $j = 1, \ldots, N-1$ and $i = 1, \ldots, j$. To see this, we use the interlacing property of the Markov chain $\lambda^{(j)}(t) \preccurlyeq \lambda^{(j+1)}(t)$, which means that, for each i,

$$\lambda_{i+1}^{(j+1)}(t) \leq \lambda_i^{(j)}(t) \leq \lambda_i^{(j+1)}(t).$$

So we have (omitting the dependence on t)

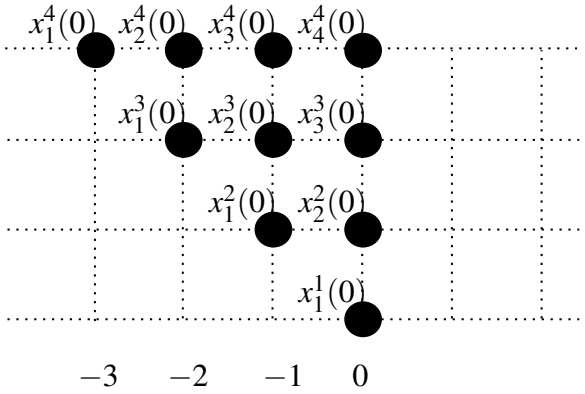
$$x_i^{j+1} = \lambda_{j+2-i}^{(j+1)} - (j+1) + i < \lambda_{j+1-i}^{(j)} - j + i = x_i^j,$$

and for the second inequality,

$$x_i^j = \lambda_{j+1-i}^{(j)} - j + i \leq \lambda_{j+1-i}^{(j+1)} - j + i = \lambda_{(j+1)+1-(i+1)}^{(j+1)} - (j+1) + (i+1) = x_{i+1}^{j+1}.$$

We can imagine that each partition $\lambda^{(j)}$ defines a level of particles $x_1^j, \ldots, x_j^j$. Since $\lambda^{(j)}(0) = \varnothing$ for each j, the positions at time $t = 0$ are given by $x_i^j(0) = i - j$. We illustrate this configuration in Figure 16.

Figure 16 – Block-Push process. Initial condition.



Source: Elaborated by the author.

The particles want to move one unit to the right, staying in the same level. The movement of the particles that can jump is decided by the flip of a coin (if it turns heads, the particles move, otherwise it does not). By the definition of the Markov chain, from the interlacing property (3.10) and from the independence of the movement of some particles, the updates follow the following rules:

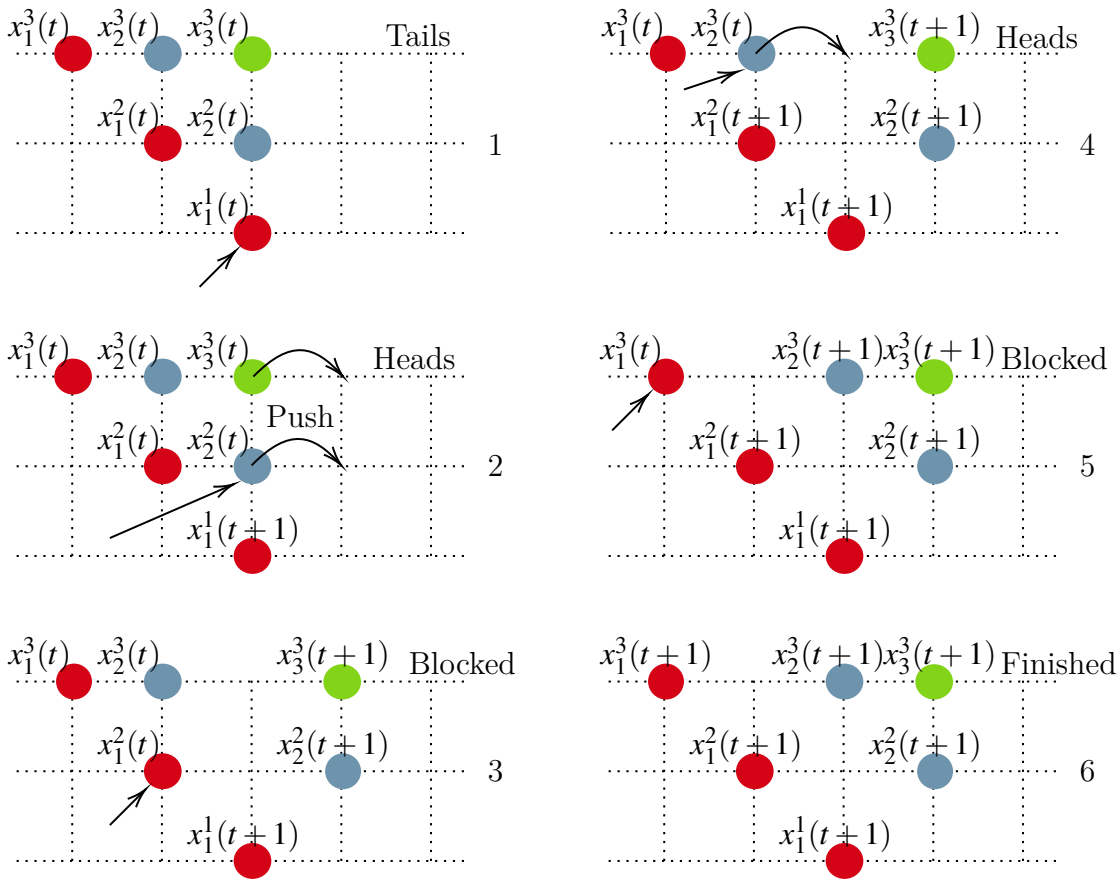
- The partial updates go from the bottom to the top and from the right to the left (a particle cannot move twice in the same update);

- (Push) When a particle moves, the adjacent particle above also moves (imagine that you have a pile of books, if you slowly push the lower one, the ones above also move);
- (Block) A particle is "blocked" by the ones in the level below, that is, the ones with same i-index can never be at same position.

In Figure 17, we see an example of simulation. Unfortunately, this text is not written in paper from Harry Potter universe, so we cannot show animated computer simulations here, but you can check Patrik Ferrari's website (link in Section 1.1.1).

Remember from the Introduction, that we want to study the totally asymmetric simple exclusion process (TASEP). In fact, we can recover the TASEP from the Block-Push process. Moreover, we can use all the results obtained for Schur processes to compute probabilities for the TASEP. We see this in the next section.

Figure 17 – Simulation of Block-Push process. Read the figures in the order of the numbers in the right of each one. Particles with same i-index have the same color, to emphasize the blocking rule. The arrows pointing to the particles indicate which particles we are updating.



Source: Elaborated by the author.

3.4.2. The discrete-time TASEP

Recall the **discrete-time TASEP**, presented in the Introduction (Section 1.1). Considering the left-most particles in the block-push process discussed above, which are given by $y_i(t) := x_1^i(t)$, we obtain the discrete-time TASEP (see Figure 18). The initial conditions are $y_i(0) = 1-i$, for each $i = 1,\dots,N$. Also, recall that b defines the probability $\frac{b}{1+b}$, which is the probability for the jump of a non-blocked particle.

The next result shows how simple is to compute a probability in the TASEP, using the Schur process framework. We want to compute the probability that all particles have passed a given point at time t, which is impracticable to do by hand. Note that this probability is the same as the probability that the last particle $y_N(t)$ has passed the point.

Theorem 3.2. Consider the discrete-time TASEP with N particles $(y_N(t) < \cdots < y_1(t))$ with initial condition $y_i(0) = 1-i$. Then for every $\ell \geq 0$, we have

$$\mathbb{P}_{\text{TASEP}}(y_N(t) - 1 + N \geq \ell) = \mathbb{S}_{\sigma_1,\sigma_2}(\nu_N \geq \ell),$$

where

$$\sigma_1 = ((\underbrace{1,\dots,1}_{N},0,0,\dots);\mathbf{0};0) \quad \text{and} \quad \sigma_2 = (\mathbf{0};(\underbrace{b,\dots,b}_{t},0,0,\dots);0).$$

Proof. Consider the block-push process discussed in previous section. By Theorem 3.1 (or Corollary 3.2), the distribution at time t is the Schur process parametrized by $\rho_1,\dots,\rho_N,\sigma_2$, where

$$\sigma_2 = \rho^- \cup \underbrace{\rho' \cup \cdots \cup \rho'}_{t} = \rho' \cup \cdots \cup \rho' = (\mathbf{0};(\underbrace{b,\dots,b}_{t},0,0,\dots);0),$$

that is,

$$\mathbb{P}\left(\left(\lambda^{(1)}(t),\dots,\lambda^{(N)}(t)\right) = \left(\mu^{(1)},\dots,\mu^{(N)}\right)\right) = \mathbb{S}_{\rho_1,\dots,\rho_N,\sigma_2}\left(\mu^{(1)},\dots,\mu^{(N)}\right).$$

By Proposition 3.3, the one-dimensional marginal with respect to the N-th coordinate is given by the Schur measure

$$\mathbb{P}\left(\lambda^{(N)}(t) = \mu\right) = \mathbb{S}_{\sigma_1,\sigma_2}(\mu),$$

where

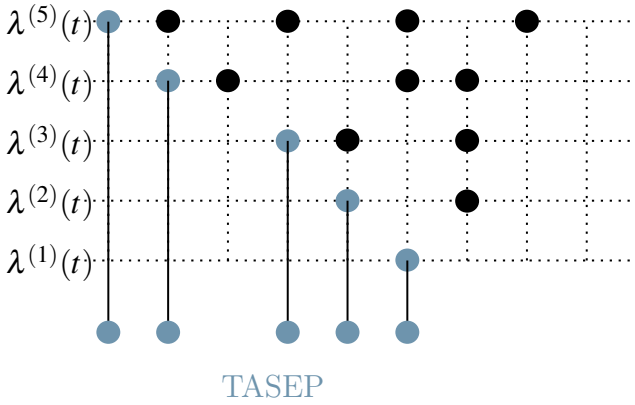
$$\sigma_1 = \rho_1 \cup \cdots \cup \rho_N = ((\underbrace{1,\dots,1}_{N},0,0,\dots);\mathbf{0};0).$$

Since $y_N(t) = x_1^N(t) = 1 - N + \lambda_N^{(N)}(t)$, we have

$$\mathbb{P}_{\text{TASEP}}\left(y_N(t) - 1 + N \geq \ell\right) = \mathbb{P}\left(\lambda_N^N(t) \geq \ell\right) = \mathbb{S}_{\sigma_1,\sigma_2}(\nu_N \geq \ell),$$

as we wanted to prove. Notice that from Example 3.2, the Schur measure $\mathbb{S}_{\sigma_1,\sigma_2}$ is well-defined for $b \in [0,1]$. □

Figure 18 – We recover the TASEP from the Block-Push process, projecting the left-most particles into the "ground".



Source: Elaborated by the author.

Example 3.6. (Continuous-time TASEP) Consider a non-blocked particle in the discrete-time TASEP. The particle jumps with probability $\frac{b}{1+b}$ and stands still with probability $\frac{1}{1+b}$. So the expected waiting-time for the jump is $\frac{1+b}{b}$. Let $\tau := bt$ be a new variable for the time, so that the expected waiting time in τ is $1+b$. Making $b \to 0$, we obtain the continuous-time TASEP, where the expected time for the jumps is 1. You can imagine that each particle has a independent clock that rings with rate 1. When the clock of a particle rings, the particle moves if not blocked.

From now on, we are interested in understanding the asymptotics of the TASEP, that is, we want to study what happens when $t \to \infty$. Nonetheless, we first need to understand a little about Determinantal Point Processes, which we present in the next chapter.

CHAPTER

4

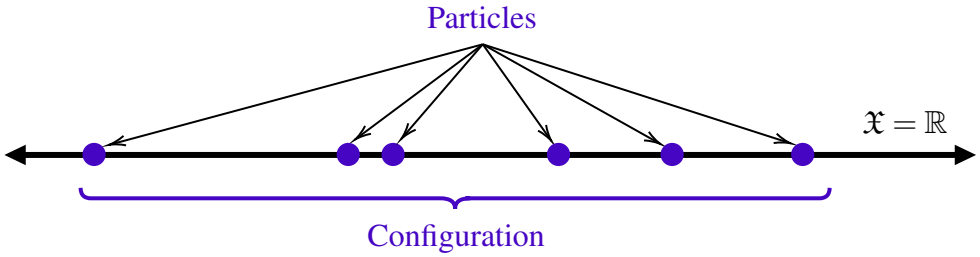
DETERMINANTAL POINT PROCESSES

In this chapter, we present the theory of random point processes and how we apply it to study the TASEP. We are interested in a particular type of random point process: the determinantal point processes. They have good properties that are applied together with the Schur process framework to compute probabilities on the TASEP model. Some theoretical aspects are beyond the scope of this work, since we are mostly interested in the physical meaning and math properties to work with. We are following (BAIK; DEIFT; SUIDAN, 2016), which is based on (JOHANSSON, 2005), so the interested reader should check these references for more details. The lecture notes by Baik (2018) are also widely used here.

4.1. Summary of the Chapter

Using few words, a **random point process** is a set of particles distributed randomly in some space, with respect to some probability measure. For example, in Figure 19, we see an example of random configuration of particles on $\mathbb{R}$.

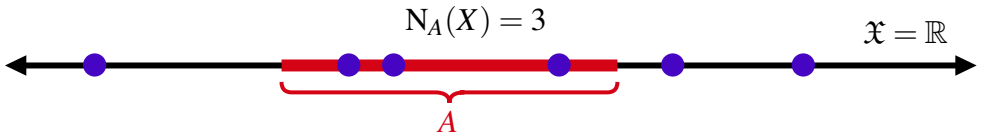
Figure 19 – Illustration of a configuration in $\mathfrak{X} = \mathbb{R}$.



Source: Elaborated by the author.

We are interested in counting particles inside sets. Given a set A and a configuration X, we define $\mathrm{N}_A(X)$ as the number of particles of X lying on A (Figure 20).

Figure 20 – The counting function N_A.



Source: Elaborated by the author.

If we freeze the time t in the TASEP model, what we have is a configuration X of random particles $y_1(t) > \cdots > y_N(t)$ on $\{N-1, N-2, \ldots\}$, that is, a random point process. Recall the probability

$$\mathbb{P}_{\mathrm{TASEP}}(y_N(t) - 1 + N \geq \ell) = \mathbb{S}_{\sigma_1,\sigma_2}(\nu_N \geq \ell), \tag{4.1}$$

that we compute using the Schur process. Consider the set $A = \{N-1, N-2, \ldots, \ell - N\}$. The probability on the left-hand side of (4.1) is exactly $\mathbb{P}(\mathrm{N}_A(X) = 0)$, which we call a **gap probability**.

The gap probabilities are related to the so-called **correlation functions**, a infinite set of functions $(\rho_n(x_1, \ldots, x_n))_{n=1}^{\infty}$ intrinsically present in a random point process (we show more details later). More precisely, we have a nice formula to compute $\mathbb{P}(\mathrm{N}_A(X) = 0)$ depending only on the correlation functions. This formula is given by

$$\mathbb{P}(\mathrm{N}_A(X) = 0) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int_{A^n} \rho_n(y_1, \ldots, y_n) \mu^{\otimes n}(\mathrm{d}^n y). \tag{4.2}$$

A special type of random point process are the **determinantal point processes**. They are characterized by the existence of a two-variable function $K(x, y)$ such that each correlation function is given by

$$\rho_n(x_1, \ldots, x_n) = \det\left[K(x_i, x_j)\right]_{i,j=1}^{n}$$

for every $n \in \mathbb{N}$. The function K is called **correlation kernel**. The correlation kernel is useful since it compresses some information in a single function. For example, the formula in (4.2) would depend just on K. We will show later that the Schur measure induces a determinantal point process in $\mathbb{Z}$, so that (4.1) turns to be

$$\mathbb{P}_{\mathrm{TASEP}}\left(y_N(t) - 1 + N \geq \ell\right) = \det\left[1 - K_{N,\ell,t}\right], \tag{4.3}$$

where $K_{N,\ell,t}(m,n) := \mathbb{1}_{A_\ell}(m) K_{N,t}(m,n) \mathbb{1}_{A_\ell}(n)$, with

$$K_{N,t}(m,n) = \frac{1}{(2\pi i)^2} \oint_{\Gamma_0} \oint_{\Gamma_{w,0}} \frac{w^{n-1}}{z^m} \frac{1}{z-w} \frac{(1-w)^N \left(1 + \frac{b}{w}\right)^t}{(1-z)^N \left(1 + \frac{b}{z}\right)^t} \mathrm{d}z \mathrm{d}w.$$

The expression $\det\left[1 - K_{N,\ell,t}\right]$ is not a matrix determinant, it is called a **Fredholm determinant** and we will define it later.

In summary, we join together formulas from Schur processes and point processes to obtain results for the TASEP.

4.2. Random Point Processes

Let us denote by $\mathfrak{X}$ the **state space** or **particle space**, it is usually a reasonable metric space, like $\mathbb{R}$, $\mathbb{R}^n$, $\mathbb{Z}$ or some subset of these ones. As the name suggests, it is the space where some particles lie on. A locally finite subset $X \subset \mathfrak{X}$ (i.e. with no accumulation points) is called a **configuration** in $\mathfrak{X}$ (see Figure 19). The set of all configurations in $\mathfrak{X}$ is denoted by $\mathbf{Conf}(\mathfrak{X})$.

Consider a bounded set $A \subset \mathfrak{X}$. Since any $X \in \mathbf{Conf}(\mathfrak{X})$ has no accumulation points, the intersection $A \cap X$ must have finite elements, so the function

$$\begin{aligned} \mathrm{N}_A : \mathbf{Conf}(\mathfrak{X}) &\to \{0,1,2,\ldots\} \\ X &\mapsto \mathrm{N}_A(X) := |A \cap X| \end{aligned}$$

is well-defined, where $|A \cap X|$ denotes the number of points in $A \cap X$. The function N_A counts the number of points of X lying on A (see Figure 20). Every time that we say *particle* the reader should understand as a *particle of some configuration*.

Note 4.1. The locally finite property of X is also necessary for physical meaning. To illustrate, take $\mathfrak{X} = [0,\infty]$ and say that each particle of a configuration X represents the time that a new person arrives at the queue for lunch at the university restaurant. If X is not locally finite, we could have the first person arriving at the time $t_1 = 1/2$, the second at $t_2 = 3/4$, the third at $t_3 = 7/8$ and so on, such that the n-th person arrives at $t_n = \frac{2^n-1}{2^n}$. So we have infinitely many people arriving in the finite interval of time $[0,1]$ (see Figure 21), which clearly is not a realistic situation.

Figure 21 – The locally finite property is important.

Source: Elaborated by the author.

We equip $\mathbf{Conf}(\mathfrak{X})$ with the smallest σ-algebra such that the map N_B is measurable for every bounded set $B \subset \mathfrak{X}$, this σ-algebra is denoted by $\mathscr{M}_{\mathfrak{X}}$ (it is well-defined since $\mathscr{P}(\mathfrak{X})$ has the latter property).

Definition 4.1. A **random point process** on $\mathfrak{X}$ is a probability space $(\mathbf{Conf}(\mathfrak{X}), \mathscr{M}_{\mathfrak{X}}, \mathbb{P})$ where $\mathbb{P}$ is a probability measure on the measurable space $(\mathbf{Conf}(\mathfrak{X}), \mathscr{M}_{\mathfrak{X}})$.

Let A be a bounded set. For each $n \in \mathbb{N}$, the set

$$\mathrm{N}_A^{-1}(\{n\}) = \{X \in \mathbf{Conf}(\mathfrak{X}) : \mathrm{N}_A(X) = n\}$$

is measurable, hence it does make sense to calculate $\mathbb{P}(\mathrm{N}_A(X) = n)$, i.e., the probability of having n points of X lying on A.

Example 4.1. (Poisson Process on $\mathbb{R}$) Let $\mathfrak{X} = \mathbb{R}$. Consider a configuration $X \subset \mathbb{R}$ of particles uniformly distributed with density 1 (that is, there is approximately one particle in each interval of length one) and a bounded set $A \subset \mathbb{R}$. Let us compute $\mathbb{P}(\xi(A) = n)$, for any $n \in \mathbb{N}$. For some particle $x \in X$, we cannot directly calculate $\mathbb{P}(x \in A)$ using the proportion of lengths $\frac{|A|}{|\mathbb{R}|}$, since $|\mathbb{R}| = \infty$. To solve this problem, take $M \in \mathbb{N}$ sufficiently large and set our particle space to be $\mathfrak{X}_M = [-M/2, M/2]$, where $A \subset \mathfrak{X}_M$. Now let $X = \{x_1, \ldots, x_M\}$ be a configuration uniformly distributed. Notice that X has density 1 and $\mathbb{P}(x_i \in A) = |A|/M$ for every i. Let $n \leq M$, so

$$\mathbb{P}(\mathrm{N}_A(X) = n) = \binom{M}{n} \left(\frac{|A|}{M}\right)^n \left(1 - \frac{|A|}{M}\right)^{M-n} \xrightarrow{M\to\infty} \mathrm{e}^{-|A|} \frac{|A|^n}{n!},$$

thus we conclude that $\mathrm{N}_A(X)$ has the Poisson distribution with parameter $|A|$. One can also prove that, if A_1 and A_2 are disjoint bounded intervals, then the random variables $\mathrm{N}_{A_1}(X)$ and $\mathrm{N}_{A_2}(X)$ are independent. For more details on the Poisson process, see (BILLINGSLEY, 1995, Chapter 23).

In the last example, the careful reader would notice that the Poisson process was not precisely defined accordingly to Definition 4.1. This is due to the difficulty to calculate probabilities for arbitrary elements of $\mathscr{M}_{\mathfrak{X}}$. Although we can compute $\mathbb{P}(\mathrm{N}_A(X) = n)$, it would be much harder to compute $\mathbb{P}(\mathbb{Z} \subset X)$ or $\mathbb{P}(\text{primes} \subset X)$. Later we discuss a little about the well-definition of point processes.

Example 4.2. (Bernoulli process) Let $\mathfrak{X} = \mathbb{Z}$. Given $p \in (0,1)$, put a particle on each point $n \in \mathbb{Z}$, independently, with probability p. That gives us a random configuration X, and consequently, a random point process.

4.2.1. Correlation Functions

Since it is impracticable to provide any example of random point process using Definition 4.1, we discuss the main tool for having well-defined random point processes: the correlations functions. The study of such functions goes back to Lenard (1973). Most of the proofs in this section are omitted as they go beyond the scope of this work, so see (LENARD, 1973; LENARD, 1975a; LENARD, 1975b) for more details about this theory.

Considering the state spaces $\mathfrak{X}$ and $\mathfrak{X}^n$, we define the map

$$\begin{aligned} \Xi_n : \mathrm{Conf}(\mathfrak{X}) &\longrightarrow \mathrm{Conf}(\mathfrak{X}^n) \\ X &\longmapsto \Xi_n(X) = X^{(n)} := \left\{(x_{i_1}, \ldots, x_{i_n}) | x_{i_k} \neq x_{i_j}\right\}, \end{aligned}$$

where $X = \{x_1, x_2, \ldots\}$. That is, $X^{(n)}$ is defined by taking all n-tuples $(x_{i_1}, \ldots, x_{i_n})$ of pairwise distinct points of X (the points are taken distinct to avoid counting problems when proving some results).

Proposition 4.1. The map Ξ_n defined above is measurable.

The importance of the last proposition is that if we have a point process on $\mathfrak{X}$, the map Ξ_n induces a point process $(\mathbf{Conf}(\mathfrak{X}^n), \mathscr{M}_{\mathfrak{X}^n}, \mathbb{P}_n)$ on $\mathfrak{X}^n$ via

$$\mathbb{P}_n(\Delta) = \mathbb{P}\left(\Xi_n^{-1}(\Delta)\right), \quad \Delta \in \mathscr{M}_{\mathfrak{X}^n}.$$

Given a bounded set B, counting the number of points of $X^{(n)}$ lying on B^n is straightforward: we have $\mathrm{N}_B(X)$ possibilities for the first coordinate, $\mathrm{N}_B(X) - 1$ for the second and so on. So we have the following proposition and an immediate corollary.

Proposition 4.2. If B is bounded, then

$$\mathrm{N}_{B^n}\left(X^{(n)}\right) = \mathrm{N}_B(X)(\mathrm{N}_B(X) - 1)\cdots(\mathrm{N}_B(X) - n + 1).$$

Corollary 4.1. Let $B_1, \ldots, B_\ell$ be disjoint bounded sets in $\mathfrak{X}$ and $n_1, \ldots, n_\ell \in \mathbb{N}$ be such that $n_1 + \cdots + n_\ell = n$, we have

$$\mathrm{N}_{B_1^{n_1} \times \cdots \times B_\ell^{n_\ell}}\left(X^{(n)}\right) = \prod_{i=1}^{\ell} \mathrm{N}_{B_i}(X)(\mathrm{N}_{B_i}(X) - 1)\cdots(\mathrm{N}_{B_i}(X) - n_i + 1).$$

Definition 4.2. Given a point process on $\mathfrak{X}$ and a bounded set $A \subset \mathfrak{X}^n$, the expected value

$$M_n(A) \coloneqq \mathbb{E}\left[\mathrm{N}_A\left(X^{(n)}\right)\right]$$

is called the ***n*-th correlation measure** of the point process.

By Corollary 4.1, the next results are immediate.

Corollary 4.2. For disjoint bounded sets $B_1, \ldots, B_\ell \subset \mathfrak{X}$ with $n_1 + \cdots + n_\ell = n$, we have

$$M_n\left(B_1^{n_1} \times \cdots \times B_\ell^{n_\ell}\right) = \mathbb{E}\left[\prod_{i=1}^{\ell} \mathrm{N}_{B_i}(X)(\mathrm{N}_{B_i}(X) - 1)\cdots(\mathrm{N}_{B_i}(X) - n_i + 1)\right].$$

Corollary 4.3. The correlations measures are symmetric, that is, for any permutation $\sigma \in \mathscr{S}_n$, we have

$$M_n(B_1 \times \cdots \times B_n) = M_n(B_{\sigma(1)} \times \cdots \times B_{\sigma(n)}).$$

Note 4.2. (BORODIN; GORIN, 2012; SOSHNIKOV, 2000) The n-th correlation measure has the following characterization: for all bounded Borel function with bounded support,

$$\int_{\mathfrak{X}} f \mathrm{d}M_n = \mathbb{E}\left[\sum_{x_{i_1} \neq \cdots \neq x_{i_n}} f(x_{i_1}, \ldots, x_{i_n})\right]$$

To see this, we first prove it for characteristic functions $f = \chi_A$ and then extend for other functions, as we usually do in Measure Theory.

Let μ be a **reference measure** on $\mathfrak{X}$ (as the Lebesgue measure on $\mathbb{R}$ or the counting measure on $\mathbb{Z}$). Suppose M_n is absolutely continuous with respect to the product measure

$$\mu^{\otimes n} = \mu \times \cdots \times \mu,$$

on $\mathfrak{X}^n$, denoting this relation by $M_n \ll \mu^{\otimes n}$. By Radon-Nikodym theorem (see (FOLLAND, 1999)), there exists a unique density ρ_n of M_n with respect to $\mu^{\otimes n}$ (that means that any other density is equal to ρ_n μ-almost everywhere). In other words, we have $\mathrm{d}M_n = \rho_n \mathrm{d}\mu^{\otimes n}$.

Definition 4.3. Given a reference measure μ on $\mathfrak{X}$ such that $M_n \ll \mu^{\otimes n}$ on $\mathfrak{X}^n$, the density ρ_n of M_n with respect to $\mu^{\otimes n}$ is called the ***n*-th correlation function** for the random point process.

Note 4.3. One can see that ρ_n is μ-almost everywhere non-negative since M_n is a measure. Moreover, it can be proven that ρ_n is symmetric (SOSHNIKOV, 2000).

Example 4.3. Suppose $\mathfrak{X} = \mathbb{R}$ and $M_1 \ll m$, where m is the Lebesgue measure. We have

$$\mathbb{E}\left[\mathrm{N}_A(X)\right] = M_1(A) = \int_A \rho_1(x)\mathrm{d}m. \tag{4.4}$$

Take a small number $\Delta x > 0$, so we can do the approximation

$$\rho_1(x)\Delta x \approx \int_x^{x+\Delta x} \rho_1(t)\mathrm{d}t = M_1\left([x, x+\Delta x]\right) = \mathbb{E}\left[\mathrm{N}_{[x,x+\Delta x]}(X)\right] = \sum_i \mathbb{P}(x_i \in [x, x+\Delta x]),$$

where $X = \{x_1, x_2, \ldots\}$. Notice that as Δx goes small, it is not possible to have more than one particle on $[x, x+\Delta x]$, since every point of X is isolated, thus

$$\begin{aligned}\sum_i \mathbb{P}(x_i \in [x, x+\Delta x]) &= \sum_i \mathbb{P}(x_i \in [x, x+\Delta x], x_j \notin [x, x+\Delta x] \text{ for } j \neq i) \\ &= \mathbb{P}(\text{there is exactly one particle in } [x, x+\Delta x]).\end{aligned}$$

With this computation, we concluded that $\rho_1(x)$ is the density for the probability to find a particle in x. This makes sense, because integrating ρ_1 in some interval A gives the expected number of particles in A (Equation (4.4)). We also observe this for ρ_n, where

$$\rho_n(x_1, \ldots, x_n)\Delta x_1 \cdots \Delta x_n = \mathbb{P}(\text{there is exactly one particle in each } [x_i, x_i + \Delta x_i]).$$

Example 4.4. When $\mathfrak{X}$ is discrete, given any finite set $A = \{x_1, \ldots, x_n\} \subset \mathfrak{X}$, we have analogously to Example 4.3,

$$\rho_n(x_1, \ldots, x_n) = \mathbb{P}(A \subset X) = \mathbb{P}(\text{there are particles in } x_1, \ldots, x_n).$$

The last examples show that correlation functions actually have a physical meaning, so it is not a surprise that they can completely determine random point processes under some conditions. For this reason, Lenard (1973, 1975a, 1975b) was able to find

necessary and sufficient conditions for a sequence of correlation functions $(\rho_n)_1^\infty$ to be associated with a unique random point process. In other words, under mild conditions, any two distinct point processes on $\mathfrak{X}$ cannot have the same sequence $(\rho_n)_1^\infty$ of correlation functions. We do not enter in such details here.

Theorem 4.1. (LENARD, 1973; LENARD, 1975a; LENARD, 1975b) A random point process on a discrete space $\mathfrak{X}$ is uniquely determined by its correlation functions.

To provide good examples of random point processes, we still need some results. Let $\phi : \mathfrak{X} \to \mathbb{C}$ be a bounded Borel function, we denote $\|\phi\|_\infty = \sup_{x\in\mathfrak{X}} |\phi(x)|$.

Proposition 4.3. Let $\phi : \mathfrak{X} \to \mathbb{C}$ be a bounded Borel function with support in a bounded set B. If a point process satisfies

$$\sum_{n=0}^{\infty} \frac{\|\phi\|_\infty^n}{n!} M_n(B^n) < \infty,$$

then for a configuration $X = \{x_1, x_2, \ldots\}$, we have

$$\mathbb{E}\left[\prod_{i=1}^{\infty}(1+\phi(x_i))\right] = \sum_{n=0}^{\infty}\frac{1}{n!}\int_{\mathfrak{X}^n}\left(\prod_{i=1}^{\infty}\phi(y_i)\right)\rho_n(y_1,\ldots,y_n)\mu^{\otimes n}(\mathrm{d}^n y).$$

Proof. (BAIK; DEIFT; SUIDAN, 2016) □

Example 4.5. (Gap Probabilities) A problem of interest in random point processes is to calculate the probability of finding no particles in a bounded set B. Proposition 4.3 helps us with this task by taking $\phi = -\chi_B$. Indeed,

$$\begin{aligned}\mathbb{P}(\mathrm{N}_B(X)=0) &= 1\cdot\mathbb{P}(\mathrm{N}_B(X)=0) + 0\cdot\mathbb{P}(\mathrm{N}_B(X)\geq 1) = \mathbb{E}\left[\prod_{i=1}^{\infty}(1-\chi_B(x_i))\right]\\ &= \sum_{n=0}^{\infty}\frac{(-1)^n}{n!}\int_{B^n}\rho_n(y_1,\ldots,y_n)\mu^{\otimes n}(\mathrm{d}^n y),\end{aligned}$$

so we have a nice formula to compute $\mathbb{P}(\mathrm{N}_B(X)=0)$ depending only on the correlation functions.

Another application for Proposition 4.3 is in point processes that have a "last particle", that is, a point process X on $\mathfrak{X}=\mathbb{R}$ (or $\mathfrak{X}=\mathbb{Z}$) that satisfies $\mathrm{N}_{(t_0,\infty)}(X) < \infty$ for some $t_0 \in \mathbb{R}$. Immediately, $\mathrm{N}_{(t,\infty)}(X) < \infty$ for any $t \in \mathbb{R}$ and there exists a particle $x_{\max}(X) \in X$ such that $x_{\max} > x$ for every particle $x \in X$. When $\mathbb{E}\left[\mathrm{N}_{(t,\infty)}(X)\right] < \infty$, we say that X has a last particle almost surely. The next proposition gives a sufficient condition for a point process to have a last particle, and again it depends on the correlation functions.

Proposition 4.4. Let X be a point process on $\mathbb{R}$ with n-th correlation functions $(\rho_n)_1^\infty$. If for each $t \in \mathbb{R}$,

$$\sum_{n=0}^{\infty} \frac{1}{n!} \int_{(t,\infty)^n} \rho_n(x_1,\ldots,x_n)\mu^{\otimes n}(\mathrm{d}^n x) < \infty,$$

then the process has a last particle almost surely and

$$\mathbb{P}(x_{\max}(X) \leq t) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int_{(t,\infty)^n} \rho_n(x_1,\ldots,x_n)\mu^{\otimes n}(\mathrm{d}^n x).$$

Proof. (BAIK; DEIFT; SUIDAN, 2016) Let $m > t$ be a natural number. Recall that

$$\mathbb{E}\left[\mathrm{N}_{(t,m)}(X)\right] = M_1((t,m)) = \int_{(t,m)} \rho_1(x)\mathrm{d}\mu(x),$$

so by continuity from below, $\mathbb{E}[\mathrm{N}_{(t,m)}(X)] \xrightarrow{m\to\infty} \mathbb{E}[\mathrm{N}_{(t,\infty)}(X)]$. Moreover, by Monotone Convergence Theorem,

$$\int_{(t,m)} \rho_1(x)\mathrm{d}\mu(x) \xrightarrow{m\to\infty} \int_{(t,\infty)} \rho_1(x)\mathrm{d}\mu(x),$$

hence, by uniqueness of limits and by hypothesis, $\mathbb{E}[\mathrm{N}_{(t,\infty)}(X)] < \infty$, so the process has a last particle almost surely. For the second claim, we have

$$\begin{aligned}
\mathbb{P}(x_{\max} \leq t) &= \mathbb{P}(\mathrm{N}_{(t,\infty)}(X) = 0) = \lim_{m\to\infty} \mathbb{P}(\mathrm{N}_{(t,m)}(X) = 0) \\
&= \lim_{m\to\infty} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int_{(t,m)^n} \rho_n(x_1,\ldots,x_n)\mu^{\otimes n}(\mathrm{d}^n x) \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int_{(t,\infty)^n} \rho_n(x_1,\ldots,x_n)\mu^{\otimes n}(\mathrm{d}^n x),
\end{aligned}$$

by the formula for gap probabilities (Example 4.5) and by Dominated Convergence Theorem together with the hypothesis. □

After all this theory, we can finally provide nice examples of point processes.

Example 4.6. (Symmetric probability densities)(BAIK; DEIFT; SUIDAN, 2016) Let $u(\mathrm{x}) = u(x_1,\ldots,x_N)$ be a continuous symmetric probability density on $\mathbb{R}^N$, that is, we have

$$\int_{\mathbb{R}^N} u(x_1,\ldots,x_N)\mathrm{dx} = 1$$

and for any Borel set B,

$$\mathbb{P}_u(B) = \int_B u(x_1,\ldots,x_N)\mathrm{dx}.$$

Consider the map defined by

$$\begin{aligned}
\Phi : \mathbb{R}^N &\to \mathrm{Conf}(\mathbb{R}) \\
(x_1,\ldots,x_N) &\mapsto \{x_1,\ldots,x_N\}.
\end{aligned}$$

If we assume Φ measurable, we can define a point process on $\mathbb{R}$ via

$$\mathbb{P}(B) := \mathbb{P}_u\left(\Phi^{-1}(B)\right)$$

and since $u(\mathrm{x})$ is continuous, we have $\mathbb{P}(x_i = x_j, i \neq j) = 0$, that is, the points of any configuration are pairwise distinct almost surely. For any bounded set B and $k \leq N$, we have

$$\mathbb{P}(\mathrm{N}_B(X) = k) = \binom{N}{k} \int_{B^k \times (B^{\mathrm{c}})^{N-k}} u(x_1, \ldots, x_N) \mathrm{d}x_1 \cdots \mathrm{d}x_N.$$

For any $\phi : \mathbb{R} \to \mathbb{R}$ with bounded support,

$$\begin{aligned}
\mathbb{E}\left[\prod_{i=1}^{N}(1+\phi(x_i))\right] &= \sum_{n=0}^{N} \frac{1}{n!} \mathbb{E}\left[\sum_{x_{i_1} \neq \cdots \neq x_{i_n}} \phi(x_{i_1}) \cdots \phi(x_{i_n})\right] \\
&= \sum_{n=0}^{N} \frac{1}{n!} \frac{N!}{(N-n)!} \int_{\mathbb{R}^N} \phi(x_1) \cdots \phi(x_n) u(x_1, \ldots, x_N) \mathrm{d}x_1 \cdots \mathrm{d}x_N \\
&= \sum_{n=0}^{N} \frac{1}{n!} \int_{\mathbb{R}^n} \phi(x_1) \cdots \phi(x_n) \rho_n(x_1, \ldots, x_n) \mathrm{d}x_1 \cdots \mathrm{d}x_n,
\end{aligned}$$

where

$$\rho_n(x_1, \ldots, x_n) = \frac{N!}{(N-n)!} \int_{\mathbb{R}^{N-n}} u(x_1, \ldots, x_N) \mathrm{d}x_{n+1} \cdots \mathrm{d}x_N,$$

which is the n-th correlation function, by Proposition 4.3.

Example 4.7. We presented in Chapter 2, the Poissonized Plancherel measure with parameter t, given by

$$\mathbb{P}_t^{\mathrm{Poisson}}(\lambda) = \frac{\mathrm{e}^{-t} |\lambda|^t \dim(\lambda)}{(|\lambda|!)^2}, \quad \lambda \in \mathbb{Y}.$$

Define the measurable map $\Phi : \mathbb{Y} \to \mathbf{Conf}(\mathbb{Z})$ by

$$\Phi(\lambda) := \{\lambda_i - i \mid i = 1, 2, \ldots, \ell(\lambda)\}.$$

Then, in analogy to the previous example, we define a probability measure on $\mathbf{Conf}(\mathbb{Z})$ by

$$\mathbb{P}(A) := \mathbb{P}_t^{\mathrm{Poisson}}(\Phi^{-1}(A)),$$

so we have a well-defined point process. Again, we notice that the particles are distinct, since $\lambda_i - i \neq \lambda_j - j$ when $i \neq j$. One can prove that the correlation functions for this process are given by

$$\rho_n(x_1, \ldots, x_n) = \det\left[K(x_i, x_j)\right]_{i,j=1}^{n}, \quad n \geq 1,$$

for a specific function $K : \mathbb{Z} \times \mathbb{Z} \to \mathbb{C}$ (see (BAIK; DEIFT; SUIDAN, 2016) for details).

Correlation functions of the form showed in Example 4.7 are also found in Random Matrix Theory, motivating the study of a special class of point processes, which we present now.

4.3. Determinantal Point Processes

We start defining a specific class of random point processes.

Definition 4.4. A point process on $\mathfrak{X}$ is called a **determinantal point process (DPP)** if there exists a function $K : \mathfrak{X} \times \mathfrak{X} \to \mathbb{C}$ such that the correlation functions (with respect to a reference measure) are given by

$$\rho_n(x_1, \dots, x_n) = \det\left[K(x_i, x_j)\right]_{i,j=1}^n$$

for all $n \in \mathbb{N}$. The function K is called **correlation kernel**.

Note 4.4. The correlation kernel of a determinantal point process is not unique. More precisely, if $K : \mathfrak{X} \times \mathfrak{X} \to \mathbb{C}$ is a correlation kernel for a DPP, then for any $f : \mathfrak{X} \to \mathbb{C} \setminus \{0\}$, the function

$$\widetilde{K}(x, y) \coloneqq \frac{f(x)}{f(y)} K(x, y)$$

is also a correlation kernel for the same point process. Indeed we have

$$\begin{aligned}
\rho_n(x_1, \dots, x_n) &= \det\left[K(x_i, x_j)\right]_{i,j=1}^n \\
&= \underbrace{\prod_{i=1}^n \frac{f(x_i)}{f(x_i)}}_{=1} \det\left[K(x_i, x_j)\right]_{i,j=1}^n \\
&= \det\left[\frac{f(x_i)}{f(x_j)} K(x_i, x_j)\right]_{i,j=1}^n \\
&= \det\left[\widetilde{K}(x_i, x_j)\right]_{i,j=1}^n,
\end{aligned}$$

so a DPP has infinitely many correlation kernels. The kernel $\widetilde{K}$ is called a **conjugation** of K.

Example 4.8. Remember the formula for gap probabilities showed in Example 4.5. For a DPP, we have

$$\begin{aligned}
\mathbb{P}(\mathrm{N}_B(X) = 0) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int_{B^n} \det\left[K(y_i, y_j)\right] \mu^{\otimes n}(\mathrm{d}^n y) \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int_{\mathfrak{X}^n} \mathbb{1}_B(y_1) \cdots \mathbb{1}_B(y_n) \det\left[K(y_i, y_j)\right] \mu^{\otimes n}(\mathrm{d}^n y) \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int_{\mathfrak{X}^n} \mathbb{1}_B^2(y_1) \cdots \mathbb{1}_B^2(y_n) \det\left[K(y_i, y_j)\right] \mu^{\otimes n}(\mathrm{d}^n y) \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int_{\mathfrak{X}^n} \det\left[\mathbb{1}_B(y_i) K(y_i, y_j) \mathbb{1}_B(y_j)\right] \mu^{\otimes n}(\mathrm{d}^n y).
\end{aligned}$$

We may want to rescale the positions of particles in point processes in the future. The next proposition suits this purpose.

Proposition 4.5. Let $\mathfrak{X} \subset \mathbb{R}$ with Lebesgue measure μ, and consider a point process $\mathbb{P}$ in $\mathfrak{X}$ with configurations X. Define the scaled point process $\widehat{\mathbb{P}}$ on the space $\widehat{\mathfrak{X}} = \frac{1}{b}(\mathfrak{X} - a)$ with scaled configurations $\widehat{X} = \frac{1}{b}(X - a)$, for $a \in \mathbb{R}$ and $b \neq 0$. If $\{\rho_n\}$ are the correlation functions for $\mathbb{P}$, then the correlation functions for $\widehat{\mathbb{P}}$ are given by

$$\widehat{\rho}_n(\widehat{x}_1, \ldots, \widehat{x}_n) = |b|^n \rho_n(b\widehat{x}_1 + a, \ldots, b\widehat{x}_n + a), \quad n \in \mathbb{N}, \quad \widehat{x}_1, \ldots, \widehat{x}_n \in \widehat{\mathfrak{X}}.$$

Proof. Let M_n and $\widehat{M}_n$ be the n-th correlation measures for the processes $\mathbb{P}$ and $\widehat{\mathbb{P}}$, respectively. For a bounded Borel set $B \in \widehat{\mathfrak{X}}^n$, define the characteristic functions $f = \chi_B$ and $g = \chi_{bB+a}$, then by definition we have

$$\begin{aligned}
\widehat{M}_n(B) = \int_{\widehat{\mathfrak{X}}^n} f \mathrm{d}\widehat{M}_n &= \mathbb{E}\left[\sum_{\widehat{x}_1 \neq \cdots \neq \widehat{x}_n} f(\widehat{x}_1, \ldots, \widehat{x}_n)\right] \\
&= \mathbb{E}\left[\sum_{x_1 \neq \cdots \neq x_n} f\left(\frac{1}{b}(x_1 - a), \ldots, \frac{1}{b}(x_n - a)\right)\right] \\
&= \mathbb{E}\left[\sum_{x_1 \neq \cdots \neq x_n} g(x_1, \ldots, x_n)\right] \\
&= \int_{\mathfrak{X}^n} g \mathrm{d}M_n = M_n(bB + a).
\end{aligned}$$

So by definition of correlation functions,

$$\widehat{M}_n(B) = M_n(bB + a) = \int_{bB+a} \rho_n(x_1, \ldots, x_n) \mathrm{d}\mu(\mathrm{x}),$$

and changing the variables to $x_i = b\widehat{x}_i + a$, for $i = 1, \ldots, n$, we obtain $\widehat{x}_i \in B$ and the Jacobian is $|b|^n$, so

$$\widehat{M}_n(B) = \int_B |b|^n \rho_n(b\widehat{x}_1 + a, \ldots, b\widehat{x}_n + a) \mathrm{d}\mu(\widehat{\mathrm{x}})$$

as we wanted to prove. □

Corollary 4.4. In addition to the assumptions of Proposition 4.5, suppose that $\mathbb{P}$ is determinantal with correlation kernel $K(x, y)$, then the process $\widehat{\mathbb{P}}$ is also determinantal and its correlation kernel is $\widehat{K}(\widehat{x}, \widehat{y}) = |b| K(b\widehat{x} + a, b\widehat{y} + a)$.

Proof. Notice that

$$|b|^n \rho_n(b\widehat{x}_1 + a, \ldots, b\widehat{x}_n + a) = \det\left[|b| K(b\widehat{x}_i + a, b\widehat{x}_j + a)\right]_{i,j=1}^n$$

and apply Proposition 4.5. □

To end the chapter, we discuss some theoretical aspects of DPPs. As we have already commented, Lenard found conditions for a point process to be determined by its correlation functions. Soshnikov (2000) did a similar work in the context of DPPs, finding

conditions for a function $K : \mathfrak{X} \times \mathfrak{X} \to \mathbb{C}$ to induce a DPP on $\mathfrak{X}$. These results appear in (BAIK; DEIFT; SUIDAN, 2016), where we also see some examples. We do not present the technical details here (in fact, we present later a simpler way to generate DPPs that will be more useful for us).

We are interested in the answer of the following question: under what conditions a function $K(x,y)$ does define a DPP? This question is important for this work. To illustrate, imagine we have a sequence $\mathbb{P}_N$ of DPPs with correlation kernels K_N, respectively. We would like to know if this sequence converges to a point process $\mathbb{P}$ with correlation kernel K, but we do not have any precise definition for convergence of point processes. If we manage to prove that some correlation kernels are associated to unique DPPs, then we may say that $\mathbb{P}_N \to \mathbb{P}$ in the sense of the pointwise convergence $K_N(x,y) \to K(x,y)$. In this regard, Soshnikov proved the following result.

Theorem 4.2. (SOSHNIKOV, 2000) Let $(\mathfrak{X}, \mu)$ be a measure space. A Hermitian locally trace-class operator $K \in L^2(\mathfrak{X}, \mu)$ defines a determinantal point process with correlation kernel K if, and only if, $0 \leq K \leq 1$. Moreover, the corresponding point process is unique.

For the convenience of the reader, we will break the conditions in Theorem 4.2 into parts. First, remember that

$$L^2(\mathfrak{X}, \mu) = \left\{ f : \mathfrak{X} \to \mathbb{C} \mid f \text{ is measurable}, \int_{\mathfrak{X}} |f(x)|^2 \, \mathrm{d}\mu(x) < \infty \right\}$$

is equipped with the inner product

$$\langle f, g \rangle = \int_{\mathfrak{X}} f(x) \overline{g(x)} \mathrm{d}\mu(x),$$

so that $\|f\|_2 = \sqrt{\langle f, f \rangle}$ is a norm. A correlation function $K : \mathfrak{X} \times \mathfrak{X} \to \mathbb{C}$ defines an operator $K : L^2(\mathfrak{X}, \mu) \to L^2(\mathfrak{X}, \mu)$ by

$$Kf(x) = \int_{\mathfrak{X}} K(x,y) f(y) \mathrm{d}\mu(y),$$

provided that

$$\int_{\mathfrak{X}} \int_{\mathfrak{X}} |K(x,y)|^2 \, \mathrm{d}\mu(x) \mathrm{d}\mu(y) < \infty.$$

The well-definition of the operator K follows by Fubini's theorem and Cauchy-Schwarz inequality, since L^2 is a Hilbert space. We say that K is **Hermitian** if $K(x,y) = \overline{K(y,x)}$ for any $x, y \in \mathfrak{X}$. The operator K is called **bounded** if there exists $C > 0$ such that $\|Kf\|_2 \leq C\|f\|_2$ for every f. A trace-class operator is a special class of bounded operators that we can obtain by the following lemma (that we can use as definition) (BAIK, 2018; BAIK; DEIFT; SUIDAN, 2016).

Lemma 4.1. Let $(\mathfrak{Y}, \nu)$ be a measure space and assume

$$K(x,y) = \int_{\mathfrak{Y}} L(x,z)R(z,y)\mathrm{d}\nu(z), \quad x,y \in \mathfrak{X},$$

for some $L : \mathfrak{X} \times \mathfrak{Y} \to \mathbb{C}$ and $R : \mathfrak{Y} \times \mathfrak{X} \to \mathbb{C}$ satisfying

$$\int_{\mathfrak{Y}} \int_{\mathfrak{X}} |L(x,y)|^2 \mathrm{d}\mu(x)\mathrm{d}\nu(y) < \infty, \quad \int_{\mathfrak{X}} \int_{\mathfrak{Y}} |R(x,y)|^2 \mathrm{d}\nu(x)\mathrm{d}\mu(y) < \infty,$$

then K is a **trace-class operator** on $L^2(\mathfrak{X}, \mu)$. If the conditions above hold with $\mathfrak{X}$ replaced by any compact B of $\mathfrak{X}$, then K is a **locally trace-class operator**.

The condition $0 \leq K \leq 1$ means that

$$0 \leq \langle Kf, f \rangle \leq \|f\|_2^2$$

for every compactly supported smooth function $f : \mathfrak{X} \to \mathbb{C}$. We will use Theorem 4.2 later to show that the *Airy point process* is a DPP.

4.3.1. The Airy Point Process

We now discuss a point process that deserves special attention in this work, the *Airy point process*. The basic object for the construction of the Airy point process is the *Airy function*, so we start discussing it.

Definition 4.5. The **Airy function** $\mathrm{Ai} : \mathbb{R} \to \mathbb{R}$ is defined by the integral

$$\mathrm{Ai}(x) = \frac{1}{2\pi} \int_{\gamma} \mathrm{e}^{i\left(\frac{z^3}{3} + xz\right)} \mathrm{d}z,$$

where the curve γ starts at $\infty \mathrm{e}^{i\theta_1}$ and ends at $\infty \mathrm{e}^{i\theta_2}$, for $\theta_1 \in \left(\frac{2\pi}{3}, \pi\right)$ and $\theta_2 \in \left(0, \frac{\pi}{3}\right)$ (see Figure 22 for the contour).

See Figure 27 for the graph of $\mathrm{Ai}(x)$. The notation $\infty \mathrm{e}^{i\theta}$ represents a point at infinite with argument θ. Let

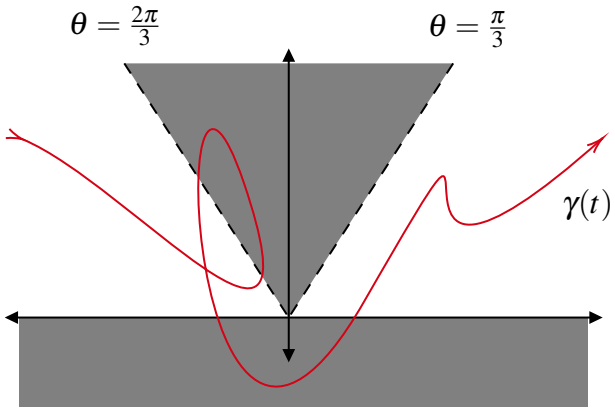
$$f(z) := \mathrm{e}^{i\left(\frac{z^3}{3} + xz\right)},$$

to ensure that $\mathrm{Ai}(x)$ is finite, we need $|f(z)|$ small when $|z|$ is large, so the real part of the exponent must be (very) negative. Write $z = R\mathrm{e}^{i\theta}$, so that $z^3 = R^3 \mathrm{e}^{3i\theta}$, we obtain

$$\Re\left[i\left(\frac{z^3}{3} + xz\right)\right] = -\frac{R^3 \sin(3\theta)}{3} - xR\sin\theta.$$

Since the cubic term R^3 dominates the expression as $|z| = R \to \infty$, we need $\sin(3\theta) > 0$. This is satisfied by our choice of θ_1 and θ_2, so we have exponential decay on the extremes of γ, hence the integral is finite.

Figure 22 – Airy function contour. As $t \to \pm\infty$ the curve must not touch the gray area.



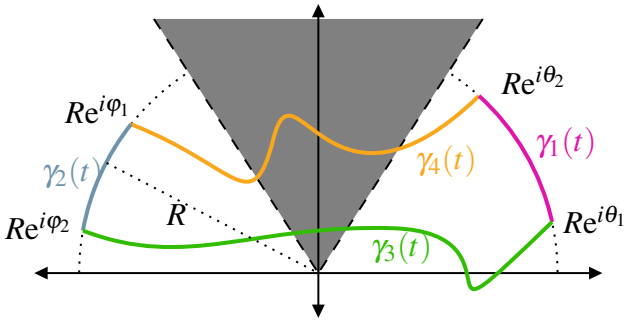
Source: Elaborated by the author.

We also have to guarantee that the choice of γ does not matter, as long as the angles in Definition 4.5 are satisfied. To show this, consider $\theta_1, \theta_2 \in (0, \frac{\pi}{3})$ and $\varphi_1, \varphi_2 \in (\frac{2\pi}{3}, \pi)$ such that $\theta_1 < \theta_2$ and $\varphi_1 < \varphi_2$. For fixed $R > 0$, take the curves

$$\begin{aligned}
\gamma_1(t) &= Re^{it}, \quad t \in [\theta_1, \theta_2]; \\
\gamma_2(t) &= Re^{it}, \quad t \in [\varphi_1, \varphi_2]; \\
\gamma_3(t) &= \text{any smooth curve connecting the points } Re^{i\theta_1} \text{ and } Re^{i\varphi_2}; \\
\gamma_4(t) &= \text{any smooth curve connecting the points } Re^{i\varphi_1} \text{ and } Re^{i\theta_2}
\end{aligned}$$

and consider the juxtaposition γ of all of them, oriented counter-clockwise (Figure 23).

Figure 23 – γ contour.



Source: Elaborated by the author.

We claim that the integral of $f(z)$ on the curves γ_1 and γ_2 vanishes as $R \to \infty$. Let us show it for γ_1. The technique is to find $c > 0$ and $R_0 > 0$, such that

$$\frac{R^3}{3}\sin(3t) + xR\sin(t) > R^3 c$$

for every $t \in [\theta_1, \theta_2]$ and $R > R_0$. First, notice that there exists a constant $d > 0$ such that

$$\forall t \in [\theta_1, \theta_2], \quad \sin t > d > 0 \quad \text{and} \quad \sin(3t) > d > 0.$$

Also, we have $\frac{xd}{R^2} \to 0$ when $R \to \infty$, so there exists $R_0 > 0$ such that

$$\forall R \geq R_0, \quad -\frac{d}{6} < \frac{xd}{R^2} < \frac{d}{6}.$$

So when $R \geq R_0$ and $t \in [\theta_1, \theta_2]$, we have

$$\begin{aligned}\frac{R^3}{3}\sin(3t) + xR\sin(t) &> \frac{R^3}{3}d + xRd = R^3\left(\frac{d}{3} + \frac{xd}{R^2}\right) \\ &> R^3\left(\frac{d}{3} - \frac{d}{6}\right) = R^3\frac{d}{6},\end{aligned}$$

thus we get the result by taking $c = \frac{d}{6}$. Now we have for z on the curve $\gamma_1(t)$ and $R \geq R_0$,

$$|f(z)| = \left|\mathrm{e}^{i\left(\frac{z^3}{3} + xz\right)}\right| = \mathrm{e}^{\Re\left(i\left(\frac{z^3}{3} + xz\right)\right)} = \mathrm{e}^{-\frac{R^3\sin(3t)}{3} - xR\sin t} < \mathrm{e}^{-R^3 c},$$

so

$$\left|\int_{\gamma_1} f(z)\mathrm{d}z\right| \leq \sup_{z=\gamma_1(t)} |f(z)|\,\ell(\gamma_1) \leq \mathrm{e}^{-R^3 c}(\theta_2 - \theta_1)R \xrightarrow{R\to\infty} 0,$$

where $\ell(\gamma_1)$ denotes the arc length of γ_1. The same happens with the curve γ_2. Now, since $f(z)$ is an entire function, we have by Cauchy Theorem that

$$\int_{\gamma} f(z) = \int_{\gamma_1} f(z)\mathrm{d}z + \int_{\gamma_2} f(z)\mathrm{d}z + \int_{\gamma_3} f(z)\mathrm{d}z + \int_{\gamma_4} f(z)\mathrm{d}z = 0,$$

and when $R \to \infty$, the integrals in γ_1 and γ_2 vanish, so if the integrals in γ_3 and γ_4 exist, they satisfy

$$\int_{\gamma_3} f(z)\mathrm{d}z = \int_{\gamma_4^{-1}} f(z)\mathrm{d}z,$$

where γ_4^{-1} is the curve γ_4 oriented in the opposite direction. After all these considerations, we conclude that the definition of the Airy function is well posed. It is also possible to prove that $\mathrm{Ai}(x)$ is real-valued.

Proposition 4.6. The Airy function satisfies $\mathrm{Ai}''(x) = x\,\mathrm{Ai}(x)$.

Proof. By Leibniz rule, we can differentiate with respect to x under the integral sign. We have

$$\mathrm{Ai}''(x) = \frac{1}{2\pi}\int_\gamma -z^2 \mathrm{e}^{i\left(\frac{z^3}{3}+xz\right)}\mathrm{d}z,$$

then

$$\begin{aligned} x\,\mathrm{Ai}(x) - \mathrm{Ai}''(x) &= \frac{1}{2\pi}\int_\gamma \left(x+z^2\right)\mathrm{e}^{i\left(\frac{z^3}{3}+xz\right)}\mathrm{d}z \\ &= -\frac{i}{2\pi}\int_\gamma i\left(x+z^2\right)\mathrm{e}^{i\left(\frac{z^3}{3}+xz\right)}\mathrm{d}z \\ &= -\frac{i}{2\pi}\int_\gamma \frac{\mathrm{d}}{\mathrm{d}z}\mathrm{e}^{i\left(\frac{z^3}{3}+xz\right)}\mathrm{d}z \\ &= -\lim_{a,b\to\infty}\frac{i}{2\pi}\mathrm{e}^{i\left(\frac{z^3}{3}+xz\right)}\Big|_{z=\gamma(-a)}^{\gamma(b)} = 0. \end{aligned}$$

The limit above is zero due to the choice of γ in Definition 4.5. □

Definition 4.6. The **Airy kernel** is the function $\mathbb{A}:\mathbb{R}^2\to\mathbb{R}$ given by

$$\mathbb{A}(x,y) = \frac{\mathrm{Ai}(x)\,\mathrm{Ai}'(y) - \mathrm{Ai}'(x)\,\mathrm{Ai}(y)}{x-y}.$$

Proposition 4.7. The Airy kernel can be expressed as

$$\mathbb{A}(x,y) = \int_0^\infty \mathrm{Ai}(x+t)\,\mathrm{Ai}(t+y)\mathrm{d}t.$$

Proof. By a direct computation and Proposition 4.6, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\mathrm{Ai}(x+t)\,\mathrm{Ai}'(t+y) - \mathrm{Ai}'(x+t)\,\mathrm{Ai}(t+y)}{y-x}\right] = \mathrm{Ai}(x+t)\,\mathrm{Ai}(t+y),$$

so by the Fundamental Theorem of Calculus,

$$\begin{aligned} &\int_0^\infty \mathrm{Ai}(x+t)\,\mathrm{Ai}(t+y)\mathrm{d}t \\ &\qquad = \lim_{b\to\infty}\frac{\mathrm{Ai}(x+b)\,\mathrm{Ai}'(b+y) - \mathrm{Ai}'(x+b)\,\mathrm{Ai}'(b+y)}{y-x} - \frac{\mathrm{Ai}(x)\,\mathrm{Ai}'(y) - \mathrm{Ai}'(x)\,\mathrm{Ai}(y)}{y-x}. \end{aligned}$$

Since $\mathrm{Ai}(s)$ and $\mathrm{Ai}'(s)$ decay as $s\to\infty$, the first parcel above is zero, hence

$$\int_0^\infty \mathrm{Ai}(x+t)\,\mathrm{Ai}(t+y)\mathrm{d}t = \frac{\mathrm{Ai}(x)\,\mathrm{Ai}'(y) - \mathrm{Ai}'(x)\,\mathrm{Ai}'(y)}{x-y} = \mathbb{A}(x,y).$$

□

Proposition 4.8. We have

$$\mathbb{A}(x,y) = \frac{i}{(2\pi)^2}\int_{\gamma_w}\int_{\gamma_z}\frac{1}{z-w}\frac{\mathrm{e}^{i\left(\frac{z^3}{3}+xz\right)}}{\mathrm{e}^{i\left(\frac{w^3}{3}+yw\right)}}\mathrm{d}z\mathrm{d}w,$$

where the γ_z is the same contour as for the Airy function and the γ_w is the reflection of γ_z on the real axis.

Proof. By the change of variables $x \mapsto x+t$ and $y \mapsto y+t$, we have

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\frac{i}{(2\pi)^2}\int\int \frac{\exp\left(i\left(\frac{z^3}{3}+(x+t)z-\frac{w^3}{3}-(y+t)w\right)\right)}{z-w}\mathrm{d}z\mathrm{d}w \\
&\qquad = \frac{i}{(2\pi)^2}\int\int i(z-w)\frac{\exp\left(i\left(\frac{z^3}{3}+(x+t)z-\frac{w^3}{3}-(y+t)w\right)\right)}{z-w}\mathrm{d}z\mathrm{d}w \\
&\qquad = -\frac{1}{(2\pi)^2}\int \exp\left(-i\left(\frac{w^3}{3}+(y+t)w\right)\right)\mathrm{d}w\int \exp\left(i\left(\frac{z^3}{3}+(x+t)z\right)\right)\mathrm{d}z.
\end{aligned}$$

Also, changing $w \mapsto -w$, the integral in w is the same that defines the Airy function, so the last expression is equal to

$$\frac{-1}{(2\pi)^2}(2\pi)^2\,\mathrm{Ai}(y+t)\,\mathrm{Ai}(x+t) = -\,\mathrm{Ai}(x+t)\,\mathrm{Ai}(y+t).$$

Finally, we apply the Fundamental Theorem of Calculus to obtain that

$$\frac{i}{(2\pi)^2}\int_{\gamma_w}\int_{\gamma_z}\frac{1}{z-w}\frac{\mathrm{e}^{i\left(\frac{z^3}{3}+xz\right)}}{\mathrm{e}^{i\left(\frac{w^3}{3}+yw\right)}}\mathrm{d}z\mathrm{d}w = \int_0^\infty \mathrm{Ai}(x+t)\,\mathrm{Ai}(y+t)\mathrm{d}t, \tag{4.5}$$

using the fact that the Airy function decays exponentially in the given contours. By Proposition 4.7, this is $\mathbb{A}(x,y)$. □

Proposition 4.9. The Airy kernel defines a unique determinantal point process in $\mathbb{R}$.

Proof. (BAIK, 2018) It suffices to check the conditions of Theorem 4.2. Since $\mathbb{A}(x,y)$ is real-valued, we have $\mathbb{A}(x,y) = \mathbb{A}(y,x) = \overline{\mathbb{A}(y,x)}$, by Definition 4.6, so $\mathbb{A}$ is Hermitian. Motivated by Proposition 4.7, we define $L : [a,\infty)\times[0,\infty) \to \mathbb{R}$ and $R : [0,\infty)\times[a,\infty) \to \mathbb{R}$ by

$$L(x,t) := \mathrm{Ai}(x+t), \quad R(t,y) := \mathrm{Ai}(t+y).$$

By the asymptotic expansion for $\mathrm{Ai}(x)$ (see Section 5.2.2, Example 5.3), we can guarantee that the integrals

$$\int_0^\infty\int_a^\infty |L(x,t)|^2\,\mathrm{d}x\mathrm{d}t, \quad \int_a^\infty\int_0^\infty |R(t,y)|^2\,\mathrm{d}t\mathrm{d}y$$

are finite, since we have exponential decay in the integrands, as $x,y,t \to \infty$. So by Lemma 4.1, $\mathbb{A}$ is a locally trace-class operator. The proof that $0 \le \mathbb{A} \le 1$ is more technical and we do not show here. □

4.3.2. Biorthogonal Ensembles

In this section, we discuss a special class of determinantal point processes.

Definition 4.7. Let $(\mathfrak{X}, \mu)$ be a measure space and consider measurable functions $\phi_1, \ldots, \phi_N, \psi_1, \ldots, \psi_N : \mathfrak{X} \to \mathbb{C}$ such that $\int_{\mathfrak{X}} \phi_i(x)\psi_j(x)\mathrm{d}\mu(x) < \infty$ for every pair i, j. Also, assume that

$$Z = \int_{\mathfrak{X}^N} \det\left[\phi_i(x_j)\right]_{i,j=1}^N \det\left[\psi_i(x_j)\right]_{i,j=1}^N \mathrm{d}\mu(x_1)\cdots\mathrm{d}\mu(x_N) \neq 0.$$

The point process on $\mathfrak{X}$ induced by the probability density

$$u(x_1,\ldots,x_N) = \frac{1}{Z}\det\left[\phi_i(x_j)\right]_{i,j=1}^N \det\left[\psi_i(x_j)\right]_{i,j=1}^N$$

is called a **biorthogonal ensemble**.

Note 4.5. The density $u(x_1,\ldots,x_N)$ is symmetric, since permutations of variables are permutations of columns in the matrices. Therefore, by Example 4.6, $u(x_1,\ldots,x_N)$ indeed induces a point process on $\mathfrak{X}$.

Given a biorthogonal ensemble, in the notation of Definition 4.7, the **Gram matrix** associated is the matrix

$$G = [G_{ij}]_{i,j=1}^N, \quad G_{ij} \coloneqq \int_{\mathfrak{X}} \phi_i(x)\psi_j(x)\mathrm{d}\mu(x). \tag{4.6}$$

Proposition 4.10 (Andréief's integration formula)**.** We have

$$\det\left[\int_{\mathfrak{X}} \phi_i(x)\psi_j(x)\mathrm{d}\mu(x)\right] = \frac{1}{N!}\int_{\mathfrak{X}^N} \det\left[\phi_i(x_j)\right]\det\left[\psi_i(x_j)\right]\mathrm{d}\mu(x_1)\cdots\mathrm{d}\mu(x_N). \tag{4.7}$$

Proof. (BAIK, 2018) Remember the property for multiple integrals of a product of functions depending on distinct variables:

$$\int\int f(x)g(y)\mathrm{d}\mu(x)\mathrm{d}\mu(y) = \int f(x)\mathrm{d}\mu(x)\int g(y)\mathrm{d}\mu(y). \tag{4.8}$$

By definition of determinant on $\left[\psi_i(x_j)\right]$, the right-hand side integral on (4.7) is

$$\begin{aligned}
&\int_{\mathfrak{X}^N} \det\left[\phi_i(x_j)\right]\left(\sum_{\sigma}\mathrm{sgn}(\sigma)\prod_{j=1}^N \psi_{\sigma(j)}(x_j)\right)\mathrm{d}\mu(x_1)\cdots\mathrm{d}\mu(x_N)\\
&= \sum_{\sigma}\mathrm{sgn}(\sigma)\int_{\mathfrak{X}^N} \det\left[\phi_i(x_j)\right]\left(\prod_{j=1}^N \psi_{\sigma(j)}(x_j)\right)\mathrm{d}\mu(x_1)\cdots\mathrm{d}\mu(x_N)\\
&= \sum_{\sigma}\mathrm{sgn}(\sigma)\int_{\mathfrak{X}^N} \det\left[\phi_i(x_j)\psi_{\sigma(j)}(x_j)\right]\mathrm{d}\mu(x_1)\cdots\mathrm{d}\mu(x_N).
\end{aligned}$$

Since the integrand is a determinant (i.e. a sum of products of functions depending on distinct variables), we apply property (4.8) and multilinearity of determinant to see that

this expression is equal to

$$\begin{aligned}
&\sum_{\sigma} \operatorname{sgn}(\sigma) \det\left[\int_{\mathfrak{X}} \phi_i(x)\psi_{\sigma(j)}(x)\mathrm{d}\mu(x)\right] \\
&= \sum_{\sigma} \operatorname{sgn}(\sigma)\operatorname{sgn}(\sigma^{-1}) \det\left[\int_{\mathfrak{X}} \phi_i(x)\psi_j(x)\mathrm{d}\mu(x)\right] \\
&= N! \det\left[\int_{\mathfrak{X}} \phi_i(x)\psi_j(x)\mathrm{d}\mu(x)\right]
\end{aligned}$$

as we wanted to prove. □

Corollary 4.5. The Gram matrix G associated to a biorthogonal ensemble is invertible.

Proof. By Andréief's formula,

$$\begin{aligned}
0 \neq Z &= \int_{\mathfrak{X}^N} \det\left[\phi_i(x_j)\right] \det\left[\psi_i(x_j)\right] \mathrm{d}\mu(x_1)\cdots\mathrm{d}\mu(x_N) \\
&= N! \det\left[\int_{\mathfrak{X}} \phi_i(x)\psi_j(x)\mathrm{d}\mu(x)\right] = N! \det G,
\end{aligned}$$

hence $\det G \neq 0$. □

Note 4.6. If the sets of functions $\{\phi_i\}$ and $\{\psi_i\}$ form a biorthogonal system, in the sense that $\int_{\mathfrak{X}} \phi_i(x)\psi_j(x)\mathrm{d}\mu(x) = c_{ij}\delta_{ij}$, then the Gram matrix is easy to invert. That is the reason for the term *biorthogonal ensemble*.

Example 4.9. (Generalized Wishart Ensemble) Consider two sets of pairwise distinct real numbers $a_1,\ldots,a_N$ and $b_1,\ldots,b_N$ such that $a_i + b_j > 0$ for every pair $i,j = 1,\ldots,N$. A simple computation shows that the functions $\phi_i, \psi_i : \mathbb{R}_+ \to \mathbb{R}$ given by

$$\phi_i(x) := \mathrm{e}^{-a_i x}, \quad \psi_i(x) := \mathrm{e}^{-b_i x},$$

for $i = 1,\ldots,N$, satisfy

$$\int_0^{\infty} \phi_i(x)\psi_j(x)\mathrm{d}x = \frac{1}{a_i + b_j} < \infty.$$

The formula

$$\det\left[\frac{1}{a_i + b_j}\right] = \frac{\prod_{k<\ell}(a_\ell - a_k)(b_\ell - b_k)}{\prod_{k,\ell}(a_k + b_\ell)} \neq 0 \tag{4.9}$$

holds and it is known as the **Cauchy determinant**. By Andréief's integration formula (Proposition 4.10) and by (4.9),

$$\begin{aligned}
\int_0^{\infty}\cdots\int_0^{\infty} \det\left[\phi_i(x_j)\right] \det\left[\psi_i(x_j)\right] \mathrm{d}x_1\cdots\mathrm{d}x_N &= N! \det\left[\int_0^{\infty} \phi_i(x)\psi_j(x)\mathrm{d}x\right] \\
&= N! \det\left[\frac{1}{a_i + b_j}\right] \neq 0,
\end{aligned}$$

so we have a biorthogonal ensemble.

The next goal is to prove that the point process induced by a biorthogonal ensemble is determinantal. For this purpose, we need the following technical lemma.

Lemma 4.2 (Reproducing kernel lemma)**.** Let $L : \mathfrak{X} \times \mathfrak{X} \to \mathbb{C}$ be a function such that

$$\forall x, z \in \mathfrak{X}, \quad \int_{\mathfrak{X}} L(x,y)L(y,z)\mathrm{d}\mu(y) = L(x,z) \tag{4.10}$$

and

$$\int_{\mathfrak{X}} L(x,x)\mathrm{d}\mu(x) = k, \tag{4.11}$$

then for every $n > 1$,

$$\int_{\mathfrak{X}} \det\left[L(x_i,x_j)\right]_{i,j=1}^{n} \mathrm{d}\mu(x_n) = (k-n+1)\det\left[L(x_i,x_j)\right]_{i,j=1}^{n-1}.$$

Proof. (BAIK, 2018) We are proving for $n = 3$, since the general case follows the same idea. Expanding the determinant of $\left[L(x_i,x_j)\right]_{i,j=1}^{N}$ along the last row, we obtain

$$\begin{aligned}\det\left[L(x_i,x_j)\right]_{i,j=1}^{N} = L(x_3,x_1)&\begin{vmatrix} L(x_1,x_2) & L(x_1,x_3) \\ L(x_2,x_2) & L(x_2,x_3)\end{vmatrix} \\ &- L(x_3,x_2)\begin{vmatrix} L(x_1,x_1) & L(x_1,x_3) \\ L(x_2,x_1) & L(x_2,x_3)\end{vmatrix} + L(x_3,x_3)\begin{vmatrix} L(x_1,x_1) & L(x_1,x_2) \\ L(x_2,x_1) & L(x_2,x_2)\end{vmatrix}.\end{aligned}$$

Multiplying the constant outside the first two parcels in the last column of the matrices, we have

$$\begin{aligned}\det\left[L(x_i,x_j)\right]_{i,j=1}^{N} = &\begin{vmatrix} L(x_1,x_2) & L(x_1,x_3)L(x_3,x_1) \\ L(x_2,x_2) & L(x_2,x_3)L(x_3,x_1)\end{vmatrix} \\ &- \begin{vmatrix} L(x_1,x_1) & L(x_1,x_3)L(x_3,x_2) \\ L(x_2,x_1) & L(x_2,x_3)L(x_3,x_2)\end{vmatrix} + L(x_3,x_3)\begin{vmatrix} L(x_1,x_1) & L(x_1,x_2) \\ L(x_2,x_1) & L(x_2,x_2)\end{vmatrix}.\end{aligned}$$

Integrating with respect to x_3, we use linearity and hypothesis (4.10) and (4.11) to obtain that

$$\begin{aligned}\int_{\mathfrak{X}} \det\left[L(x_i,x_j)\right]_{i,j=1}^{N} \mathrm{d}\mu(x_3) = &\begin{vmatrix} L(x_1,x_2) & L(x_1,x_1) \\ L(x_2,x_2) & L(x_2,x_1)\end{vmatrix} \\ &- \begin{vmatrix} L(x_1,x_1) & L(x_1,x_2) \\ L(x_2,x_1) & L(x_2,x_2)\end{vmatrix} + k\det\left[L(x_i,x_j)\right]_{i,j=1}^{2}\end{aligned}$$

and this is equal to

$$k\det\left[L(x_i,x_j)\right]_{i,j=1}^{2} - 2\det\left[L(x_i,x_j)\right]_{i,j=1}^{2},$$

as we wanted to prove. □

Theorem 4.3. A biorthogonal ensemble is a determinantal point process. More precisely, the function $K : \mathfrak{X} \times \mathfrak{X} \to \mathbb{C}$ defined by

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

is the correlation kernel of the point process, where G is the Gram matrix associated to this ensemble.

Proof. (BAIK, 2018) By Corollary 4.5, we have $\det G \neq 0$ and $Z = N! \det G$, so we can write

$$\begin{aligned} u(x_1,\ldots,x_N) &= \frac{\det\left[\phi_i(x_j)\right] \det\left[\psi_i(x_j)\right]}{N! \det G} \\ &= \frac{\det\left[\phi_i(x_j)\right] \det\left(G^{-1}\right) \det\left[\psi_i(x_j)\right]}{N!} \\ &= \frac{\det\left[K(x_i,x_j)\right]}{N!} \end{aligned} \tag{4.12}$$

The last equality above is due to the fact that

$$\left[K(x_i,x_j)\right] = \left[\phi_i(x_j)\right]^T G^{-T} \left[\psi_i(x_j)\right],$$

simply by definition of multiplication of matrices. Remember from Example 4.6 that the correlation functions for a symmetric density function is

$$\rho_n(x_1,\ldots,x_N) = \frac{N!}{(N-n)!} \int_{\mathfrak{X}^{N-n}} u(x_1,\ldots,x_N) \mathrm{d}\mu\left(x_{n+1}\right) \cdots \mathrm{d}\mu\left(x_N\right),$$

for $n = 1,2,3,\ldots$. By (4.12),

$$\rho_n(x_1,\ldots,x_N) = \frac{1}{(N-n)!} \int_{\mathfrak{X}^{N-n}} \det\left[K(x_i,x_j)\right] \mathrm{d}\mu\left(x_{n+1}\right) \cdots \mathrm{d}\mu\left(x_N\right). \tag{4.13}$$

For each $i = 1,\ldots,N$, define $\Psi_i(x) = \sum_{j=1}^{N} \left(G^{-T}\right)_{ij} \psi_j(x)$, so that $K(x,y) = \sum_{i=1}^{N} \phi_i(x) \Psi_i(y)$. Since $G^{-T} G^T = \mathrm{Id}$,

$$\begin{aligned} \delta_{ij} &= \sum_{\ell=1}^{N} \left(G^{-T}\right)_{i\ell} \left(G^T\right)_{\ell j} \\ &= \sum_{\ell=1}^{N} \left(G^{-T}_{i\ell}\right) \int_{\mathfrak{X}} \phi_j(x) \psi_\ell(x) \mathrm{d}\mu(x) \\ &= \int_{\mathfrak{X}} \phi_j(x) \Psi_i(x) \mathrm{d}\mu(x), \end{aligned}$$

thus the sets of functions $\{\phi_i\}$ and $\{\Psi_i\}$ form a biorthonormal system. From this, we have

$$\int_{\mathfrak{X}} K(x,x) \mathrm{d}\mu(x) = N \tag{4.14}$$

and

$$\int_{\mathfrak{X}} K(x,y)\,K(y,z)\,\mathrm{d}\mu(y) = \sum_{i,j=1}^{N} \phi_i(x)\Psi_j(z) \underbrace{\int_{\mathfrak{X}} \phi_j(y)\,\Psi_i(y)\,\mathrm{d}\mu(y)}_{=\delta_{ij}}$$

$$= \sum_{i=1}^{N} \phi_i(x)\,\Psi_i(z) = K(x,z)\,. \tag{4.15}$$

Equations (4.14) and (4.15) are exactly the hypotheses for Lemma 4.2, therefore we have

$$\int_{\mathfrak{X}} \det\left[K\left(x_i,x_j\right)\right]_{i,j=1}^{N} \mathrm{d}\mu\left(x_N\right) = \det\left[K\left(x_i,x_j\right)\right]_{i,j=1}^{N-1}.$$

We can apply the lemma again, integrating with respect to x_{N-1}, obtaining

$$\int_{\mathfrak{X}^2} \det\left[K\left(x_i,x_j\right)\right]_{i,j=1}^{N} \mathrm{d}\mu\left(x_{N-1}\right)\mathrm{d}\mu\left(x_N\right) = 2\det\left[K\left(x_i,x_j\right)\right]_{i,j=1}^{N-2}.$$

Applying the lemma $N-n$ times, we obtain

$$\int_{\mathfrak{X}^{N-n}} \det\left[K(x_i,x_j)\right]_{i,j=1}^{N} \mathrm{d}\mu\left(x_{n+1}\right)\cdots\mathrm{d}\mu\left(x_N\right) = (N-n)!\det\left[K\left(x_i,x_j\right)\right]_{i,j=1}^{n},$$

hence, from (4.13), the nth correlation kernel is given by

$$\rho_n\left(x_1,\ldots,x_n\right) = \det\left[K\left(x_i,x_j\right)\right]_{i,j=1}^{n},$$

so the point process is determinantal. □

In the following sections, we show how to use this theory, together with Schur process, to study the TASEP model.

4.4. The Schur Measure as a Determinantal Point Process

4.4.1. Shifted Schur Measure

We showed in Section 2.5 that any sequence $\alpha = (\alpha_1 \geq \alpha_2 \geq \cdots \geq 0)$ satisfying $\sum_i \alpha_i < \infty$ defines a specialization $\rho = (\alpha;\mathbf{0};0)$, which acts on symmetric functions as the evaluation map $x_i \mapsto \alpha_i$ (Example 2.35). Take a finite sequence $a = (a_1 \geq \cdots \geq a_N)$ and consider the Schur-positive specialization

$$a = ((a_1,\ldots,a_N,0,0,\ldots);\mathbf{0};0). \tag{4.16}$$

By the definition of Schur functions (Definition 2.23), evaluating a Schur function in a is the same as evaluating the Schur polynomial in N variables on the values $a_1,\ldots,a_N$, that is,

$$\mathrm{s}_\lambda\left(a_1,\ldots a_N,0,0,\ldots\right) = \mathrm{s}_\lambda\left(a_1,\ldots,a_N\right).$$

Let $a=(a_1,\dots,a_N)$ and $b=(b_1,\dots,b_N)$ with $a_i,b_j\geq 0$ and $a_ib_j<1$ for all i,j (we will justify this assumption later). Considering the corresponding specializations a and b, as in (4.16), the Schur measure parametrized by a and b is now given in terms of Schur polynomials, i.e.,

$$\mathbb{S}_{a,b}(\lambda)=\frac{s_\lambda(a)s_\lambda(b)}{\prod(a,b)}=\frac{s_\lambda(a_1,\dots,a_N)s_\lambda(b_1,\dots,b_N)}{\prod(a,b)},\quad \lambda\in\mathbb{Y},$$

where $\prod(a,b)=\prod_{1\leq i,j\leq N}\frac{1}{1-a_ib_j}$. Observe that $\mathbb{S}_{a,b}$ is supported on $\{\lambda:\ell(\lambda)\leq N\}$ by definition of Schur polynomials.

We would like to define a point process using the Schur measure. For this purpose, we could assume that each term of a partition λ is a particle on $\mathbb{Z}_+=\{0,1,2,\dots\}$. But there is a problem: we may have $\lambda_i=\lambda_j$ for $i\neq j$, so there would be two particles on the same position. To avoid this situation, define

$$\mu_i\coloneqq\lambda_i+N-i,\quad i=1,\dots,N,$$

where $\ell(\lambda)\leq N$, so that $\mu_1>\mu_2>\cdots>\mu_N\geq 0$, thus

$$X_N(\lambda)\coloneqq(\mu_1,\dots,\mu_N)$$

is a configuration in $\mathbb{Z}_+$.

Definition 4.8. Let $a=(a_1,\dots,a_N)$ and $b=(b_1,\dots,b_N)$ with $a_i,b_j>0$ and $a_ib_j<1$ for every pair i,j. The **shifted Schur measure** parametrized by a and b is the probability measure on $\mathbb{Z}_+^N$ given by

$$\mathbb{S}^{\text{Shift}}_{a,b}(x_1,\dots,x_N)=\frac{\det\left[a_i^{x_j}\right]\det\left[b_i^{x_j}\right]}{N!\det\left[a_i^{j-1}\right]\det\left[b_i^{j-1}\right]\prod(a,b)},\quad (x_1,\dots,x_N)\in\mathbb{Z}_+^N,$$

where all the matrices above are $N\times N$.

We have some remarks about the above definition.

Note 4.7. The function $\mathbb{S}^{\text{Shift}}_{a,b}$ is non-negative and symmetric, since permutations of variables are permutations on columns of the matrices.

Note 4.8. If $x_i=x_j$ for some $i\neq j$, them column vectors a^{x_i} and a^{x_j} are equal, so $\mathbb{S}^{\text{Shift}}_{a,b}(x_1,\dots,x_N)=0$. This is an essential property since it means that we almost surely do not have two particles at the same position.

Note 4.9. Given $X_N(\lambda)=(\mu_1,\dots,\mu_N)$ for $\ell(\lambda)\leq N$, we have by Note 4.7,

$$\sum_{x\in\sigma(X_N(\lambda))}\mathbb{S}^{\text{Shift}}_{a,b}(x_1,\dots,x_N)=N!\mathbb{S}^{\text{Shift}}_{a,b}(\mu_1,\dots,\mu_N)=\frac{s_\lambda(a)s_\lambda(b)}{\prod(a,b)}=\mathbb{S}_{a,b}(\lambda)$$

where $\sigma(X_N(\lambda))$ denotes the set of all permutations of $X_N(\lambda)$. In other words, the symmetrized version of $\mathbb{S}^{\text{Shift}}_{a,b}$ is the Schur measure.

By the remarks above, we can guarantee that the shifted Schur measure is indeed a probability measure in $\mathbb{Z}_+^N$, since

$$\sum_{x\in\mathbb{Z}_+^N}\mathbb{S}_{a,b}^{\text{Shift}}(x_1,\dots,x_N)=\sum_{x_1\neq\cdots\neq x_N}\mathbb{S}_{a,b}^{\text{Shift}}(x_1,\dots,x_N)$$
$$=\sum_{\lambda}\sum_{x\in\sigma(X_N(\lambda))}\mathbb{S}_{a,b}^{\text{Shift}}(x_1,\dots,x_N)=\sum_{\lambda}\mathbb{S}_{a,b}(\lambda)=1.$$

Since it is a symmetric measure, it induces a N-point process on $\mathbb{Z}_+$ as we showed in Example 4.6. This process is a biorthogonal ensemble (compare the Definitions 4.7 and 4.8) and consequently it is a DPP (Theorem 4.3). Moreover, we obtain a nice formula for the correlation kernel, given by a complex double integral, as we show next.

Proposition 4.11. The N-point process induced by the shifted Schur measure $\mathbb{S}_{a,b}^{\text{Shift}}$ is determinantal and the correlation kernel is given by

$$K(x,y)=\frac{1}{(2\pi\mathrm{i})^2}\oint_{\Gamma_b}\oint_{\Gamma_a}\frac{z^xw^y}{1-zw}\prod_{i=1}^{N}\frac{(1-b_kz)(1-a_kw)}{(z-a_k)(w-b_k)}\mathrm{d}z\mathrm{d}w \tag{4.17}$$

where the contours Γ_a and Γ_b are disjoint and enclose all the singularities a_i's and b_j's, respectively. More precisely, we take $w=re^{\mathrm{i}t}$ and $z=se^{\mathrm{i}t}$, for $r>\max\{b_i\}$ and $s>\max\{a_i\}$, such that $rs<1$ and $s<r$.

Proof. Assume $\mathbb{Z}_+$ is equipped with the standard counting measure. By definition,

$$\mathbb{S}_{a,b}^{\text{Shift}}(x_1,\dots,x_N)=\frac{1}{Z}\det\left[a_i^{x_j}\right]\det\left[b_i^{x_j}\right],$$

where $Z=N!\det\left[a_i^{j-1}\right]\det\left[b_i^{j-1}\right]\prod(a,b)$. Considering the functions $\phi_i(x)=a_i^x$ and $\psi_i(x)=b_i^x$, for $i,j=1,\dots,N$, we see that $\mathbb{P}_{a,b}^{\text{Shift}}$ is a biorthogonal ensemble in $\mathbb{Z}_+$, since

$$\forall i,j,\quad \int_{\mathbb{Z}_+}\phi_i(x)\psi_j(x)\mathrm{d}\mu(x)=\sum_{n=0}^{\infty}a_i^nb_j^n<\infty.$$

by the assumption that $\left|a_ib_j\right|<1$ for every i,j. Therefore, by Theorem 4.3, this point process is determinantal with correlation kernel given by

$$K(x,y)=\sum_{i,j=1}^{N}a_i^xb_j^y(G^{-t})_{ij} \tag{4.18}$$

where $G=(G_{ij})$ with

$$G_{ij}=\sum_{n=0}^{\infty}a_i^nb_j^n=\frac{1}{1-a_ib_j}.$$

To obtain an explicit formula for $K(x,y)$ we need to compute G^{-t}. By Andréief's formula (Proposition 4.10),

$$Z=\int_{\mathbb{Z}_+^N}\det\left[\phi_i(x_j)\right]\det\left[\psi_i(x_j)\right]\mathrm{d}\mu^{\otimes N}(x)$$
$$=N!\det\left[\int_{\mathbb{Z}}\phi_i(x)\psi_j(x)\mathrm{d}\mu(x)\right]=N!\det G,$$

so

$$\det\left[G_{ij}\right] = \det\left[a_i^{j-1}\right]\det\left[b_i^{j-1}\right]\prod(a,b) = \frac{\prod_{1\le k<\ell\le N}(a_k-a_\ell)(b_k-b_\ell)}{\prod_{1\le k,\ell\le N}(1-a_kb_\ell)}. \tag{4.19}$$

By Cramer's rule,

$$\left(G^{-T}\right)_{ij} = (-1)^{i+j}\frac{\det G^{(i,j)}}{\det G}$$

where $G^{(i,j)}$ is the matrix obtained removing the i-th row and the j-th column of the matrix G. Notice that $G^{(i,j)}$ is a matrix of the same type of G, so we can compute $\det G^{(i,j)}$ with a similar formula as seen in (4.19). After some computation, we conclude that

$$\left(G^{-T}\right)_{ij} = \frac{\prod_{k=1}^{N}(1-a_ib_k)(1-a_kb_j)}{(1-a_ib_j)\prod_{k\neq i}(a_i-a_k)\prod_{k\neq j}(b_j-b_k)}.$$

Now we apply the Residue Theorem to compute the integral in (4.17), obtaining exactly the kernel in (4.18). The contours are possible to find, since $a_ib_j<1$ for every pair i,j. □

Corollary 4.6. The correlation kernel obtained in Proposition 4.11 can be also written as

$$K(x,y) = \frac{1}{(2\pi\mathrm{i})^2}\oint_{\Gamma_b}\oint_{\Gamma_{w,0}}\frac{w^{y-N}}{v^{x-N+1}}\frac{1}{v-w}\frac{\mathrm{H}(a,v)\mathrm{H}(b,w^{-1})}{\mathrm{H}(a,w)\mathrm{H}(b,v^{-1})}\mathrm{d}v\mathrm{d}w$$

where Γ_b encloses all the b_j's and $\Gamma_{w,0}$ encloses the point 0 and the variable w (which is fixed when integrating in v). Also, we denote $\mathrm{H}(a,u)=\prod_{k=1}^{N}\frac{1}{1-a_ku}$ and the same for $\mathrm{H}(b,u)$.

Proof. In the kernel obtained in Proposition 4.11, consider the change of variable $z\mapsto\frac{1}{v}$, so that $\mathrm{d}z=-\frac{\mathrm{d}v}{v^2}$ and the v-contour is a circle of radius $\frac{1}{s}>r$ oriented clockwise. The w-contour is in the interior of the v-contour, so the integral is well-posed. Since the v-contour is oriented clockwise, we change its orientation, multiplying the integral by -1, hence canceling the sign of $-\frac{\mathrm{d}v}{v^2}$. We have

$$\begin{aligned}K(x,y) &= \frac{1}{(2\pi\mathrm{i})^2}\oint_{\Gamma_b}\oint_{\Gamma_{w,0}}\frac{v^{-x}w^y}{1-wv^{-1}}\prod_{k=1}^{N}\frac{(1-b_kv^{-1})(1-a_kw)}{(v^{-1}-a_k)(w-b_k)}\frac{\mathrm{d}v}{v^2}\mathrm{d}w\\ &= \frac{1}{(2\pi\mathrm{i})^2}\oint_{\Gamma_b}\oint_{\Gamma_{w,0}}\frac{w^y}{v^{x+1}}\frac{1}{v-w}\frac{w^{-N}}{v^{-N}}\prod_{k=1}^{N}\frac{(1-b_kv^{-1})(1-a_kw)w}{(v^{-1}-a_k)(w-b_k)v}\mathrm{d}v\mathrm{d}w\end{aligned}$$

and we get the result by noticing that

$$\frac{\mathrm{H}(a,v)\mathrm{H}(b,w^{-1})}{\mathrm{H}(a,w)\mathrm{H}(b,v^{-1})} = \frac{(1-b_kv^{-1})(1-wa_k)w}{(v^{-1}-a_k)(w-b_k)v}.$$

□

4.4.2. The Schur measure is a DPP

In this section, we show that the Schur measure induces a determinantal point process and we obtain a nice formula for the correlation kernels, which is a indispensable tool for the asymptotic analysis on the next chapter.

Consider the state space $\mathfrak{X} = \mathbb{Z}$ with the standard counting measure. Let λ range over $\mathbb{Y}$ and define the map $\Phi : \mathbb{Y} \to \mathbf{Conf}(\mathbb{Z})$ by

$$\lambda \mapsto \Phi(\lambda) := X(\lambda) = \{\lambda_i - i + 1\}_{i=1}^{\infty},$$

which is measurable and the points of $X(\lambda)$ are distinct. Given specializations ρ_1 and ρ_2, the probability measure given by

$$\mathbb{P}(A) = \mathbb{S}_{\rho_1,\rho_2}\left(\Phi^{-1}(A)\right), \quad A \in \mathscr{M}_{\mathrm{Conf}(\mathbb{Z})}$$

defines a point process on $\mathbb{Z}$. From now on, we will refer to this process just as configurations $X(\lambda) = \{\lambda_i - i + 1\}_{i=1}^{\infty}$.

Theorem 4.4. (OKOUNKOV, 2001) The point process $X(\lambda) = \{\lambda_i - i + 1\}_{i=1}^{\infty} \subset \mathbb{Z}$ induced by $\mathbb{S}_{\rho_1,\rho_2}$ is determinantal and the correlation kernel is

$$K(i,j) = \frac{1}{(2\pi \mathrm{i})^2} \oint_{\Gamma_w} \oint_{\Gamma_z} \frac{w^{j-1}}{z^i} \frac{1}{z-w} \frac{\mathrm{H}(\rho_1, z)\mathrm{H}(\rho_2, w^{-1})}{\mathrm{H}(\rho_1, w)\mathrm{H}(\rho_2, z^{-1})} \mathrm{d}z\mathrm{d}w,$$

where

$$\mathrm{H}(\rho, v) = \sum_{k=0}^{\infty} h_k(\rho) v^k = \mathrm{e}^{\gamma z} \prod_i \frac{1+\beta_i v}{1-\alpha_i v}, \qquad \rho = (\alpha, \beta, \gamma).$$

Also, the contours are $|w| = R_1$ and $|z| = R_2$ with $R_1 < R_2$ so that the functions $H(\rho, v^{-1})$ and $H(\rho, v)$ are analytic in $R_1 - \varepsilon < |v| < R_2 + \varepsilon$.

Proof. We present the main ideas of a proof by Johansson (2001b). By the expression for correlation functions on discrete case (see Example 4.4), we need to prove that, for every finite set $A = \{a_1, \ldots, a_n\} \subset \mathbb{Z}$,

$$\begin{aligned} \det\left[K(a_i, a_j)\right]_{i,j=1}^{n} &= \rho_n(a_1, \ldots, a_n) = \mathbb{P}(A \subset X(\lambda)) \\ &= \sum_{\lambda : A \subset X(\lambda)} \mathbb{S}_{\rho_1,\rho_2}(\lambda) = \sum_{\lambda : A \subset X(\lambda)} \frac{\mathrm{s}_\lambda(\rho_1)\mathrm{s}_\lambda(\rho_2)}{\prod(\rho_1, \rho_2)}. \end{aligned}$$

For this purpose, it suffices to show that, for indeterminates $\mathrm{x} = \{x_1, x_2, \ldots\}$ and $\mathrm{y} = \{y_1, y_2, \ldots\}$, the formal identity

$$\det\left[\widehat{K}(a_i, a_j)\right]_{i,j=1}^{n} = \sum_{\lambda : A \subset X(\lambda)} \frac{\mathrm{s}_\lambda(x)\mathrm{s}_\lambda(y)}{\prod(\mathrm{x}, \mathrm{y})} \tag{4.20}$$

is valid, where $\widehat{K}$ is K with the specializations $\rho_1 = (\mathrm{x};\mathbf{0};0)$ and $\rho_2 = (\mathrm{y};\mathbf{0};0)$. One can show that it suffices proving (4.20) for finite variables, that is, we assume $\mathrm{x} = \{x_1,\dots,x_N,0,0,\dots\}$ and $\mathrm{y} = \{y_1,\dots,y_N,0,0,\dots\}$. Since $\mathrm{s}_\lambda(x_1,\dots,x_N) = 0$ when $\ell(\lambda) > N$, we can also suppose that we have only N particles, so the configuration is $\{\lambda_i - i + 1\}_{i=1}^N$. The reason why this assumption works is the fact that the particles of $X(\lambda)$, with index $i > N$, are deterministically distributed when $\ell(\lambda) \le N$, so they do not affect the correlation function. By a shift of $N-1$, we may even work with the configuration $X_N(\lambda) = \{\lambda_i + N - i\}_{i=1}^N \subset \mathbb{Z}_+$. This one is exactly the point process induced by the shifted Schur measure $\mathbb{S}^{\mathrm{Shift}}_{x,y}$, which we know by Corollary 4.6, that it is a determinantal point process with correlation kernel

$$\widetilde{K}(i,j) = \frac{1}{(2\pi\mathrm{i})^2}\oint_{\Gamma_y}\oint_{\Gamma_{w,0}} \frac{w^{j-N}}{v^{i-N+1}}\frac{1}{v-w}\frac{\mathrm{H}(x,v)\mathrm{H}(y,w^{-1})}{\mathrm{H}(x,w)\mathrm{H}(y,v^{-1})}\mathrm{d}v\mathrm{d}w$$

This means that the n-th correlation functions for the process $X_N(\lambda)$ induced by $\mathbb{S}^{\mathrm{Shift}}_{x,y}$ are given by

$$\widetilde{\rho}_n(a_1,\dots,a_n) = \det\left[\widetilde{K}(a_i,a_j)\right]_{i,j=1}^N,$$

therefore, the nth correlation functions for the process $\{\lambda_i - i + 1\}_{i=1}^N$ is

$$\begin{aligned}\widehat{\rho}_n(a_1,\dots,a_n) &= \mathbb{P}\left(\{a_1,\dots,a_n\} \subset \{\lambda_i - i + 1\}_{i=1}^N\right)\\ &= \mathbb{P}\left(\{a_1+N-1,\dots,a_n+N-1\} \subset X_N(\lambda)\right)\\ &= \widetilde{\rho}_n\left(a_1+N-1,\dots,a_n+N-1\right)\\ &= \det\left[\widetilde{K}(s_i+N-1,s_j+N-1)\right]_{i,j=1}^n,\end{aligned}$$

hence,

$$\begin{aligned}\widehat{K}(i,j) &= \widetilde{K}(i+N-1,j+N-1)\\ &= \frac{1}{(2\pi\mathrm{i})^2}\oint_{\Gamma_y}\oint_{\Gamma_{w,0}} \frac{w^{j-1}}{v^{i}}\frac{1}{v-w}\frac{\mathrm{H}(x,v)\mathrm{H}(y,w^{-1})}{\mathrm{H}(x,w)\mathrm{H}(y,v^{-1})}\mathrm{d}v\mathrm{d}w\end{aligned}$$

as we wanted to prove in (4.20). □

4.4.3. TASEP and DPPs

Now we come back to the TASEP. Remember the last results obtained in Chapter 3. We consider the following definition that for now just simplifies the notation.

Definition 4.9. Let $(\mathfrak{X},\mu)$ be a measure space and $K : \mathfrak{X}\times\mathfrak{X} \to \mathbb{C}$. The **Fredholm determinant** of K is defined by

$$\det[1+zK] := 1 + \sum_{n=1}^{\infty}\frac{z^n}{n!}\int_{\mathfrak{X}^n}\det\left[K(x_i,x_j)\right]_{i,j=1}^n \mathrm{d}\mu^{\otimes n}(\mathrm{x}),$$

when this sum converges.

We can now rewrite Theorem 3.2 in the following form.

Theorem 4.5. Let $t \geq N$ and $A_\ell = \{-N+1, -N+2, \ldots, \ell-N-1, \ell-N\}$. We have

$$\mathbb{P}_{\text{TASEP}}\left(y_N(t) - 1 + N \geq \ell\right) = \det\left[1 - K_{N,\ell,t}\right],$$

where $K_{N,\ell,t}(m,n) \coloneqq \mathbb{1}_{A_\ell}(m) K_{N,t}(m,n) \mathbb{1}_{A_\ell}(n)$, with

$$K_{N,t}(m,n) = \frac{1}{(2\pi i)^2} \oint_{\Gamma_0} \oint_{\Gamma_{w,0}} \frac{w^{n-1}}{z^m} \frac{1}{z-w} \frac{(1-w)^N \left(1+\frac{b}{w}\right)^t}{(1-z)^N \left(1+\frac{b}{z}\right)^t} \mathrm{d}z \mathrm{d}w.$$

The contours are $|w| = R_1 < R_2 = |z|$, such that the z-contour contains the singularities w and 0 (and not $-b$), and the w-contour contains 0.

Proof. (BAIK, 2018) Since $\mathrm{s}_\lambda(\sigma_1) = \mathrm{s}_\lambda(1, \ldots, 1)$ and $t \geq N$, the Schur measure $\mathbb{S}_{\sigma_1,\sigma_2}$ in Theorem 3.2 is supported on partitions with $\ell(\lambda) \leq N$. So we may assume that the point process $\{\lambda_i - i + 1\}_1^\infty \subset \mathbb{Z}$ induced by $\mathbb{S}_{\sigma_1,\sigma_2}$ is a N-point process in $\{\lambda_i - i + 1\}_1^N$ on $\mathfrak{X} = \{-N+1, -N+2, \ldots\}$, because the particles on the left of $1-N$ are deterministically distributed. Now $\lambda_N - N + 1$ is the left-most particle, so $\mathbb{P}(\lambda_N - N + 1 \geq \ell - N + 1)$ is the gap probability for having no particles in $A_\ell = \{-N+1, -N+2, \ldots, \ell-N\}$. So from Example 4.8,

$$\begin{aligned}\mathbb{P}(\lambda_N - N + 1 \geq \ell - N + 1) &= \mathbb{P}\left(\mathrm{N}_{A_\ell}(X(\lambda)) = 0\right) \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int_{\mathfrak{X}^n} \det\left[\mathbb{1}_{A_\ell}(m_i) K_{N,t}(m_i, m_j) \mathbb{1}_{A_\ell}(m_j)\right]_{i,j=1}^n \mathrm{d}\mu^{\otimes n}(\mathrm{m})\end{aligned}$$

where

$$K_{N,t}(m,n) = \frac{1}{(2\pi i)^2} \oint \oint \frac{w^{n-1}}{z^m} \frac{1}{z-w} \frac{\mathrm{H}(\sigma_1, z)\mathrm{H}(\sigma_2, w^{-1})}{\mathrm{H}(\sigma_1, w)\mathrm{H}(\sigma_2, z^{-1})} \mathrm{d}z \mathrm{d}w.$$

Notice that

$$\mathrm{H}(\sigma_1, z) = \left(\frac{1}{1-z}\right)^N \quad \text{and} \quad \mathrm{H}(\sigma_2, z^{-1}) = \left(1 + \frac{b}{z}\right)^t,$$

so

$$K_{N,t}(m,n) = \frac{1}{(2\pi i)^2} \oint_{\Gamma_0} \oint_{\Gamma_{w,0}} \frac{w^{n-1}}{z^m} \frac{1}{z-w} \frac{(1-w)^N (1+\frac{b}{w})^t}{(1-z)^N (1+\frac{b}{z})^t} \mathrm{d}z \mathrm{d}w.$$

Here, the contours are $|w| = R_1$ and $|z| = R_2$, where $R_1 < R_2$, such that the z-contour contains the singularities w and 0 (and not $-b$), and the w-contour contains 0. Thus, by definition of Fredholm determinant with $z = -1$,

$$\mathbb{P}\left(\lambda_N(t) - N + 1 \geq \ell - N + 1\right) = \det\left[1 - K_{N,\ell,t}\right].$$

Noticing that $y_N(t) = \lambda_N(t) - N + 1$, we finally obtain

$$\mathbb{P}\left(y_N(t) - 1 + N \geq \ell\right) = \mathbb{P}\left(\lambda_N(t) - N + 1 \geq \ell - N + 1\right) = \det\left[1 - K_{N,\ell,t}\right],$$

as we wanted. □

Scaling the time to be $\tau = bt$ and making $b \to 0$, we obtain the continuous-time TASEP (Example 3.6). For any $v \in \mathbb{C}$, we have

$$\left(1+\frac{b}{v}\right)^{t} = \left(1+\frac{v^{-1}}{b^{-1}}\right)^{\frac{\tau}{b}} \xrightarrow{b\to 0} \mathrm{e}^{\frac{\tau}{v}} \tag{4.21}$$

and this remark leads to the next proposition.

Proposition 4.12. For continuous-time TASEP, the gap probability computed in Theorem 4.5 holds for the kernel

$$K_{N,\tau}(m,n) = \frac{1}{(2\pi i)^2}\oint_{\Gamma_0}\oint_{\Gamma_1} \frac{w^{n-1}}{z^m}\frac{1}{w-z}\frac{(1-w)^N}{(1-z)^N}\mathrm{e}^{\frac{\tau}{w}-\frac{\tau}{z}}\mathrm{d}z\mathrm{d}w.$$

Proof. By (4.21), we have

$$K_{N,\tau}(m,n) = \frac{1}{(2\pi i)^2}\oint_{\Gamma_0}\oint_{\Gamma_{w,0}} \frac{w^{n-1}}{z^m}\frac{1}{z-w}\frac{(1-w)^N}{(1-z)^N}\mathrm{e}^{\frac{\tau}{w}-\frac{\tau}{z}}\mathrm{d}z\mathrm{d}w,$$

so we just need to deform the z-contour. The only singularity outside the z-contour is 1, so taking a contour Γ_1 containing 1 oriented clockwise does not affect the integral (provided that the residue at infinity is zero), that is

$$K_{N,\tau}(m,n) = \frac{1}{(2\pi i)^2}\oint_{\Gamma_0}\oint_{\Gamma_1} \frac{w^{n-1}}{z^m}\frac{1}{z-w}\frac{(1-w)^N}{(1-z)^N}\mathrm{e}^{\frac{\tau}{w}-\frac{\tau}{z}}\mathrm{d}z\mathrm{d}w.$$

Now, changing the orientation of Γ_1 we get

$$K_{N,\tau}(m,n) = \frac{1}{(2\pi i)^2}\oint_{\Gamma_0}\oint_{\Gamma_1} \frac{w^{n-1}}{z^m}\frac{1}{w-z}\frac{(1-w)^N}{(1-z)^N}\mathrm{e}^{\frac{\tau}{w}-\frac{\tau}{z}}\mathrm{d}z\mathrm{d}w.$$

□

The expression for this kernel is similar to the correlation kernel for the Laguerre Unitary Ensemble (LUE), a biorthogonal ensemble from Random Matrix Theory. This emsemble, when rescaled, converges to the Airy point process (BAIK, 2018). Motivated by this, we wonder if we can find a scaled version for the continuous-time TASEP that converges to the Airy point process, in the sense of convergence of the underlying kernels. The answer is yes! We do this in the final chapter.

CHAPTER

5

ASYMPTOTICS

The goal of this final chapter is simple but not easy to achieve: we want to study the behavior of the particles in the TASEP model in large time scale. As a consequence of all the work we have done until now, this task is reduced to understanding a kernel $K(x,y)$ that depends on time. We rescale the positions of the particles and then use asymptotic methods to prove that this scaled version of the model is closely related to the Airy point process, giving rise to the Tracy-Widom distribution.

5.1. Summary of the Chapter

In this chapter, the last one, we use asymptotic methods to understand the TASEP in large time scale. We first study some useful asymptotic methods: the **Laplace method** and the **Steepest Descent method**. They are used to approximate a function $f(x)$, defined by a "strange" formula, by a simpler function $g(x)$, for x large. For example, the Stirling's formula for factorial, given by

$$n! \simeq \sqrt{2\pi n}\left(\frac{n}{\mathrm{e}}\right)^n,$$

for n large, is obtained by Laplace's method.

After understanding the asymptotic methods, we explore their usage to see what happens to the formula (4.3), as $t \to \infty$. Since the formula depends just on the kernel $K_{N,t}$, we must analyze it.

If we simply send t to infinity in $K_{N,\ell,t}$, it will probably diverge. So we have to rescale the positions of the particles, defining new particles $\widehat{x} = \frac{1}{\beta}(x-\alpha)$, for suitable constants α and β. With this procedure, we obtain a rescaled point process with a rescaled kernel $\widehat{K}_{\xi,t}$ (where ξ depends on ℓ), so that equation (4.3) can be rewritten as

$$\mathbb{P}\left(\frac{y_N(t)-\alpha}{\beta} \leq \xi\right) = \det\left[1-\widehat{K}_{\xi,t}\right].$$

A proper choice for α and β is fundamental. They are defined by

$$\alpha = c_1 t^{2/3} + c_3 t^{1/3} \quad \text{and} \quad \beta = c_2 t^{1/3}.$$

The specific powers $t^{1/3}$ and $t^{2/3}$ are due to the so-called **KPZ scaling**, they do not appear here simply by chance, there is a physical interpretation behind them. The constants c_1, c_2 and c_3 will be showed later. Now, for t large, we have

$$\mathbb{P}\left(\frac{y_N(t)-\alpha}{\beta} \leq \xi\right) \simeq \mathbb{P}\left(\frac{y_N(t)-c_1 t^{2/3}}{c_2 t^{1/3}} \leq \xi\right), \tag{5.1}$$

where the notation $\simeq$ indicates that for $t \to \infty$ the term with $t^{2/3}$ is much larger than the term with $t^{1/3}$, so we discard the latter one.

Now we invite the reader to recall the Central Limit Theorem, which states that, for a sequence $\{X_1, X_2, \ldots\}$ of i.i.d. random variables, the quantity $\frac{S_n-\mu}{\sigma}$ converges in distribution to the Gaussian distribution, where $S_n = X_1 + \cdots + X_n$ and μ and σ^2 are the mean and the variance of each X_i, respectively. Looking back at (5.1), we notice a expression similar to $\frac{S_n-\mu}{\sigma}$. So we are trying to find a "Central Limit Theorem" for fluctuations of the particles in the TASEP model.

The main result we obtain in this chapter (using the steepest descent method) is the pointwise convergence of the scaled kernel $\widehat{K}_{\xi,t}$ to the kernel $\mathbb{A}_\xi$ of the **Airy point process**, a well-known point process in the literature. At the end we show some ideas on how to improve this convergence to obtain the convergence of Fredholm determinants

$$\det\left[1-\widehat{K}_{\xi,t}\right] \to \det\left[1-\mathbb{A}_\xi\right].$$

The expression on the right-hand side on the latter equation is the celebrated **Tracy-Widom distribution** $\mathrm{F}_2(\xi)$.

5.2. Asymptotic Methods

Given complex valued sequences $u(n)$ and $v(n)$, we write $u(n) \simeq v(n)$ if $\lim \frac{u(n)}{v(n)} = 1$ as $n \to \infty$, and we say that $u(n)$ converges to $v(n)$ asymptotically.

The reader should also understand the big $\mathscr{O}$ notation, used to describe the limiting behavior of a function close to a point (or close to infinity). For functions $f, g : \mathbb{R} \to \mathbb{R}$, we write $f(x) = \mathscr{O}(g(x))$ as $x \to c$, if there exist $M > 0$ and $\delta > 0$ such that

$$\forall x \in (c-\delta, c+\delta), \quad |f(x)| \leq M\,|g(x)|.$$

Analogously, we have $f(x) = \mathscr{O}(g(x))$ as $x \to \infty$, if there exist $M > 0$ and $x_0 \geq 0$ such that

$$\forall x \geq x_0, \quad |f(x)| \leq M\,|g(x)|.$$

For complex functions, the definitions are analogous.

Example 5.1. Let $f(x) = 13x^4 + 2x^3 - 54x + 2$. For $M = 71$ and $x_0 = 1$, we have for every $x > x_0$,

$$\begin{aligned} |f(x)| &\leq 13\left|x^4\right| + 2\left|x^3\right| + 54|x| + 2 \\ &\leq 13\left|x^4\right| + 2\left|x^4\right| + 54\left|x^4\right| + 2\left|x^4\right| \\ &= 71\left|x^4\right|, \end{aligned}$$

thus $f(x) = \mathscr{O}\left(x^4\right)$ as $x \to \infty$.

Example 5.2. We can write the exponential near 0 as

$$\begin{aligned} \mathrm{e}^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots \\ &= 1 + x + \frac{x^2}{2!} + \mathscr{O}\left(x^3\right) \quad \text{as } x \to 0 \\ &= 1 + x + \mathscr{O}\left(x^2\right) \quad \text{as } x \to 0 \\ &= 1 + \mathscr{O}(x) \quad \text{as } x \to 0. \end{aligned}$$

We now proceed to an overview of the main asymptotic methods we need.

5.2.1. Laplace Method

Some functions are defined by integrals, as the Gamma function and the Airy function discussed in Section 4.3.1, so it may be not easy to understand the behavior of such functions. For this reason, **Laplace's method** helps us to approximate such functions by simpler ones. More precisely, we are interested in approximating integrals of the type

$$I(N) = \int_a^b \mathrm{e}^{Nf(x)}\mathrm{d}x \tag{5.2}$$

as $N \to \infty$, where a and b can be $\pm\infty$. We are looking for a less complicated function g such that $I(N) \simeq g(N)$. The function g is called an **asymptotic expansion** for $I(N)$. A well-known asymptotic expansion is the Stirling's formula for factorial, which states that

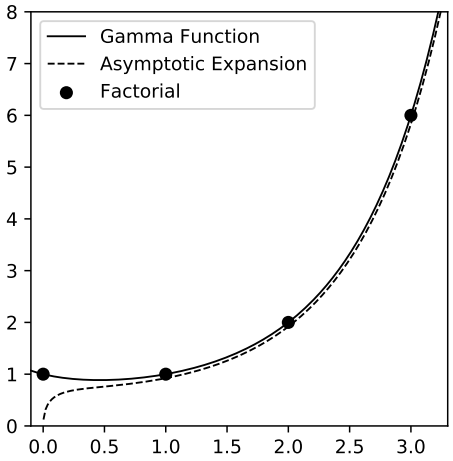
$$n! \simeq \sqrt{2\pi n}\left(\frac{n}{\mathrm{e}}\right)^n.$$

This is obtained by Laplace's method, Figure 24 shows how good this approximation is.

Let us briefly review how the method works. Assume f (in (5.2)) has a unique maximum at $x_c \in (a,b)$. Since $f(x) < f(x_c)$ for any $x \neq x_c$, the number $\mathrm{e}^{Nf(x)}$ becomes exponentially smaller than $\mathrm{e}^{Nf(x_c)}$, as $N \to \infty$. Fix a small neighborhood $V = [x_c - \delta, x_c + \delta]$, then the integral is much smaller in $[a,b] \setminus V$ than in V, thus we expect

$$I(N) = \int_a^b \mathrm{e}^{Nf(x)}\mathrm{d}x \simeq \int_{x_c-\delta}^{x_c+\delta} \mathrm{e}^{Nf(x)}\mathrm{d}x.$$

Figure 24 – Comparision between $\Gamma(x+1)$ and Stirling's formula. Remember that $n! = \Gamma(n+1)$.



Source: Elaborated by the author.

Since the integral is computed on a small interval around x_c, we can approximate f near x_c by its Taylor series. We also have $f'(x_c) = 0$ and $f''(x_c) \leq 0$ (because x_c is a maximum point). Consider for now $f''(x_c) < 0$, so that

$$f(x) \approx f(x_c) + \frac{f''(x_c)}{2}(x - x_c)^2$$

for x near x_c, thus

$$I(N) \simeq \mathrm{e}^{Nf(x_c)} \int_{x_c-\delta}^{x_c+\delta} \mathrm{e}^{N\frac{f''(x_c)}{2}(x-x_c)^2} \mathrm{d}x.$$

One may notice that the last integral is an approximation for the gaussian integral. Changing the variable to $u = \sqrt{-Nf''(x_c)}(x - x_c)$, we obtain

$$I(N) \simeq \mathrm{e}^{Nf(x_c)} \frac{1}{\sqrt{-Nf''(x_c)}} \int_{-\infty}^{\infty} \mathrm{e}^{\frac{-u^2}{2}} \mathrm{d}u = \mathrm{e}^{Nf(x_c)} \sqrt{\frac{2\pi}{-Nf''(x_c)}}.$$

Of course we would need rigorous estimations to check the precision of these approximations, this is just the intuition about the method. Also, if we had $f''(x_c) = 0$, we would have to approximate f by a third-degree polynomial and find another formula, so there are no clear rules to follow. See (BAIK; DEIFT; SUIDAN, 2016) or the books (BENDER; ORSZAG; ORSZAG, 1999; WONG, 2001) for more details.

5.2.2. Steepest Descent Method

We now extend Laplace's method for complex valued functions. Consider the integral (very similar to (5.2)), given by

$$I(N) = \int_\gamma g(z)\mathrm{e}^{Nf(z)}\mathrm{d}z \tag{5.3}$$

where f and g are complex analytic functions and γ is a contour. The main idea here is to deform γ to a new contour $\widetilde{\gamma}$, which is a level curve $\operatorname{Im} f(z) = c$ for some $c \in \mathbb{R}$. If we find such contour, we obtain

$$I(N) = \mathrm{e}^{iNc} \int_{\widetilde{\gamma}} g(z) \mathrm{e}^{N \operatorname{Re} f(z)} \mathrm{d}z$$

and then we apply Laplace's method in the last integral, in both real and imaginary parts. As we saw before, to apply Laplace's method, $\operatorname{Re} f(z)$ must achieve a maximum point on $\widetilde{\gamma}$. It may not be easy to find $\widetilde{\gamma}$, so the **method of steepest descent** helps us with this task.

We may try to understand what properties $\widetilde{\gamma}$ must have, to know what we are looking for. Consider the parametrization $\widetilde{\gamma}(t) = x(t) + \mathrm{i}y(t)$ and write

$$f(z) = f(x + \mathrm{i}y) = u(x,y) + \mathrm{i}v(x,y),$$

where $\operatorname{Re} f = u$ and $\operatorname{Im} f = v$. Assuming $\operatorname{Im} f$ constant on $\widetilde{\gamma}$, we have for all t,

$$0 = \frac{\mathrm{d}}{\mathrm{d}t} v(x(t), y(t)) = v_x x' + v_y y'.$$

Moreover, assuming that $\operatorname{Re} f$ achieves a maximum on $\widetilde{\gamma}$ at some point $z_c = \widetilde{\gamma}(t_0)$, we have

$$0 = \frac{\mathrm{d}}{\mathrm{d}t} u(x(t_0), y(t_0)) = u_x x' + u_y y'.$$

We then have the linear system

$$\begin{cases} u_x(z_c)x'(t_0) + u_y(z_c)y'(t_0) = 0; \\ v_x(z_c)x'(t_0) + v_y(z_c)y'(t_0) = 0. \end{cases}$$

Requiring $\widetilde{\gamma}$ to be smooth at z_c, we look for non-trivial solutions (x', y'). Such solutions exist when

$$\det \begin{bmatrix} u_x & u_y \\ v_x & v_y \end{bmatrix} = u_x v_y - u_y v_x = 0.$$

By Cauchy-Riemann equations (recall that $f = u + \mathrm{i}v$ is analytic), the above condition holds when $v_y^2 + v_x^2 = 0$, which is equivalent to $v_x = v_y = u_x = u_y = 0$. So we finally obtain $f'(z_c) = u_x(z_c) + \mathrm{i}v_x(z_c) = 0$, that is, z_c is a critical point of f.

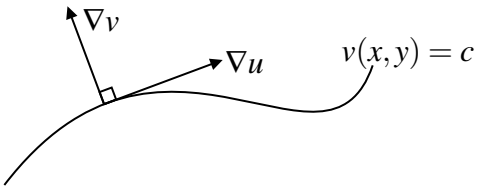
We concluded that $\widetilde{\gamma}$ must pass through a critical point (or more than one). In some cases it may be easy to find such level curves for $\operatorname{Im} f(z)$ explicitly, but in other cases it may not. It depends on the expression of f. So now we enter in the essence of the method.

Recall the definition of gradient of multivariable functions, denoted by $\nabla u = (u_x, u_y)$ and $\nabla v = (v_x, v_y)$. By Cauchy-Riemann equations,

$$\nabla u \cdot \nabla v = u_x v_x + u_y v_y = 0,$$

so ∇u and ∇v are orthogonal vectors in $\mathbb{R}^2$. A well-known result from Calculus says that the gradient ∇v on the points of a level curve $v(x,y) = c$ must be perpendicular to the curve, so ∇u is tangent to this level curve (see Figure 25).

Figure 25 – Gradients ∇u and ∇v on the level curve $v(x,y) = c$.



Source: Elaborated by the author.

So to obtain the level curves of $v(x,y)$, it suffices to follow the streamlines of the gradient field ∇u. Remember we want the curve $v(x,y) = c$ passing through z_c in such a way that z_c is a maximum point for $u(x,y)$, therefore, we should look for streamlines of the gradient field ∇u pointing towards z_c, since the gradient points towards the direction of greatest increase of u. This is the so-called **steepest descent path**, because we are choosing the path passing at z_c such that $u(x,y)$ has the fastest rate of decrease outwards z_c. So if we cannot compute the level curves $v(x,y) = c$ explicitly, we may plot the gradient field ∇u to visualize it and find the proper deformation for γ.

To summarize, we have the following steps:

1. Find the critical points of f;
2. Deform the contour of integration to a level curve of $\operatorname{Im} f$ passing through a critical point z_c. If such curve is hard to compute explicitly, we use the fact that they are streamlines of ∇u;
3. Apply Laplace's method.

Example 5.3. Let us compute the Airy function asymptotics as $x \to \infty$, using the steepest descent method. Remember from Definition 4.5, that

$$\operatorname{Ai}(x) = \frac{1}{2\pi} \int_{\infty e^{i\theta_1}}^{\infty e^{i\theta_2}} e^{i\left(\frac{z^3}{3} + xz\right)} dz,$$

where $\theta_1 \in \left(\frac{2\pi}{3}, \pi\right)$ and $\theta_2 \in \left(0, \frac{\pi}{3}\right)$. First we define $z = x^{\frac{1}{2}} w$ and $t = x^{\frac{3}{2}}$ to obtain

$$\operatorname{Ai}\left(t^{\frac{2}{3}}\right) = \frac{t^{\frac{1}{3}}}{2\pi} \int_\gamma e^{i\left(\frac{tw^3}{3} + tw\right)} dw = \frac{t^{\frac{1}{3}}}{2\pi} \int_\gamma e^{tf(w)} dw,$$

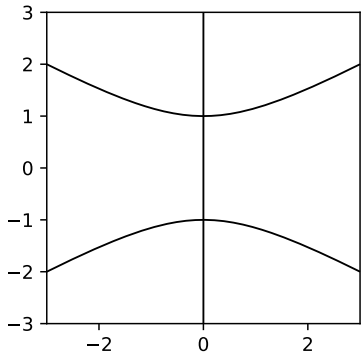
where $f(w) = i\left(\frac{w^3}{3} + w\right)$ and the contour γ is the anterior one scaled by $x^{-\frac{1}{2}}$. Since $f'(w) = i(w^2+1)$, the critical points are $w_\pm = \pm i$. We have

$$\operatorname{Re} f(w) = \frac{y^3}{3} - y - x^2 y, \quad \operatorname{Im} f(w) = \frac{x^3}{3} + x - xy^2,$$

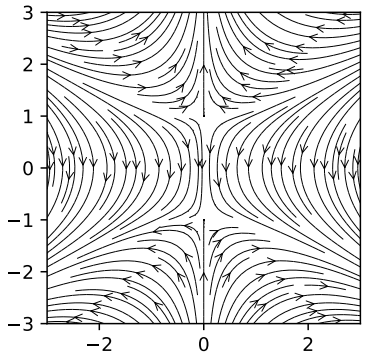
so the level curves of $\operatorname{Im} f(z)$ passing the critical points are given by $\operatorname{Im} f(z) = \operatorname{Im} f(\pm \mathrm{i}) = 0$ and we obtain the explicit formulas

$$x = 0, \qquad \frac{x^2}{3} - y^2 + 1 = 0.$$

The second equation is the equation for a hyperbola with two leaves (Figure 26a). Plotting the gradient field $\operatorname{Re} f(w)$ (Figure 26b), we notice streamlines near the upper leaf of the hyperbola pointing to the critical point $w_1 = \mathrm{i}$, so this is the level curve $\widetilde{\gamma}$ we want. Also,



(a) Level curves $x = 0$ and $\frac{x^2}{3} - y^2 + 1 = 0$.



(b) Gradient field for $\operatorname{Re} f(w)$.

we can deform γ to $\widetilde{\gamma}$ without changing the value of the integral, by the definition of the Airy function. Finally, we apply Laplace's method. Since $f(\mathrm{i}) = -\frac{2}{3}$ and $f''(\mathrm{i}) = -2$, we have

$$f(w) \approx -\frac{2}{3} - (w - \mathrm{i})^2,$$

for w near i. So

$$\int_{\widetilde{\gamma}} \mathrm{e}^{tf(w)} \mathrm{d}w \simeq \mathrm{e}^{-\frac{2}{3}t} \int_{\widetilde{\gamma}} \mathrm{e}^{-(w-\mathrm{i})^2 t} \mathrm{d}w, \quad \text{as } t \to \infty.$$

Since $\mathrm{e}^{-(w-\mathrm{i})^2}$ is analytic and it decays exponentially fast as $|w| \to \infty$, we can deform $\widetilde{\gamma}$ to $\mathbb{R}$ (using Cauchy's Theorem), then

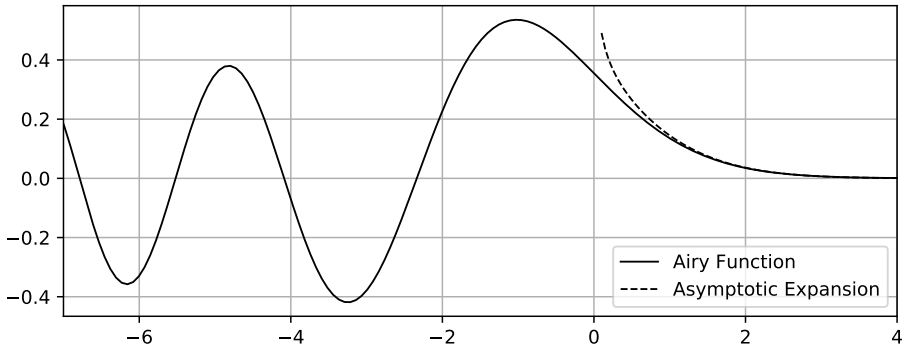
$$\int_{\widetilde{\gamma}} \mathrm{e}^{tf(w)} \simeq \mathrm{e}^{-\frac{2}{3}t} \int_{-\infty}^{\infty} \mathrm{e}^{-(w-1)^2} \mathrm{d}w = \mathrm{e}^{-\frac{2}{3}t} \sqrt{\frac{\pi}{t}},$$

using the change of variable $z = \sqrt{t}(w - i)$ and the Gaussian integral. Returning to the variable $x = t^{\frac{2}{3}}$, we obtain the asymptotic expansion for the Airy function, which is given by

$$\operatorname{Ai}(x) \simeq \frac{1}{2\sqrt{\pi} x^{1/4}} \mathrm{e}^{-\frac{2}{3}x^{3/2}}, \quad \text{as } x \to \infty.$$

Notice again that we did not provide precise estimations for these approximations, but in Figure 27 we can see how good this approximation is as $x \to \infty$. With appropriate detailed analysis, it is also possible to turn the formal calculations above to rigorous statements, but our goal here was to illustrate the methods.

Figure 27 – Airy function and its asymptotic expansion as $x \to \infty$.



Source: Elaborated by the author.

The way we apply this method changes according to the function f in (5.3), so it is hard to provide general results: we have to analyze each case individually (and carefully). In the next section, we apply the steepest descent method to study the large-time asymptotics in TASEP, providing more rigorous details.

5.3. Large-time asymptotics in TASEP

5.3.1. Rescaled TASEP and the Airy Point Process

For continuous-time TASEP, we have by Proposition 4.12,

$$\mathbb{P}\left(y_N(t) - 1 + N \geq \ell\right) = \det\left[1 - K_{N,\ell,t}\right],$$

where $K_{N,\ell,t}(m,n) = \mathbb{1}_{A_\ell}(m) K_{N,t}(m,n) \mathbb{1}_{A_\ell}(n)$, for $A_\ell = \{-N+1, \ldots, \ell - N\}$ and

$$K_{N,t}(m,n) = \frac{1}{(2\pi i)^2} \oint_{\Gamma_0} \oint_{\Gamma_1} \frac{w^{n-1}}{z^m} \frac{1}{w-z} \frac{(w-1)^N}{(z-1)^N} \mathrm{e}^{\frac{t}{w} - \frac{t}{z}} \mathrm{d}z \mathrm{d}w. \tag{5.4}$$

We now discuss what happens to $K_{N,t}$ when time gets large. If we simply send t to infinity the expression will diverge, so we rescale the variables. First, consider a new variable u such that

$$N = \frac{t}{4} + u\left(\frac{t}{2}\right)^{2/3}.$$

Note that $N = \mathscr{O}(t)$ as $t \to \infty$, so the number of particles increases to infinity. We rescale the positions of the particles, transforming a configuration X to a new configuration $\widehat{X} = \frac{1}{\beta}(X - \alpha)$, where

$$\beta = -\left(\frac{t}{2}\right)^{1/3} \quad \text{and} \quad \alpha = -2u\left(\frac{t}{2}\right)^{2/3} + u^2\left(\frac{t}{2}\right)^{1/3}. \tag{5.5}$$

As it turns out, this scaling will provide a non-trivial limit. The precise powers $t^{1/3}$ and $t^{2/3}$ also admit interpretation on physical grounds, see (FERRARI, 2019).

By Corollary 4.4, the rescaled point process $\widehat{X}$ has kernel

$$\widehat{K}_t(x,y) = |\beta|\, K_{N,t}(\beta x+\alpha, \beta y+\alpha)$$

and after a direct computation, we obtain

$$\widehat{K}_t(x,y) = \left(\frac{t}{2}\right)^{1/3} \frac{1}{(2\pi \mathrm{i})^2} \oint_{\Gamma_0}\oint_{\Gamma_1} \frac{\mathrm{e}^{t f_0(w)+\left(\frac{t}{2}\right)^{1/3} f_1(w,u,y)+\left(\frac{t}{2}\right)^{2/3} f_2(w,u)}}{\mathrm{e}^{t f_0(z)+\left(\frac{t}{2}\right)^{1/3} f_1(z,u,x)+\left(\frac{t}{2}\right)^{2/3} f_2(z,u)}} \frac{w^{-1}}{w-z} \mathrm{d}z\mathrm{d}w, \tag{5.6}$$

where

$$\begin{aligned} f_0(v) &= \frac{1}{v} + \frac{1}{4}\log(v-1); \\ f_1(v,u,s) &= (u^2-s)\log v; \\ f_2(v,u) &= u\log(v-1) - 2u\log v. \end{aligned}$$

The next theorem relates the TASEP to the Airy point process. Recall from Note 4.4 that correlation kernels are not unique and we can conjugate $\widehat{K}_t$ without changing the correlation function of the process $\widehat{X}$. Also, recall Definition 4.6 (the Airy kernel).

Theorem 5.1. The kernel $\widehat{K}_t$ converges pointwise (after a conjugation) to the Airy kernel, as $t \to \infty$.

The proof for the above theorem follows the notes from Ferrari (2019). We break the proof in several parts (Lemma 5.1; Propositions 5.1, 5.2 and 5.3), which we present later.

By Theorem 5.1, we can say that the scaled version of continuous-time TASEP converges to the Airy point process, in the sense of convergence of the kernel. Its proof is quite long, so we break it in some steps. We first present the main ideas and later we provide rigorous details. The goal is to compute the asymptotics for $\widehat{K}_t$ as $t \to \infty$, applying the steepest descent method.

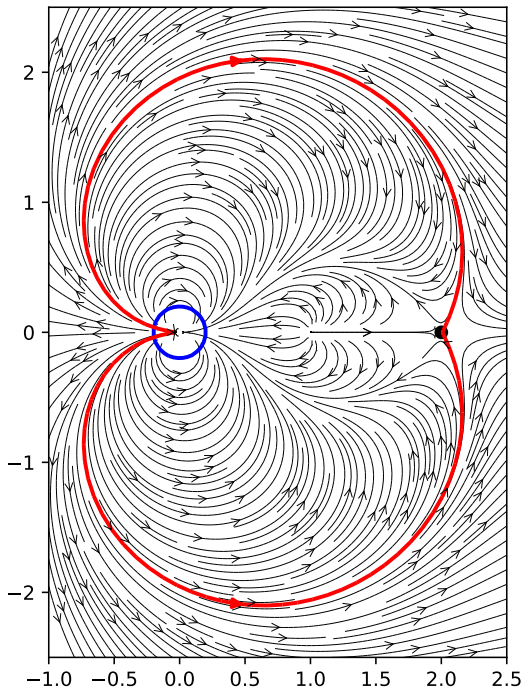
(Step 1) The term $\exp(t f_0(v))$ in the integral 5.6 suggests the use of the steepest descent method, so we apply it considering the function $f_0(v)$. Even though this integral is different from the one in (5.3) (it is even a double integral), it is still possible to use the ideas of the method. This fact illustrates how versatile and powerful the steepest descent method is, having no exact rules to follow (here mathematicians can do their art). We have

$$f_0'(v) = -\frac{1}{v^2} + \frac{1}{4}\frac{1}{v-1} = \frac{(v-2)^2}{4v^2(v-1)}$$

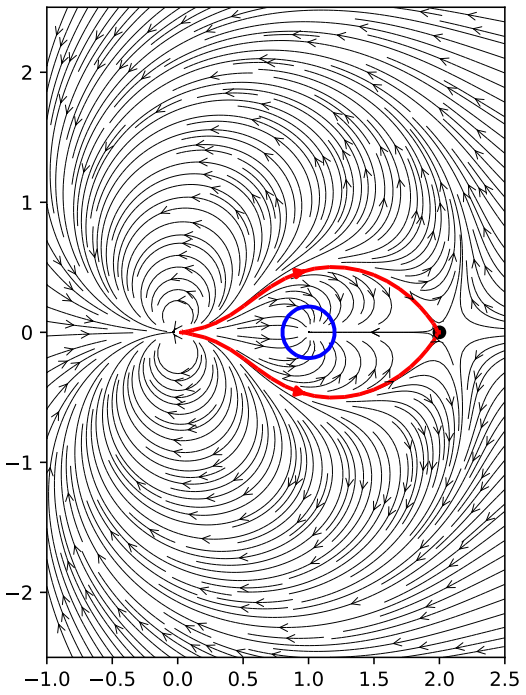
so $v_c = 2$ is the only critical point. We need to find the paths of steepest descent for w and z, passing through $v_c = 2$ (note that for z we consider $-f_0(z)$). Plotting the gradient field for

$$\nabla \operatorname{Re} f_0(x+\mathrm{i}y) = \frac{x}{x^2+y^2} + \frac{1}{8}\ln\left((x-1)^2+y^2\right)$$

(see Figures 28a and 28b), we find such paths. It is possible to deform the contours Γ_0 (resp. Γ_1) to a new contour γ_0 (resp. γ_1) in such a way that the value of the integral stays the same, the function $\operatorname{Im} f_0(w)$ is constant on γ_0 (resp. $\operatorname{Im} -f_0(z)$ is constant on γ_1) and $\operatorname{Re} f_0(w)$ achieves a unique maximum on γ_0 at $v_c = 2$ (resp. $\operatorname{Re} -f_0(z)$ achieves a unique maximum on γ_1 at v_c).



(a) Streamlines for $\nabla \operatorname{Re} f_0(w)$. The black dot is the critical point, the blue circle is the contour Γ_0 and the red curve is the new deformed contour γ_0 on a path of steepest descent.



(b) Streamlines for $\nabla - \operatorname{Re} f_0(z)$. The black dot is the critical point, the blue circle is the contour Γ_1 and the red curve is the new deformed contour γ_1 on a path of steepest descent.

In this part, we have some technical issues:

- The term $(w-z)^{-1}$ introduces a problem: to deform Γ_0 to γ_0 we pass through the singularity $w = z$, since γ_0 contains Γ_1, so we must subtract a residue. At this point, the kernel in (5.4) turns to be

$$\frac{1}{(2\pi\mathrm{i})^2}\oint_{\bar{\Gamma}_1}\oint_{\bar{\Gamma}_0}\frac{w^{n-1}}{z^m}\frac{1}{w-z}\frac{(w-1)^N}{(z-1)^N}\mathrm{e}^{\frac{t}{w}-\frac{t}{z}}\mathrm{d}w\mathrm{d}z - \frac{1}{2\pi\mathrm{i}}\oint_{\bar{\Gamma}_1} z^{n-m-1}\mathrm{d}z.$$

- The contour γ_0 does not contain the singularity $w = 0$, so we have to redefine γ_0 in a small neighborhood of zero, to maintain the value of the integral.

- Since γ_0 and γ_1 pass through the critical point v_c, we again have a problem with the term $(w-z)^{-1}$ when $w=z$. So we must deform the curves in a small neighborhood of 2 so that they do not intersect.

In summary, we obtain contours $\bar{\Gamma}_0$ and $\bar{\Gamma}_1$ that mostly agree with the steepest descent paths, except for small neighborhoods of 0 and 2 (see 31). Moreover, we claim that it is possible to deform Γ_0 and Γ_1 to these contours without changing the value of the integral.

Following the steepest descent method, we should then prove that the absolute value of the integral is exponentially small in t for w and z far from $v_c=2$. More precisely, for any small $\delta>0$, we write $\bar{\Gamma}_0=\bar{\Gamma}_0^{\delta}\cup\bar{\Gamma}_0'$, where $\bar{\Gamma}_0^{\delta}$ (resp. $\bar{\Gamma}_0'$) consists of the points $w\in\bar{\Gamma}_0$ such that $|w-2|\leq\delta$ (resp. $|w-2|>\delta$). Analogously, we define $\bar{\Gamma}_1^{\delta}$ and $\bar{\Gamma}_1'$. So we claim that the integral in $\bar{\Gamma}_0'\times\bar{\Gamma}_1'$ is $\mathscr{O}(\mathrm{e}^{-ct})$ for some c depending on δ. Now the kernel in (5.6) is

$$\mathscr{O}\left(\mathrm{e}^{-ct}\right)+\left(\frac{t}{2}\right)^{1/3}\frac{1}{(2\pi\mathrm{i})^2}\oint_{\bar{\Gamma}_0^{\delta}}\oint_{\bar{\Gamma}_1^{\delta}}\frac{\mathrm{e}^{tf_0(w)+\left(\frac{t}{2}\right)^{1/3}f_1(w,u,y)+\left(\frac{t}{2}\right)^{2/3}f_2(w,u)}}{\mathrm{e}^{tf_0(z)+\left(\frac{t}{2}\right)^{1/3}f_1(z,u,x)+\left(\frac{t}{2}\right)^{2/3}f_2(z,u)}}\frac{w^{-1}}{w-z}\mathrm{d}z\mathrm{d}w.$$

(Step 2) Now we estimate the contribution in $\bar{\Gamma}_0^{\delta}\times\bar{\Gamma}_1^{\delta}$ (near the critical point) for t large. For this purpose, we approximate f_0, f_1 and f_2 by their Taylor expansions around $v_c=2$. We have

$$\begin{aligned}
f_0(v)-f_0(2)&=\tfrac{1}{48}(v-2)^3+\mathscr{O}\left((v-2)^4\right)\quad\text{as } v\to 2;\\
f_1(v,u,s)-f_1(2,u,s)&=\tfrac{1}{2}\left(u^2-s\right)(v-2)+\mathscr{O}\left((v-2)^2\right)\quad\text{as } v\to 2;\\
f_2(v,u)-f_2(2,u)&=\tfrac{u}{4}(v-2)^2+\mathscr{O}\left((v-2)^3\right)\quad\text{as } v\to 2,
\end{aligned}$$

so

$$\begin{aligned}
t(f_0(v)-f_0(2))&=\tfrac{t}{48}(v-2)^3+\mathscr{O}\left(t(v-2)^4\right)\quad\text{as } v\to 2;\\
\left(\tfrac{t}{2}\right)^{1/3}(f_1(v,u,s)-f_1(2,u,s))&=\tfrac{1}{2}\left(\tfrac{t}{2}\right)^{1/3}\left(u^2-s\right)(v-2)+\mathscr{O}\left(t^{1/3}(v-2)^2\right)\quad\text{as } v\to 2;\\
\left(\tfrac{t}{2}\right)^{2/3}(f_2(v,u)-f_2(2,u))&=\left(\tfrac{t}{2}\right)^{2/3}\tfrac{u}{4}(v-2)^2+\mathscr{O}\left(t^{2/3}(v-2)^3\right)\quad\text{as } v\to 2.
\end{aligned}$$

Defining

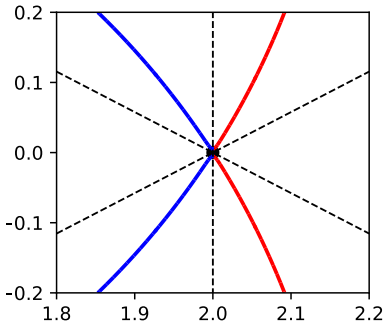
$$h(s,u):=\mathrm{e}^{tf_0(2)+(t/2)^{1/3}f_1(2,u,s)+(t/2)^{2/3}f_2(u,2)},\tag{5.7}$$

and having in mind, from Note 4.4, that we can conjugate the kernel by $\frac{h(y,u)}{h(x,u)}$, we obtain the kernel

$$\left(\tfrac{t}{2}\right)^{1/3}\frac{1}{(2\pi\mathrm{i})^2}\int_{\bar{\Gamma}_0^{\delta}}\int_{\bar{\Gamma}_1^{\delta}}\frac{\mathrm{e}^{t(f_0(w)-f_0(2))+\left(\frac{t}{2}\right)^{1/3}(f_1(w,u,y)-f_1(2,u,y))+\left(\frac{t}{2}\right)^{2/3}(f_2(w,u)-f_2(2,u))}}{\mathrm{e}^{t(f_0(z)-f_0(2))+\left(\frac{t}{2}\right)^{1/3}(f_1(z,u,x)-f_1(2,u,x))+\left(\frac{t}{2}\right)^{2/3}(f_2(z,u)-f_2(2,u))}}\frac{w^{-1}}{w-z}\mathrm{d}z\mathrm{d}w.$$

The angles for which the steepest descent paths arrive at the critical point suggest that we may find Airy asymptotics (Figure 29).

Figure 29 – The steepest descent path γ_0 (in red) "arrives" at the critical point with angles between $\left(\frac{\pi}{6},\frac{\pi}{2}\right)$ and $\left(-\frac{\pi}{2},-\frac{\pi}{6}\right)$. For the path γ_1 (in blue), we have the same but reflected. That is why we find Airy asymptotics.



Source: Elaborated by the author.

So in order to find the term $v^3/3$ (that appears in the Airy function), we do the change of variables

$$v-2 = 2V\left(\tfrac{t}{2}\right)^{-1/3} \tag{5.8}$$

for $v = w,z$ and $V = W,Z$, so that

$$\begin{aligned} t(f_0(v)-f_0(2)) &= \tfrac{V^3}{3} + \mathscr{O}\left(V^4 t^{-1/3}\right) \quad \text{as } V\to 0;\\ \left(\tfrac{t}{2}\right)^{1/3}(f_1(v,u,s)-f_1(2,u,s)) &= (u^2-s)V + \mathscr{O}\left(V^2 t^{-1/3}\right) \quad \text{as } V\to 0;\\ \left(\tfrac{t}{2}\right)^{2/3}(f_2(v,u)-f_2(2,u)) &= -uV^2 + \mathscr{O}\left(V^3 t^{-1/3}\right) \quad \text{as } V\to 0. \end{aligned}$$

With the new variables, we obtain

$$\frac{w^{-1}}{w-z} = \frac{\left(1+W\left(\frac{t}{2}\right)^{-1/3}\right)^{-1}}{4\left(\frac{t}{2}\right)^{-1/3}(W-Z)} = \frac{\left(1+\mathscr{O}\left(Wt^{-1/3}\right)\right)^{-1}}{4\left(\frac{t}{2}\right)^{-1/3}(W-Z)} \quad \text{as } W\to 0,$$

and we also have

$$\mathrm{d}z\mathrm{d}w = 4\left(\tfrac{t}{2}\right)^{-2/3}\mathrm{d}Z\mathrm{d}W.$$

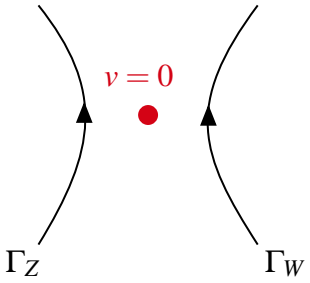
To understand the new contours, notice that since $|v-2|\le\delta$, we have $|V|\le c\delta t^{1/3}$, where $c = 2^{-4/3}$. Replacing all these expressions in the kernel, we obtain

$$\frac{1}{(2\pi\mathrm{i})^2}\int_{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\theta_1}}^{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\theta_2}}\int_{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\phi_1}}^{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\phi_2}} \frac{\left(1+\mathscr{O}\left(Wt^{-1/3}\right)\right)^{-1}}{W-Z} \times \frac{\mathrm{e}^{\frac{W^3}{3}+\mathscr{O}\left(W^4t^{-1/3}\right)-uW^2+\mathscr{O}\left(W^3t^{-1/3}\right)+\left(u^2-y\right)W+\mathscr{O}\left(W^2t^{-1/3}\right)}}{\mathrm{e}^{\frac{Z^3}{3}+\mathscr{O}\left(Z^4t^{-1/3}\right)-uZ^2+\mathscr{O}\left(Z^3t^{-1/3}\right)+\left(u^2-x\right)Z+\mathscr{O}\left(Z^2t^{-1/3}\right)}}\mathrm{d}Z\mathrm{d}W.$$

The contours are translated to be centered around zero and multiplied by a factor $ct^{1/3}$ (see Figure 30). The terms with $\mathscr{O}$ can be removed, obtaining a error of order $\mathscr{O}\left(t^{-1/3}\right)$ (we will prove this later), so the kernel is

$$\mathscr{O}\left(t^{-1/3}\right)+\frac{1}{(2\pi \mathrm{i})^2}\int_{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\theta_1}}^{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\theta_2}}\int_{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\phi_1}}^{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\phi_2}}\frac{\mathrm{e}^{\frac{W^3}{3}-uW^2-\left(y-u^2\right)W}}{\mathrm{e}^{\frac{Z^3}{3}-uZ^2-\left(x-u^2\right)Z}}\frac{\mathrm{d}Z\mathrm{d}W}{W-Z}.$$

Figure 30



(Step 3) Finally, we are able to complete the proof. First, a direct computation shows that

$$\frac{V^3}{3}-uV^2-\left(y-u^2\right)V=\frac{(V-u)^3}{3}-y(V-u)+\frac{u^3}{3}-yu.$$

So after the change of variables $W=w+u$ and $Z=z+u$, we have

$$\frac{\mathrm{e}^{\frac{u^3}{3}-yu}}{\mathrm{e}^{\frac{u^3}{3}-xu}}\frac{1}{(2\pi i)^2}\int_{c\delta t^{1/3}\mathrm{e}^{i\theta_1}-u}^{c\delta t^{1/3}\mathrm{e}^{i\theta_2}-u}\int_{c\delta t^{1/3}\mathrm{e}^{i\phi_1}-u}^{c\delta t^{1/3}\mathrm{e}^{i\phi_2}-u}\frac{\mathrm{e}^{\frac{w^3}{3}-yw}}{\mathrm{e}^{\frac{z^3}{3}-xz}}\frac{\mathrm{d}z\mathrm{d}w}{w-z}. \tag{5.9}$$

Sending t to infinity and after the change $z\mapsto \mathrm{i}z$ and $w\mapsto \mathrm{i}w$, the last integral becomes the Airy kernel $\mathbb{A}(x,y)$ (up to a conjugation term, recall Proposition 4.8). The translation of contours by u does not affect the angles of the curves at infinity.

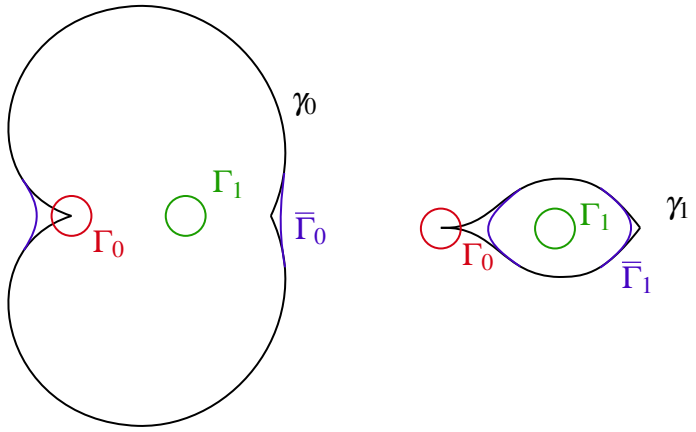
Now we present a rigorous argument in a series of propositions. We use the same notation we have presented so far in this section. First, we solve the technical issues about the deformations of contours (Step 1) in the following lemma.

Lemma 5.1. The contours Γ_0 and Γ_1 can be deformed to the contours $\bar{\Gamma}_0$ and $\bar{\Gamma}_1$ of Figure 31 without changing the value of $K_{N,t}(m,n)$.

Proof. Before going through the proof, see Figure 31 to see how the contours $\bar{\Gamma}_0$ and $\bar{\Gamma}_1$ look like. This deformation is not straightforward: because of the term $\frac{1}{w-z}$, we must subtract a residue to deform Γ_0 to $\bar{\Gamma}_0$. Changing the order of integration, we write

$$K_{N,t}(m,n)=\frac{1}{(2\pi \mathrm{i})^2}\oint_{\Gamma_1}\frac{\mathrm{e}^{-\frac{t}{z}}}{z^m(z-1)^N}\mathrm{d}z\oint_{\Gamma_0}\frac{w^{n-1}(w-1)^N}{w-z}\mathrm{e}^{\frac{t}{w}}\mathrm{d}w,$$

Figure 31 – Contours $\bar{\Gamma}_0$ and $\bar{\Gamma}_1$.



Source: Elaborated by the author.

so that we fix $z \in \Gamma_1$ to solve the integral in w. Since $w = z$ is a simple pole, we obtain

$$\operatorname{Res}\left(\frac{w^{n-1}(w-1)^N}{w-z} \mathrm{e}^{\frac{t}{w}} \middle| w = z \right) = \lim_{w\to z}(w-z)\frac{w^{n-1}(w-1)^N}{w-z}\mathrm{e}^{\frac{t}{w}} = z^{n-1}(z-1)^N \mathrm{e}^{\frac{t}{z}}.$$

The deformation from Γ_1 to $\bar{\Gamma}_1$ does not have any residue to subtract, so we have

$$K_{N,t}(m,n) = \frac{1}{(2\pi\mathrm{i})^2}\oint_{\bar{\Gamma}_1}\oint_{\bar{\Gamma}_0} \frac{w^{n-1}}{z^m}\frac{1}{w-z}\frac{(w-1)^N}{(z-1)^N}\mathrm{e}^{\frac{t}{w}-\frac{t}{z}}\mathrm{d}w\mathrm{d}z - \frac{1}{2\pi\mathrm{i}}\oint_{\bar{\Gamma}_1} z^{n-m-1}\mathrm{d}z.$$

Now z^{n-m-1} only admits singularities at $z = 0$, which is outside the contour $\bar{\Gamma}_1$, so the second integral vanishes. Note that $\bar{\Gamma}_0$ and $\bar{\Gamma}_1$ do not intersect, so the factor $(w-z)^{-1}$ is not a problem. □

Now we present a proposition that completes Step 1, proving that the integral $\widehat{K}_t(x,y)$ is exponentially small on the parts of the contours far from the critical point.

Proposition 5.1. For δ sufficiently small, the integral

$$\int_{\bar{\Gamma}_0\setminus\bar{\Gamma}_0^\delta}\int_{\bar{\Gamma}_1\setminus\bar{\Gamma}_1^\delta} \frac{\mathrm{e}^{t(f_0(w)-f_0(2))+\left(\frac{t}{2}\right)^{1/3}(f_1(w,u,y)-f_1(2,u,y))+\left(\frac{t}{2}\right)^{2/3}(f_2(w,u)-f_2(2,u))}}{\mathrm{e}^{t(f_0(z)-f_0(2))+\left(\frac{t}{2}\right)^{1/3}(f_1(z,u,x)-f_1(2,u,x))+\left(\frac{t}{2}\right)^{2/3}(f_2(z,u)-f_2(2,u))}}\frac{w^{-1}}{w-z}\mathrm{d}z\mathrm{d}w$$

is $\mathscr{O}(\mathrm{e}^{-c_0 t})$ for some c_0 depending on δ.

Proof. Since $v_c = 2$ is the unique global maximum for $\operatorname{Re} f_0(w)$ in γ_0, for any $\delta > 0$ sufficiently small, we can find η_1 such that $0 < \eta_1 < \frac{1}{2} = \operatorname{Re} f_0(2)$ and

$$w \in \gamma_0, |w-2| > \delta \quad \Rightarrow \quad \operatorname{Re} f_0(w) \le \operatorname{Re} f_0(2) - \eta_1. \tag{5.10}$$

Remember we have some freedom to choose the curve $\bar{\Gamma}_0$, so we assume $\gamma_0 \equiv \bar{\Gamma}_0$ on $\bar{\Gamma}_0 \setminus \bar{\Gamma}_0^\delta$ and we replace γ_0 by $\bar{\Gamma}_0$ in (5.10). Also, even though $\bar{\Gamma}_0 \not\equiv \gamma_0$ in a neighborhood of $w = 0$, we have $\operatorname{Re} f_0(w)$ is negative in this part of the curve, so that (5.10) still holds. Indeed, γ_0

arrives at $w=0$ with an angle $\theta \in \left(\frac{\pi}{2}+\kappa, \frac{3\pi}{2}-\kappa\right)$, so $-1 \leq \cos\theta < -\eta$ for some $\eta > 0$. In polar coordinates, we have

$$\operatorname{Re} f_0\left(re^{i\theta}\right) = \frac{\cos\theta}{r}\left[1+\frac{1}{8}\frac{r}{\cos\theta}\log\left((r\cos\theta-1)^2+r^2\sin^2\theta\right)\right],$$

so for small r, the term $\frac{\cos\theta}{r}$ goes to $-\infty$ while the term inside the square brackets remains bounded and positive, so $\operatorname{Re} f_0(w) < 0$ in a region of γ_0 close to $w=0$.

Analogously, there exists $\eta_2 > 0$ such that

$$z \in \gamma_1, |z-2| > \delta \quad \Rightarrow \quad -\operatorname{Re} f_0(z) \leq -\operatorname{Re} f_0(2) - \eta_2.$$

The functions $f_1(w,u,y)$ and $f_2(w,u)$ are bounded in $\bar{\Gamma}_0$, so assuming y in a compact set, there exists a constant $M > 0$ such that

$$\forall w \in \bar{\Gamma}_0, \quad |\operatorname{Re} f_1(w,u,y) - \operatorname{Re} f_1(2,u,y)|, |\operatorname{Re} f_1(w,u) - \operatorname{Re} f_1(2,u)| \leq M.$$

A similar result holds for $-f_1(z,u,x)$ and $-f_2(z,u)$. The term $\left|\frac{w^{-1}}{w-z}\right|$ is also bounded in $\bar{\Gamma}_0 \times \bar{\Gamma}_1$. Now for $w \in \bar{\Gamma}_0$ such that $|w-2| > \delta$, we have

$$\begin{aligned}&\operatorname{Re}\left(t(f_0(w)-f_0(2)) + \left(\tfrac{t}{2}\right)^{1/3}(f_1(w)-f_1(2)) + \left(\tfrac{t}{2}\right)^{2/3}(f_2(w)-f_2(2))\right)\\ &\quad \leq -t\eta_1 + M_1\left(t^{1/3}+t^{2/3}\right) = t\left(-\eta + M_1\left(t^{-2/3}+t^{-1/3}\right)\right) = t\left(-\eta + \mathscr{O}\left(t^{-1/3}\right)\right),\end{aligned}$$

so for t sufficiently large, we obtain $\eta_3 > 0$ such that the last expression is $\leq -t\eta_3$, thus

$$\left|\exp\left(t(f_0(w)-f_0(2)) + \left(\tfrac{t}{2}\right)^{1/3}(f_1(w)-f_1(2)) + \left(\tfrac{t}{2}\right)^{2/3}(f_2(w)-f_2(2))\right)\right| \leq e^{-t\eta_3}.$$

Analogously, for t large, $z \in \bar{\Gamma}_1$ and $|z-2| > \delta$, we find $\eta_4 > 0$ such that

$$\left|\exp\left(-t(f_0(z)-f_0(2)) - \left(\tfrac{t}{2}\right)^{1/3}(f_1(z)-f_1(2)) - \left(\tfrac{t}{2}\right)^{2/3}(f_2(z)-f_2(2))\right)\right| \leq e^{-t\eta_4}.$$

Finally, by all remarks above, we obtain

$$\left|\int_{\bar{\Gamma}_0\setminus\bar{\Gamma}_0^\delta}\int_{\bar{\Gamma}_1\setminus\bar{\Gamma}_1^\delta} \frac{e^{t(f_0(w)-f_0(2)) + \left(\frac{t}{2}\right)^{1/3}(f_1(w,u,y)-f_1(2,u,y)) + \left(\frac{t}{2}\right)^{2/3}(f_2(w,u)-f_2(2,u))}}{e^{t(f_0(z)-f_0(2)) + \left(\frac{t}{2}\right)^{1/3}(f_1(z,u,x)-f_1(2,u,x)) + \left(\frac{t}{2}\right)^{2/3}(f_2(z,u)-f_2(2,u))}} \frac{w^{-1}}{w-z} dz dw\right|$$
$$\leq C' e^{-t\eta_3} e^{-t\eta_4} = \mathscr{O}\left(e^{-c_0 t}\right),$$

as we wanted to prove. □

We also need to show that the contribution of the integral in $\bar{\Gamma}_0^\delta \times \left(\bar{\Gamma}_1 \setminus \bar{\Gamma}_1^\delta\right)$ and $\left(\bar{\Gamma}_0 \setminus \bar{\Gamma}_0^\delta\right) \times \bar{\Gamma}_1^\delta$ is exponentially small.

Proposition 5.2. For δ sufficiently small, the integral

$$\int_{\bar\Gamma_0\setminus\bar\Gamma_0^\delta}\int_{\bar\Gamma_1^\delta}\frac{\mathrm{e}^{t(f_0(w)-f_0(2))+\left(\frac{t}{2}\right)^{1/3}(f_1(w,u,y)-f_1(2,u,y))+\left(\frac{t}{2}\right)^{2/3}(f_2(w,u)-f_2(2,u))}}{\mathrm{e}^{t(f_0(z)-f_0(2))+\left(\frac{t}{2}\right)^{1/3}(f_1(z,u,x)-f_1(2,u,x))+\left(\frac{t}{2}\right)^{2/3}(f_2(z,u)-f_2(2,u))}}\frac{w^{-1}}{w-z}\mathrm{d}z\mathrm{d}w$$

is $\mathscr{O}(\mathrm{e}^{-c_1 t})$ for some c_1 depending on δ. The same holds for the integral in $\bar\Gamma_0^\delta\times\left(\bar\Gamma_1\setminus\bar\Gamma_1^\delta\right)$.

Proof. The analysis is similar to what we have done in Proposition 5.1, so we do not do it again. □

The last proposition shows that the relevant contribution to the integral of the kernel comes from the part of the contours $\bar\Gamma_0$ and $\bar\Gamma_1$ near the critical point $v_c=2$, as t gets large. It also shows how we can adapt the steepest descent method to even more complicated integrals, despite some technical issues.

Lemma 5.2. For every $z\in\mathbb{C}$, we have $|\mathrm{e}^z-1|\le|z|\,\mathrm{e}^{|z|}$.

Proof. We have

$$|\mathrm{e}^z-1|=\left|\int_0^z\mathrm{e}^w\mathrm{d}w\right|\le|z-0|\sup_{w\in[0,z]}|\mathrm{e}^w|=|z|\,\mathrm{e}^{|z|},$$

where the integral is over the segment $[0,z]$ that connects 0 and z. □

Now let us make Step 2 rigorous. We obtain an asymptotic evaluation for the same integral on the contours $\bar\Gamma_0^\delta$ and $\bar\Gamma_1^\delta$ (close to the critical point).

Proposition 5.3. For $\delta>0$ sufficiently small, the integral

$$I_\delta(t)=\left(\tfrac{t}{2}\right)^{\frac13}\frac{1}{(2\pi\mathrm{i})^2}\int_{\bar\Gamma_0^\delta}\int_{\bar\Gamma_1^\delta}\frac{\mathrm{e}^{t(f_0(w)-f_0(2))+\left(\frac{t}{2}\right)^{\frac13}(f_1(w,u,y)-f_1(2,u,y))+\left(\frac{t}{2}\right)^{\frac23}(f_2(w,u)-f_2(2,u))}}{\mathrm{e}^{t(f_0(z)-f_0(2))+\left(\frac{t}{2}\right)^{\frac13}(f_1(z,u,x)-f_1(2,u,x))+\left(\frac{t}{2}\right)^{\frac23}(f_2(z,u)-f_2(2,u))}}\frac{w^{-1}}{w-z}\mathrm{d}z\mathrm{d}w$$

is

$$\mathscr{O}\left(t^{-1/3}\right)+\frac{1}{(2\pi\mathrm{i})^2}\int_{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\theta_1}}^{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\theta_2}}\int_{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\phi_1}}^{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\phi_2}}\frac{\mathrm{e}^{\frac{W^3}{3}-uW^2-\left(y-u^2\right)W}}{\mathrm{e}^{\frac{Z^3}{3}-uZ^2-\left(x-u^2\right)Z}}\frac{\mathrm{d}Z\mathrm{d}W}{W-Z},\qquad\text{as } t\to\infty.$$

Proof. (This proof is an adaptation of ideas found in (BORODIN; FERRARI; SASAMOTO, 2008)) After the change of variables (5.8), we showed before that

$$I_\delta(t)=\frac{1}{(2\pi\mathrm{i})^2}\int_{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\theta_1}}^{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\theta_2}}\int_{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\phi_1}}^{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\phi_2}}\frac{\left(1+\mathscr{O}\left(Wt^{-1/3}\right)\right)^{-1}}{W-Z}$$
$$\times\frac{\mathrm{e}^{\frac{W^3}{3}+\mathscr{O}\left(W^4t^{-1/3}\right)-uW^2+\mathscr{O}\left(W^3t^{-1/3}\right)+\left(u^2-y\right)W+\mathscr{O}\left(W^2t^{-1/3}\right)}}{\mathrm{e}^{\frac{Z^3}{3}+\mathscr{O}\left(Z^4t^{-1/3}\right)-uZ^2+\mathscr{O}\left(Z^3t^{-1/3}\right)+\left(u^2-x\right)Z+\mathscr{O}\left(Z^2t^{-1/3}\right)}}\mathrm{d}Z\mathrm{d}W,$$

where $\theta_1 \in \left(-\frac{\pi}{2}, -\frac{\pi}{6}\right)$ and $\theta_2 \in \left(\frac{\pi}{6}, \frac{\pi}{2}\right)$. The angles ϕ_1 and ϕ_2 are the reflection of θ_1 and θ_2 on the y-axis, respectively. We show now that it is possible to remove the terms with $\mathscr{O}$ from the integral, obtaining an error of order $\mathscr{O}\left(t^{-1/3}\right)$.

Since $1 + \mathscr{O}\left(Wt^{-1/3}\right) = \mathrm{e}^{\mathscr{O}\left(Wt^{-1/3}\right)}$ as $W \to 0$, we have

$$I_\delta(t) = \frac{1}{(2\pi\mathrm{i})^2} \int_{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\theta_1}}^{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\theta_2}} \int_{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\phi_1}}^{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\phi_2}} \frac{\mathrm{e}^{\frac{W^3}{3}+\mathscr{O}\left(W^4t^{-1/3}\right)-uW^2+\mathscr{O}\left(W^3t^{-1/3}\right)+\left(u^2-y\right)W+\mathscr{O}\left(W^2t^{-1/3}\right)+\mathscr{O}\left(Wt^{-1/3}\right)}}{\mathrm{e}^{\frac{Z^3}{3}+\mathscr{O}\left(Z^4t^{-1/3}\right)-uZ^2+\mathscr{O}\left(Z^3t^{-1/3}\right)+\left(u^2-x\right)Z+\mathscr{O}\left(Z^2t^{-1/3}\right)}} \frac{\mathrm{d}Z\mathrm{d}W}{W-Z}.$$

The error is given by

$$\begin{aligned} R &= I_\delta(t) - \frac{1}{(2\pi\mathrm{i})^2} \int_{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\theta_1}}^{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\theta_2}} \int_{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\phi_1}}^{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\phi_2}} \frac{\mathrm{e}^{\frac{W^3}{3}-uW^2-\left(y-u^2\right)W}}{\mathrm{e}^{\frac{Z^3}{3}-uZ^2-\left(x-u^2\right)Z}} \frac{\mathrm{d}Z\mathrm{d}W}{W-Z} \\ &= \frac{1}{(2\pi\mathrm{i})^2} \int_{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\theta_1}}^{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\theta_2}} \int_{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\phi_1}}^{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\phi_2}} \frac{\mathrm{e}^{\frac{W^3}{3}-uW^2-\left(y-u^2\right)W}}{\mathrm{e}^{\frac{Z^3}{3}-uZ^2-\left(x-u^2\right)Z}} \left(\mathrm{e}^{t^{-\frac{1}{3}}\mathscr{O}\left(W^4,W^3,W^2,W,Z^4,Z^3,Z^2\right)} - 1\right) \frac{\mathrm{d}Z\mathrm{d}W}{W-Z}. \end{aligned}$$

Using the inequality $|\mathrm{e}^z - 1| \leq |z|\,\mathrm{e}^{|z|}$ from Lemma 5.2, we obtain

$$|R| \leq C \int_{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\theta_1}}^{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\theta_2}} \int_{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\phi_1}}^{c\delta t^{1/3}\mathrm{e}^{\mathrm{i}\phi_2}} \left| \frac{\mathrm{e}^{\frac{W^3}{3}-uW^2-\left(y-u^2\right)W}}{\mathrm{e}^{\frac{Z^3}{3}-uZ^2-\left(x-u^2\right)Z}} t^{-1/3}\mathscr{O}(\star)\,\mathrm{e}^{t^{-1/3}\mathscr{O}(\star)} \right| \frac{|\mathrm{d}Z|\,|\mathrm{d}W|}{|W-Z|},$$

where

$$\mathscr{O}(\star) = \mathscr{O}\left(W^4,W^3,W^2,W,Z^4,Z^3,Z^2\right).$$

Now we bound the integrand by a factor that does not depend on t and decays at the extremes. Since $|W| \leq c\delta t^{1/3}$, we have

$$t^{-1/3}|W| \leq c\delta, \quad t^{-1/3}\left|W^2\right| \leq c\delta\,|W|, \quad t^{-1/3}\left|W^3\right| \leq c\delta\left|W^2\right|, \quad t^{-1/3}\left|W^4\right| \leq c\delta\left|W^3\right|,$$

so

$$\mathscr{O}\left(t^{-\frac{1}{3}}W\right) \leq c\delta,\ \mathscr{O}\left(t^{-\frac{1}{3}}W^2\right) \leq c\delta\,|W|,\ \mathscr{O}\left(t^{-\frac{1}{3}}W^3\right) \leq c\delta\left|W^2\right|,\ \mathscr{O}\left(t^{-\frac{1}{3}}W^4\right) \leq c\delta\left|W^3\right|.$$

For Z we have similar inequalities. Now let $W = r\mathrm{e}^{\mathrm{i}\theta}$, we have for instance

$$\left|\mathrm{e}^{\frac{W^3}{3}+\mathscr{O}\left(t^{-1/3}W^4\right)}\right| \leq \mathrm{e}^{\mathrm{Re}\,\frac{W^3}{3}+c\delta\left|W^3\right|} = \mathrm{e}^{r^3\left(\frac{1}{3}\cos(3\theta)+c\delta\right)}.$$

and we have analogous boundaries for the other exponentials in W. More precisely, we can find $X_1(y,u,\delta,\theta)$ and $X_2(u,\delta,\theta)$, such that the part of the integrand depending on W is bounded by

$$\mathrm{e}^{r^3\left(\frac{1}{3}\cos(3\theta)+c\delta\right)+r^2X_2+rX_1+c\delta}.$$

Analogously, for $Z = se^{i\phi}$, we find Y_1 and Y_2 such that the exponential in Z is bounded by

$$e^{-s^3\left(\frac{1}{3}\cos(3\phi)-c\delta\right)-s^2Y_2-sY_1}.$$

After these remarks, we obtain

$$|R| \leq Ct^{-1/3}\int_{c\delta t^{1/3}e^{i\theta_1}}^{c\delta t^{1/3}e^{i\theta_2}}\int_{c\delta t^{1/3}e^{i\phi_1}}^{c\delta t^{1/3}e^{i\phi_2}}\frac{e^{r^3\left(\frac{1}{3}\cos(3\theta)+c\delta\right)+r^2X_2+rX_1+c\delta}}{e^{s^3\left(\frac{1}{3}\cos(3\phi)-c\delta\right)+s^2Y_2+sY_1}}\left|\frac{1}{W-Z}\mathscr{O}(\star)\right||dZ|\,|dW|,$$

where $r = |W|$, $s = |Z|$, $\theta = \arg(W)$ and $\phi = \arg(Z)$. The last integral depends on t only in the contours of integration. We can remove this dependence by sending t to ∞, so that

$$|R| \leq Ct^{-1/3}\int_{\infty e^{i\theta_1}}^{\infty e^{i\theta_2}}\int_{\infty e^{i\phi_1}}^{\infty e^{i\phi_2}}\frac{e^{r^3\left(\frac{1}{3}\cos(3\theta)+c\delta\right)+r^2X_2+rX_1+c\delta}}{e^{s^3\left(\frac{1}{3}\cos(3\phi)-c\delta\right)+s^2Y_2+sY_1}}\left|\frac{1}{W-Z}\mathscr{O}(\star)\right||dZ|\,|dW|.$$

however, we need to guarantee that this integral is well-defined. For instance, notice that when $W \to \infty e^{i\theta_1}$, we have $r \to \infty$ and $\theta \to \theta_1$. In this case, we have $\cos(3\theta) < 0$, so if we choose δ small enough, we obtain $\frac{1}{3}\cos(3\theta) + c\delta < 0$. Now for large r, the term in W becomes exponentially small, since the exponent r^3 dominates the decayment. We have a similar argument for Z, choosing δ sufficiently small such that $\frac{1}{3}\cos(3\phi) - c\delta > 0$. We then obtain the convergence for the integral, since the integrand decays exponentially in the boundaries of integration. Therefore, we proved that $|R| \leq c_2 t^{-1/3}$, so $R = \mathscr{O}\left(t^{-1/3}\right)$. □

To end the section, we enunciate Theorem 5.1 in a quantified version.

Theorem 5.2. For δ sufficiently small,

$$\frac{h(y,u)}{h(x,u)}\widehat{K}_t(x,y) = \mathscr{O}\left(t^{-1/3}\right) + \frac{1}{(2\pi i)^2}\int_{c\delta t^{1/3}e^{i\theta_1}}^{c\delta t^{1/3}e^{i\theta_2}}\int_{c\delta t^{1/3}e^{i\phi_1}}^{c\delta t^{1/3}e^{i\phi_2}}\frac{e^{\frac{w^3}{3}-yw}}{e^{\frac{z^3}{3}-xz}}\frac{dzdw}{w-z},$$

as $t \to \infty$, where

$$h(s,u) = e^{tf_0(2)+(t/2)^{1/3}f_1(2,u,s)+(t/2)^{2/3}f_2(u,2)+su}$$

defines the conjugation term.

Proof. The proof follows from the explicit expression for $\widehat{K}_t$ in (5.6) and a combination of Lemma 5.1, with Propositions 5.1, 5.2 and 5.3. The conjugation term is obtained from (5.7) and (5.9). □

Note 5.1. After the change $z \mapsto iz$ and $w \mapsto iw$, the integral in Theorem 5.2 becomes the Airy kernel $\mathbb{A}(x,y)$ when $t \to \infty$, recovering Theorem 5.1.

5.3.2. Convergence of Fredholm Determinants

Let us understand a little more about this rescaled version of the TASEP. Before the scaling, we had from Proposition 4.12,

$$\mathbb{P}\left(y_N(t)-1+N\geq \ell\right)=\det\left[1-K_{N,\ell,t}\right],$$

for

$$K_{N,\ell,t}(m,n)=\mathbb{1}_{[-N+1,\ell-N]}(m)K_{N,t}(m,n)\mathbb{1}_{[-N+1,\ell-N]}(n),\quad m,n\in\mathbb{Z}. \tag{5.11}$$

Here $y_N(t)$ is the position of the left-most particle at time t. After the scaling, the position corresponding to $y_N(t)$ is given by

$$\widehat{y}_N(t)=\tfrac{1}{\beta}(y_N(t)-\alpha),$$

and $\widehat{y}_N$ is the right-most particle, since $\beta<0$ (see (5.5)), so we denote $\widehat{y}_N=\widehat{y}_{\max}$.

The scaled state space is $\widehat{\mathfrak{X}}=\frac{1}{\beta}(\mathbb{Z}-\alpha)$ (still discrete), so $y_N(t)=\beta\widehat{y}_{\max}+\alpha$. Now, we have

$$\begin{aligned} y_N(t)-1+N\geq\ell \quad &\Longleftrightarrow \quad \beta\widehat{y}_{\max}(t)+\alpha-1+N\geq\ell \\ &\Longleftrightarrow \quad \widehat{y}_{\max}(t)\leq\frac{\ell-\alpha+1-N}{\beta} \\ &\Longleftrightarrow \quad \widehat{y}_{\max}(t)\leq\xi, \end{aligned}$$

where we are choosing $\ell=\beta\xi+\alpha+N-1$, for any $\xi\in\mathbb{R}$, since ℓ is a free parameter. We showed in the last section that $K_{N,t}$ turns to be $\widehat{K}_t$ after the scaling (see (5.6)). The last remark is for the indicators functions in (5.11). For $m=\beta x+\alpha$, we have

$$-N+1\leq m\leq\ell-N \iff \xi-\beta^{-1}\leq x\leq\frac{-N+1-\alpha}{\beta}=:M_N, \tag{5.12}$$

so $\mathbb{1}_{[-N+1,\ell-N]}(m)=\mathbb{1}_{[\xi-\beta^{-1},M_N]}(x)$. Notice that $M_N>\xi-\beta^{-1}$ for N sufficiently large, so the interval is well defined. We conclude that the scaled kernel is

$$\widehat{K}_{\xi,t}(x,y)=\mathbb{1}_{\left[\xi-\beta^{-1},M_N\right]}(x)\widehat{K}_t(x,y)\mathbb{1}_{\left[\xi-\beta^{-1},M_N\right]}(y) \tag{5.13}$$

So for the scaled TASEP, we have by the gap probability formula (Example 4.5),

$$\mathbb{P}\left(\widehat{y}_{\max}(t)\leq\xi\right)=\mathbb{P}\left(\mathrm{N}_{(\xi,\infty)}\left(\widehat{X}\right)=0\right)=\det\left[1-\widetilde{K}_{\xi,t}\right],$$

for

$$\widetilde{K}_{\xi,t}(x,y)=\mathbb{1}_{(\xi,\infty)}(x)\widehat{K}_{\xi,t}(x,y)\mathbb{1}_{(\xi,\infty)}(y),\quad x,y\in\widehat{\mathfrak{X}}.$$

However, since $\left[\xi-\beta^{-1},M_N\right]\subset(\xi,\infty)$, we have $\widetilde{K}_{\xi,t}=\widehat{K}_{\xi,t}$, so

$$\mathbb{P}\left(\widehat{y}_{\max}(t)\leq\xi\right)=\det\left[1-\widehat{K}_{\xi,t}\right].$$

Lastly, remember that $N = \mathscr{O}(t)$, $\beta = -\left(\frac{t}{2}\right)^{1/3}$ and $\widehat{K}_t \to \mathbb{A}$ pointwise as $t \to \infty$ (Theorem 5.1), so for every $x, y \in \widehat{\mathfrak{X}}$,

$$\widehat{K}_{\xi,t}(x,y) \xrightarrow{t\to\infty} \mathbb{A}_\xi(x,y) := \mathbb{1}_{(\xi,\infty)}(x)\mathbb{A}(x,y)\mathbb{1}_{(\xi,\infty)}(y).$$

Now we wonder if we have the convergence of Fredholm determinants

$$\det\left[1 - \widehat{K}_{\xi,t}\right] \xrightarrow{t\to\infty} \det\left[1 - \mathbb{A}_\xi\right], \tag{5.14}$$

since the expression on the right-hand side is a well-known distribution from Random Matrix Theory.

Definition 5.1. The function $\mathrm{F}(x) := \det[1 - \mathbb{A}_x]$ is called the **Tracy-Widom distribution**.

The Tracy-Widom distribution was first introduced in (TRACY; WIDOM, 1994) and it describes fluctuations in several models that do not seem to be related to the TASEP (JOHANSSON, 2000; BORODIN; OKOUNKOV; OLSHANSKI, 2000; JOHANSSON, 2001a; BAIK; DEIFT; JOHANSSON, 1999; PRäHOFER; SPOHN, 2002).

The convergence in (5.14) is not straightforward, since $\widehat{K}_{\xi,t}$ is defined in a discrete space, while $\mathbb{A}_\xi$ is in a continuous space. Remember from the definition of Fredholm determinants (Definition 4.9) that we have

$$\det[1 - K] = 1 + \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \int_{\mathfrak{X}^n} \det\left[K(x_i, x_j)\right]_{i,j=1}^n \mathrm{d}\mu^{\otimes n}(\mathrm{x}),$$

so for $K = \widehat{K}_{\xi,t}$, the integrals become sums over $\left[\frac{1}{\beta}(\mathbb{Z} - \alpha)\right]^n$. For $K = \mathbb{A}_\xi$, we have usual integrals in $\mathbb{R}^n$, thus, we need to figure out a way to "immerse" the discrete space into the continuous one. To solve this problem and also prove the convergence (5.14), we use some results in the theory of trace-class operators.

5.3.3. Trace-Class Operators and Fredholm Determinants

In this section, we present basic properties about Trace-class operators and Fredholm determinants. We show their close relation to kernels. This theory is quite advanced and the details are beyond the scope of this work, so we just present the necessary language and results, without proofs. We follow the results presented in (BAIK; DEIFT; SUIDAN, 2016), so see it for more details.

First we establish some notation. We denote by $\mathscr{H}$ a separable Hilbert space, as $\ell^2(I)$ (for a index set I) or $L^2(M,\mu)$, for a measure space (M,μ). A linear operator $A : \mathscr{H}_1 \to \mathscr{H}_2$ is called **bounded** if there exists $C > 0$ such that $\|Af\|_{\mathscr{H}_2} \leq C\|f\|_{\mathscr{H}_1}$ for all $f \in \mathscr{H}_1$. The space of all bounded linear operators from $\mathscr{H}_1$ to $\mathscr{H}_2$ is denoted by $\mathscr{L}(\mathscr{H}_1, \mathscr{H}_2)$, or simply $\mathscr{L}(\mathscr{H})$ when $\mathscr{H} = \mathscr{H}_1 = \mathscr{H}_2$.

We are specially interested in trace-class operators and their relation with Hilbert-Schmidt operators. The definition of such objects is evolved for the purposes of this work, so we use results from the established theory as definitions. Furthermore, we present such results considering only spaces as L^2 or ℓ^2, the ones that we need.

We denote $\mathscr{B}_2(\mathscr{H}_1,\mathscr{H}_2) \subset \mathscr{L}(\mathscr{H}_1,\mathscr{H}_2)$ as the space of **Hilbert-Schmidt operators** from $\mathscr{H}_1$ to $\mathscr{H}_2$ (Definition 5.2). The subspace $\mathscr{B}_1(\mathscr{H}) \subset \mathscr{L}(\mathscr{H})$ is the space of **Trace-class operators** in $\mathscr{H}$, the definition of such operators is evolved, so we use Proposition 5.4 to obtain them. Both $\mathscr{B}_1(\mathscr{H})$ and $\mathscr{B}_2(\mathscr{H}_1,\mathscr{H}_2)$ have norms $\|\cdot\|_1$ and $\|\cdot\|_2$.

Definition 5.2. Let $\mathscr{H}_1 = L^2(M_1,\mu_1)$ and $\mathscr{H}_2 = L^2(M_2,\mu_2)$. If $K : M_1 \times M_2 \to \mathbb{C}$ is such that

$$\int_{M_2}\int_{M_1} |K(x,y)|^2 \,\mathrm{d}\mu_1(x)\mathrm{d}\mu_2(y) < \infty \tag{5.15}$$

then the operator $K : \mathscr{H}_2 \to \mathscr{H}_1$ defined by

$$Kf(x) := \int_{M_2} K(x,y)f(y)\mathrm{d}\mu_2(y), \quad f \in \mathscr{H}_2 \tag{5.16}$$

is called a Hilbert-Schmidt operator. Moreover, we have $\|K\|_2^2$ given by the expression (5.15). The same is true if we replace L^2 by ℓ^2

The next proposition shows the relation between trace-class operators and Hilbert-Schmidt operators. There is also a useful relation between the norms $\|\cdot\|_1$ and $\|\cdot\|_2$. We use this as definition of trace-class operators.

Proposition 5.4. Let $B : \mathscr{H}_2 \to \mathscr{H}_1$ and $C : \mathscr{H}_1 \to \mathscr{H}_2$ be Hilbert-Schmidt operators, then the composition $A = BC$ is trace-class in $\mathscr{H}_1$ ($A \in \mathscr{B}_1(\mathscr{H}_1)$). Moreover, these operators satisfy $\|A\|_1 \leq \|B\|_2 \|C\|_2$.

Note 5.2. Proposition 5.4 is basically the same as Lemma 4.1. In the proposition above, we have for $f \in \mathscr{H}_2$ and $g \in \mathscr{H}_1$,

$$Bf(x) = \int_{M_2} B(x,s)f(s)\mathrm{d}\mu_2(s), \quad Cg(s) = \int_{M_1} C(s,y)g(y)\mathrm{d}\mu_1(y),$$

so

$$\begin{aligned} BCf(x) &= \int_{M_2} B(x,s)Cf(s)\mathrm{d}\mu_2(s) \\ &= \int_{M_2} B(x,s)\int_{M_1} C(s,y)f(y)\mathrm{d}\mu_1(y)\mathrm{d}\mu_2(s) \\ &= \int_{M_2}\int_{M_1} B(x,s)C(s,y)f(y)\mathrm{d}\mu_2(s)\mathrm{d}\mu_1(y). \end{aligned}$$

So A is in the form $Af(x) = \int_{M_1} A(x,y)f(y)\mathrm{d}\mu_1(y)$ for $A(x,y) = \int_{M_2} B(x,s)C(s,y)f(y)\mathrm{d}\mu_2(s)$.

Example 5.4. The Airy kernel $\mathbb{A}$ defines a trace-class operator in $L^2(\xi,\infty)$, for every $\xi \in \mathbb{R}$. The proof is the same as in Proposition 4.9.

Example 5.5. The kernel $\widehat{K}_{\xi,t}$ for the scaled TASEP (see (5.13)) is a trace-class operator in $\ell^2\left(\widehat{\mathfrak{X}}_{\xi,N}\right)$, where $\widehat{\mathfrak{X}}_{\xi,N} \coloneqq \widehat{\mathfrak{X}} \cap \left[\xi - \beta^{-1}, M_N\right]$ (see (5.12) and remember that N depends on t). We write the kernel in the form

$$\widehat{K}_t(x,y) = \left(\frac{t}{2}\right)^{1/3} \frac{1}{(2\pi \mathrm{i})^2} \oint_{\Gamma_0} \oint_{\Gamma_1} \frac{F(y,w)}{F(x,z)} \frac{w^{-1}}{w-z} \mathrm{d}z \mathrm{d}w,$$

where

$$F(s,v) = \mathrm{e}^{-t f_0(v) - \left(\frac{t}{2}\right)^{1/3} f_1(v,u,s) - \left(\frac{t}{2}\right)^{2/3} f_2(v,u)}.$$

Since $\operatorname{Re} z > \operatorname{Re} w$ in the contours, we have

$$\frac{1}{w-z} = -\int_0^\infty \mathrm{e}^{-s(z-w)} \mathrm{d}s = \mathrm{i}^2 \int_0^\infty \mathrm{e}^{-s(z-\delta-(w-\delta))},$$

for some $\delta \in (0,1)$ such that $\operatorname{Re} w < \delta < \operatorname{Re} z$. Inserting this in the expression of $\widehat{K}_t$, we obtain

$$\widehat{K}_t(x,y) = \int_0^\infty \left[\left(\tfrac{t}{2}\right)^{1/6} \tfrac{1}{2\pi} \oint_{\Gamma_1} \frac{1}{F(x,z)} \mathrm{e}^{-s(z-\delta)} \mathrm{d}z \right] \left[\left(\tfrac{t}{2}\right)^{1/6} \tfrac{1}{2\pi} \oint_{\Gamma_0} F(y,w) w^{-1} \mathrm{e}^{s(w-\delta)} \mathrm{d}w \right] \mathrm{d}s.$$

Define the functions $\widehat{L}_{\xi,t} : \widehat{\mathfrak{X}}_{\xi,N} \times [0,\infty) \to \mathbb{C}$ and $\widehat{R}_{\xi,t} : [0,\infty) \times \widehat{\mathfrak{X}}_{\xi,N} \to \mathbb{C}$ by

$$\widehat{L}_{\xi,t}(x,s) \coloneqq \left(\tfrac{t}{2}\right)^{\frac{1}{6}} \tfrac{1}{2\pi} \oint_{\Gamma_1} \frac{1}{F(x,z)} \mathrm{e}^{-s(z-\delta)} \mathrm{d}z \quad \text{and} \quad \widehat{R}_{\xi,t}(s,y) \coloneqq \left(\tfrac{t}{2}\right)^{\frac{1}{6}} \tfrac{1}{2\pi} \oint_{\Gamma_0} \frac{F(y,w)}{w} \mathrm{e}^{s(w-\delta)} \mathrm{d}w,$$

so that

$$\widehat{K}_{\xi,t}(x,y) = \int_0^\infty \widehat{L}_{\xi,t}(x,s) \widehat{R}_{\xi,t}(s,y) \mathrm{d}s, \quad x,y \in \widehat{\mathfrak{X}}_{\xi,N}.$$

Now we show that $\widehat{L}_{\xi,t}$ and $\widehat{R}_{\xi,t}$ are Hilbert-Schmidt operators, that is,

$$\int_0^\infty \int_{\widehat{\mathfrak{X}}_{\xi,N}} \left|\widehat{L}_{\xi,t}(x,s)\right|^2 \mathrm{d}x \mathrm{d}s < \infty \quad \text{and} \quad \int_{\widehat{\mathfrak{X}}_{\xi,N}} \int_0^\infty \left|\widehat{R}_{\xi,t}(s,y)\right|^2 \mathrm{d}s \mathrm{d}y < \infty.$$

The integrals in $\widehat{\mathfrak{X}}_{\xi,N}$ are finite sums, so we only need to prove that, for all $x,y \in \widehat{\mathfrak{X}}_{\xi,N}$,

$$\int_0^\infty \left|\widehat{L}_{\xi,t}(x,s)\right|^2 \mathrm{d}s < \infty \quad \text{and} \quad \int_0^\infty \left|\widehat{R}_{\xi,t}(s,y)\right|^2 \mathrm{d}s < \infty.$$

Fix some x, since $F(s,v)$ is bounded in both contours, there exists a constant $C > 0$ such that

$$\left|\widehat{L}_{\xi,t}(x,s)\right|^2 \leq C \int_{\Gamma_1} \mathrm{e}^{-s(\operatorname{Re} z - \delta)} \left|\mathrm{d}z\right|,$$

and the integrand above decays exponentially as $s \to \infty$, because we have chosen δ such that $\operatorname{Re} z - \delta > 0$, so the integral for $\widehat{L}_{\xi,t}$ is finite. The same happens for $\widehat{R}_{\xi,t}$. So we conclude that $\widehat{K}_{\xi,t}$ is trace-class in $\ell^2(\widehat{\mathfrak{X}}_{\xi,N})$, since we have decomposed it as a product of Hilbert-Schmidt operators.

The definition of Fredholm determinants can be extended for trace-class operators. For $A \in \mathscr{B}_1(\mathscr{H})$ we have a more general Fredholm determinant $\det[1-A]_{\mathscr{H}}$ that preserves Definition 4.9 in some cases. The next results are definitely useful for us.

Proposition 5.5. The map $A \mapsto \det[1+A]_{\mathscr{H}}$ is continuous in $\mathscr{B}_1(\mathscr{H})$.

Proposition 5.6. Let $\mathscr{H}_1$ and $\mathscr{H}_2$ be two Hilbert spaces. If $A : \mathscr{H}_1 \to \mathscr{H}_2$ and $B : \mathscr{H}_2 \to \mathscr{H}_1$ are bounded operators such that $AB \in \mathscr{B}_1(\mathscr{H}_2)$ and $BA \in \mathscr{B}_1(\mathscr{H}_1)$, then $\det[1-AB]_{\mathscr{H}_2} = \det[1-BA]_{\mathscr{H}_1}$.

Now in order to obtain the convergence (5.14), we need to improve the pointwise convergence $\widehat{K}_{\xi,t} \to \mathbb{A}_\xi$ to a convergence in trace-class norm $\|\cdot\|_1$, then we apply Proposition 5.5 to obtain the convergence $\det\left[1-\widehat{K}_{\xi,t}\right] \to \det\left[1-\mathbb{A}_\xi\right]$. The careful reader may have noticed that $\widehat{K}_{\xi,t}$ is defined in a discrete space, so it cannot be seen as an operator in $L^2(\xi,\infty)$, the space where $\mathbb{A}_\xi$ is defined. We provide ideas to handle this problem in the next section.

5.3.4. Next Step: The Convergence to the Tracy-Widom Distribution

In this final section, we indicate how to improve the convergence $\widehat{K}_{\xi,t} \to \mathbb{A}_\xi$ to trace-class convergence. By doing this, we reproduce the convergence of the distribution of fluctuations in the TASEP to the Tracy-Widom distribution, a well established result in the literature (JOHANSSON, 2000; BORODIN; GORIN, 2012).

We saw that a kernel $K : \mathfrak{X} \times \mathfrak{X} \to \mathbb{C}$ of a DPP can be seen as an integral operator in $L^2(\mathfrak{X})$ (continuous case). In discrete case, take an enumeration $\mathfrak{X} = \{x_n, n \in \mathbb{N}\}$ and denote each $f \in \ell^2(\mathfrak{X})$ by $f = (f(x_n))_{x_n \in \mathfrak{X}}$ or simply $f = (f(x_n))$. Considering the standard counting measure, $Kf = (Kf(x_n))_{x_n \in \mathfrak{X}}$ where (5.16) turns to be

$$Kf(x_n) = \sum_{x_m \in \mathfrak{X}} K(x_n, x_m) f(x_m)$$

for each $x_n \in \mathfrak{X}$.

In $\widehat{\mathfrak{X}} = \frac{1}{\beta}(\mathbb{Z} - \alpha)$, consider an ordered enumeration $\widehat{\mathfrak{X}} = \{x_k, k \in \mathbb{Z}\}$. The distance between any two consecutive points x_k and x_{k+1} is $|\beta|^{-1}$ (remember that $\beta = -\left(\frac{t}{2}\right)^{1/3}$). For any $\xi \in \mathbb{R}$, define $\widehat{\mathfrak{X}}_\xi := \widehat{\mathfrak{X}} \cap (\xi, \infty)$. Let $U : \ell^2(\widehat{\mathfrak{X}}_\xi) \to L^2(\xi,\infty)$ be the linear map given for each $f = (f(x_n))_{x_n \in \widehat{\mathfrak{X}}_\xi}$ by

$$Uf(x) := \sum_{x_n \in \widehat{\mathfrak{X}}_\xi} |\beta|^{1/2} \mathbb{1}_{x_k}(x) f(x_k),$$

where $\mathbb{1}_{x_k} := \mathbb{1}_{[x_k, x_{k+1})}$.

Proposition 5.7. The map U is well-defined and it is an isometry. In particular, U is continuous and injective.

Proof. We have

$$\begin{aligned}\int_{(\xi,\infty)} |Uf(x)|^2 \mathrm{d}x &= \int_{(\xi,\infty)} \left(\sum_{x_k \in \widehat{\mathfrak{X}}_\xi} |\beta|^{1/2} \mathbb{1}_{x_k}(x) f(x_k) \right)^2 \mathrm{d}x \\ &= \sum_{x_k \in \widehat{\mathfrak{X}}_\xi} |\beta| f(x_k)^2 \underbrace{\int_{(\xi,\infty)} \mathbb{1}_{x_k}(x) \mathrm{d}x}_{=|\beta|^{-1}} \\ &= \sum_{x_k \in \widehat{\mathfrak{X}}_\xi} f(x_k)^2 < \infty,\end{aligned}$$

so $\|Uf\|_{L^2} < \infty$ and $\|Uf\|_{L^2} = \|f\|_{\ell^2}$. □

Consider the inner product $\langle f, g \rangle_{L^2} = \int_\xi^\infty f(x)g(x)\mathrm{d}x$. The set of functions $\left\{ |\beta|^{1/2} \mathbb{1}_{x_k} \right\}_{x_k \in \widehat{\mathfrak{X}}_\xi}$ is orthonormal in $L^2(\xi,\infty)$ with respect to this inner product, so it provides a basis for the subspace $E = \mathrm{Span} \left\{ |\beta|^{1/2} \mathbb{1}_{x_k} \right\}_{x_k \in \widehat{\mathfrak{X}}_\xi}$. Also, let $E^\perp$ be the orthogonal complement of E, that satisfies $L^2(\xi,\infty) = E \oplus E^\perp$, that is, every $g \in L^2(\xi,\infty)$ is uniquely written as $g = f + f^\perp$, for $f \in E$ and $f^\perp \in E^\perp$. Remember that $\left\langle f, f^\perp \right\rangle_{L^2} = 0$.

Since U is injective, we now look for a map $V : L^2(\xi,\infty) \to \ell^2(\widehat{\mathfrak{X}}_\xi)$ such that $VU = \mathrm{Id}_{\ell^2}$. For this purpose, we define V by

$$g = \sum_{x_k \in \widehat{\mathfrak{X}}_\xi} f(x_k) |\beta|^{1/2} \mathbb{1}_{x_k} + f^\perp \longmapsto Vg := \left(f(x_k) \right)_{x_k \in \widehat{\mathfrak{X}}_\xi}.$$

It is straightforward that $\mathrm{Im} U = E$ and $VU = \mathrm{Id}_{\ell^2}$, so $V|_E = U^{-1}$. Therefore $V|_E$ is an isometry, since it is the inverse of the isometry U. We can at least guarantee that V is continuous, since for any $g = f + f^\perp$,

$$\|Vg\|_{\ell^2} = \|Vf\|_{\ell^2} = \|f\|_{L^2} \leq \|f\|_{L^2} + \left\| f^\perp \right\|_{L^2} = \|g\|_{L^2}.$$

Now let us explain the idea behind the operators U and V. For a kernel operator K in $\ell^2(\widehat{\mathfrak{X}}_\xi)$, we obtain the operator UKV in $L^2(\xi,\infty)$. We are "conjugating" the discrete operator K by U and V to obtain an "equivalent" operator in the continuous space (Figure 32).

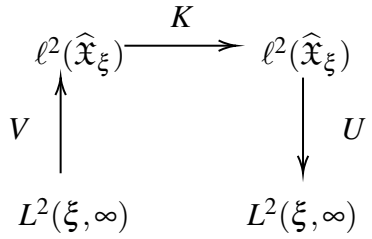
Proposition 5.8. If K is a kernel operator in $\ell^2(\widehat{\mathfrak{X}}_\xi)$, then UKV is a kernel operator in $L^2(\xi,\infty)$. More precisely, we have for all $g \in L^2(\xi,\infty)$,

$$UKVg(x) = \int_\xi^\infty \widetilde{K}(x,y) g(y) \mathrm{d}y \tag{5.17}$$

with

$$\widetilde{K}(x,y) := \sum_{x_m, x_n} |\beta| \mathbb{1}_m(x) K(x_m, x_n) \mathbb{1}_{x_n}(y).$$

Figure 32 – Action of operators U and V.



Source: Elaborated by the author.

Proof. For $g = f + f^{\perp}$, we have

$$\int_\xi^\infty \widetilde{K}(x,y)g(y)\mathrm{d}y = \int_\xi^\infty \widetilde{K}(x,y)f(y)\mathrm{d}y + \int_\xi^\infty \widetilde{K}(x,y)f^{\perp}(y)\mathrm{d}y,$$

where

$$\int_\xi^\infty \widetilde{K}(x,y)f^{\perp}(y)\mathrm{d}y = \sum_{x_m,x_n} |\beta|^{1/2}\,\mathbb{1}_{x_m}(x)K(x_m,x_n)\underbrace{\int_\xi^\infty |\beta|^{1/2}\,\mathbb{1}_n(y)f^{\perp}(y)\mathrm{d}y}_{=0,\text{ since } f^{\perp}\perp|\beta|^{1/2}\mathbb{1}_{x_n}} = 0.$$

Also, writing $f = \sum_{x_k} f(x_k)\,|\beta|^{1/2}\,\mathbb{1}_{x_k}$, we obtain

$$\begin{aligned}\int_\xi^\infty \widetilde{K}(x,y)f(y)\mathrm{d}y &= \int_{\mathbb{R}} \sum_{x_k,x_m,x_n} \underbrace{|\beta|\,\mathbb{1}_{x_m}(x)K(x_m,x_n)\mathbb{1}_{x_n}(y)\,|\beta|^{1/2}\,\mathbb{1}_{x_k}(y)f(x_k)}_{\neq 0 \iff x_n = x_k}\mathrm{d}y \\ &= \int_\xi^\infty \sum_{x_m,x_n} |\beta|\,\mathbb{1}_{x_m}(x)K(x_m,x_n)\mathbb{1}_{x_n}(y)\,|\beta|^{1/2} f(x_k)\mathrm{d}y.\end{aligned}$$

So

$$\int_\xi^\infty \widetilde{K}(x,y)g(y)\mathrm{d}y = \int_\xi^\infty \sum_{x_m,x_n} |\beta|\,\mathbb{1}_{x_m}(x)K(x_m,x_n)\mathbb{1}_{x_n}(y)\,|\beta|^{1/2} f(x_k)\mathrm{d}y.$$

On the other hand, we have

$$\begin{aligned}(UKV)\,g &= UK\left((f(x_k))_{x_k}\right) \\ &= U\left(\sum_{x_n} K(x_m,x_n)f(x_n)\right)_{x_m} \\ &= \sum_{x_m} |\beta|^{1/2}\,\mathbb{1}_{x_m}\sum_{x_n} K(x_m,x_n)f(x_n),\end{aligned}$$

so for each $x \in \mathbb{R}$,

$$\begin{aligned}(UKV)\,g(x) &= \sum_{x_m} |\beta|^{1/2}\, \mathbb{1}_{x_m}(x) \sum_{x_n} K(x_m,x_n) f(x_n) \\ &= \sum_{x_m} |\beta|^{1/2}\, \mathbb{1}_{x_m}(x) \sum_{x_n} K(x_m,x_n) \underbrace{\int_\xi^\infty |\beta|\, \mathbb{1}_{x_n}(y) f(x_n)\mathrm{d}y}_{=f(x_n)} \\ &= \int_\xi^\infty \sum_{x_m,x_n} |\beta|\, \mathbb{1}_{x_m}(x) K(x_m,x_n) \mathbb{1}_{x_n}(y)\, |\beta|^{1/2} f(x_k)\mathrm{d}y \\ &= \int_\xi^\infty \widetilde{K}(x,y) g(y)\mathrm{d}y\end{aligned}$$

as we wanted to prove. □

This construction is quite technical but very necessary for our approach to prove the convergence (5.14). In our case, since $VU = \mathrm{Id}_{\ell^2}$, we have by Proposition 5.6,

$$\det\left[1-\widehat{K}_{\xi,t}\right]_{\ell^2} = \det\left[1-VU\widehat{K}_{\xi,t}\right]_{\ell^2} = \det\left[1-U\widehat{K}_{\xi,t}V\right]_{L^2}. \tag{5.18}$$

This means that our determinant, which is in $\ell^2(\widehat{\mathfrak{X}}_\xi)$, can be written as a determinant in $L^2(\xi,\infty)$, the same space where the Airy operator is defined. Therefore, the convergence (5.14) can be studied from the point of view of trace-class operators.

We have already proved in Examples 5.4 and 5.5 that $\mathbb{A}_\xi$ and $\widehat{K}_{\xi,t}$ are trace-class in $L^2(\xi,\infty)$ and $\ell^2(\widehat{\mathfrak{X}}_\xi)$, respectively. Now we want to prove that $\widetilde{K}_{\xi,t} := U\widehat{K}_{\xi,t}V$ is also trace-class in $L^2(\xi,\infty)$. By Proposition 5.8, we have for every $f \in L^2(\xi,\infty)$ and $x \in (\xi,\infty)$,

$$\widetilde{K}_{\xi,t} f(x) = \int_\xi^\infty \widetilde{K}_{\xi,t}(x,y) f(y)\mathrm{d}y$$

where

$$\widetilde{K}_{\xi,t}(x,y) = \sum_{x_m,x_n} |\beta|\, \mathbb{1}_{x_m}(x) \widehat{K}_{\xi,t}(x_m,y_m) \mathbb{1}_{x_n}(y).$$

Notice that we are using the same notation for the function $\widetilde{K}_{\xi,t}$ and for the operator $\widetilde{K}_{\xi,t}$.

In Example 5.5, we obtained the decomposition $\widehat{K}_{\xi,t} = \widehat{L}_{\xi,t}\widehat{R}_{\xi,t}$, for Hilbert-Schmidt operators $\widehat{L}_{\xi,t} : L^2(0,\infty) \to \ell^2(\widehat{\mathfrak{X}}_\xi)$ and $\widehat{R}_{\xi,t} : \ell^2(\widehat{\mathfrak{X}}_\xi) \to L^2(0,\infty)$. So we have

$$\widetilde{K}_{\xi,t} = U\widehat{K}_{\xi,t}V = U\widehat{L}_{\xi,t}\widehat{R}_{\xi,t}V,$$

and we now denote $\widetilde{L}_{\xi,t} := U\widehat{L}_{\xi,t}$ and $\widetilde{R}_{\xi,t} := \widehat{R}_{\xi,t}V$.

To conclude the objective, we would like to prove the next two propositions, but we were not able to prove them yet.

Proposition 5.9. The operators $\widetilde{L}_{\xi,t} : L^2(0,\infty) \to L^2(\xi,\infty)$ and $\widetilde{R}_{\xi,t} : L^2(\xi,\infty) \to L^2(0,\infty)$ are Hilbert-Schmidt. Consequently, $\widetilde{K}_{\xi,t}$ is also trace-class in $L^2(\xi,\infty)$.

Proposition 5.9 is expected to be true, because U and V are isometries.

Proposition 5.10. Consider the decomposition $\mathbb{A}_\xi = L_\xi R_\xi$ given in Proposition 4.9. We have $\widetilde{L}_{\xi,t} \to L_\xi$ and $\widetilde{R}_{\xi,t} \to R_\xi$ pointwise as $t \to \infty$. Moreover, such convergences hold in Hilbert-Schmidt norm $\|\cdot\|_2$.

To prove Proposition 5.10, we would have to adapt the arguments of Theorem 5.1, but the details are not clear for us yet. With the last two propositions, we would obtain the trace-class convergence and them the convergence of Fredholm determinants, as we see next.

Proposition 5.11. We have $\widetilde{K}_{\xi,t} \xrightarrow{t\to\infty} \mathbb{A}_\xi$ in trace-class norm $\|\cdot\|_1$.

Proof. By Proposition 5.4,

$$\begin{aligned}
\left\|\widetilde{K}_{\xi,t} - \mathbb{A}_\xi\right\|_1 &= \left\|\widetilde{L}_{\xi,t}\widetilde{R}_{\xi,t} - L_\xi R_\xi\right\|_1 \\
&= \left\|\widetilde{L}_{\xi,t}\left(\widetilde{R}_{\xi,t} - R_\xi\right) + \left(\widetilde{L}_{\xi,t} - L_\xi\right)R_\xi\right\|_1 \\
&\le \left\|\widetilde{L}_{\xi,t}\left(\widetilde{R}_{\xi,t} - R_\xi\right)\right\|_1 + \left\|\left(\widetilde{L}_{\xi,t} - L_\xi\right)R_\xi\right\|_1 \\
&\le \left\|\widetilde{L}_{\xi,t}\right\|_2 \left\|\widetilde{R}_{\xi,t} - R_\xi\right\|_2 + \left\|\widetilde{L}_{\xi,t} - L_\xi\right\|_2 \|R_\xi\|_2 .
\end{aligned}$$

Since $\left\|\widetilde{R}_{\xi,t} - R_\xi\right\|_2 \to 0$ and $\left\|\widetilde{L}_{\xi,t} - L_\xi\right\|_2 \to 0$ (Proposition 5.10), we obtain $\left\|\widetilde{K}_{\xi,t} - \mathbb{A}_\xi\right\|_1 \to 0$. □

Corollary 5.1. The Fredholm determinant $\det\left[1 - \widehat{K}_{\xi,t}\right]_{\ell^2}$ converges to the Tracy-Widom distribution, as $t \to \infty$.

Proof. Since $\widetilde{K}_{\xi,t} \to \mathbb{A}_\xi$ in trace-class norm, we use Equation (5.18) and apply Proposition 5.5 to obtain the convergence

$$\det\left[1 - \widehat{K}_{\xi,t}\right]_{\ell^2} = \det\left[1 - \widetilde{K}_{\xi,t}\right]_{L^2} \to \det\left[1 - \mathbb{A}_\xi\right],$$

where the term on the right-hand side is the Tracy-Widom distribution $\mathrm{F}(\xi)$ (see Definition 5.1). □

BIBLIOGRAPHY

ADLER, M.; MOERBEKE, P. van. PDEs for the joint distributions of the Dyson, Airy and sine processes. **Ann. Probab.**, v. 33, n. 4, p. 1326–1361, 2005. ISSN 0091-1798. Available: <https://doi.org/10.1214/009117905000000107>. Citation on page 22.

BAIK, J. **Topics in Analysis: Integrable Probability Preliminary Lecture Notes**. 2018. Unpublished. Citations on pages 19, 51, 56, 57, 61, 63, 85, 96, 101, 102, 104, 105, 112, and 113.

BAIK, J.; DEIFT, P.; JOHANSSON, K. On the distribution of the length of the longest increasing subsequence of random permutations. **J. Amer. Math. Soc.**, v. 12, n. 4, p. 1119–1178, 1999. ISSN 0894-0347. Available: <https://doi.org/10.1090/S0894-0347-99-00307-0>. Citations on pages 20, 22, and 134.

BAIK, J.; DEIFT, P.; SUIDAN, T. **Combinatorics and random matrix theory**. [S.l.]: American Mathematical Soc., 2016. Citations on pages 27, 85, 91, 92, 93, 96, 118, and 134.

BAIK, J.; LIU, Z. Fluctuations of TASEP on a ring in relaxation time scale. **Comm. Pure Appl. Math.**, v. 71, n. 4, p. 747–813, 2018. ISSN 0010-3640. Available: <https://doi.org/10.1002/cpa.21702>. Citation on page 25.

BENDER, C. M.; ORSZAG, S.; ORSZAG, S. A. **Advanced mathematical methods for scientists and engineers I: Asymptotic methods and perturbation theory**. [S.l.]: Springer Science & Business Media, 1999. Citation on page 118.

BETHE, H.; METALLE, I. Zur Theorie der. Eigenwerte und eigenfunktionen der linearen atomkette. **Physik**, v. 71, p. 205–231, 1931. Citation on page 25.

BILLINGSLEY, P. **Probability and measure**. [S.l.]: John Wiley & Sons, 1995. Citation on page 88.

BLEHER, P.; ITS, A. Semiclassical asymptotics of orthogonal polynomials, Riemann-Hilbert problem, and universality in the matrix model. **Ann. of Math. (2)**, v. 150, n. 1, p. 185–266, 1999. ISSN 0003-486X. Available: <https://doi.org/10.2307/121101>. Citation on page 19.

BORODIN, A. Discrete gap probabilities and discrete Painlevé equations. **Duke Math. J.**, v. 117, n. 3, p. 489–542, 2003. ISSN 0012-7094. Available: <https://doi.org/10.1215/S0012-7094-03-11734-2>. Citation on page 22.

BORODIN, A.; FERRARI, P. L.; PRäHOFER, M.; SASAMOTO, T. Fluctuation properties of the TASEP with periodic initial configuration. **J. Stat. Phys.**, v. 129, n. 5-6, p. 1055–1080, 2007. ISSN 0022-4715. Available: <https://doi.org/10.1007/s10955-007-9383-0>. Citation on page 22.

BORODIN, A.; FERRARI, P. L.; SASAMOTO, T. Transition between $\mathbf{Airy}_1$ and $\mathbf{Airy}_2$ processes and TASEP fluctuations. **Comm. Pure Appl. Math.**, v. 61, n. 11, p. 1603–1629, 2008. ISSN 0010-3640. Available: <https://doi.org/10.1002/cpa.20234>. Citation on page 130.

BORODIN, A.; GORIN, V. **Lectures on integrable probability**. arXiv, 2012. Available: <https://arxiv.org/abs/1212.3351>. Citations on pages 19, 20, 22, 51, 61, 63, 89, and 137.

BORODIN, A.; OKOUNKOV, A.; OLSHANSKI, G. Asymptotics of Plancherel measures for symmetric groups. **J. Amer. Math. Soc.**, v. 13, n. 3, p. 481–515, 2000. ISSN 0894-0347. Available: <https://doi.org/10.1090/S0894-0347-00-00337-4>. Citations on pages 20 and 134.

BUFETOV, A.; GORIN, V. Fluctuations of particle systems determined by Schur generating functions. **Adv. Math.**, v. 338, p. 702–781, 2018. ISSN 0001-8708. Available: <https://doi.org/10.1016/j.aim.2018.07.009>. Citation on page 22.

CORWIN, I. Kardar-Parisi-Zhang universality. **Notices Amer. Math. Soc.**, v. 63, n. 3, p. 230–239, 2016. ISSN 0002-9920. Available: <https://doi.org/10.1090/noti1334>. Citations on pages 20 and 22.

DEIFT, P.; GIOEV, D. Universality at the edge of the spectrum for unitary, orthogonal, and symplectic ensembles of random matrices. **Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences**, Wiley Online Library, v. 60, n. 6, p. 867–910, 2007. Citation on page 19.

EDREI, A. On the generating function of a doubly infinite, totally positive sequence. **Transactions of the American Mathematical Society**, American Mathematical Society, v. 74, n. 3, p. 367–383, 1953. ISSN 00029947. Available: <http://www.jstor.org/stable/1990808>. Citation on page 60.

ERDőS, L.; SCHLEIN, B.; YAU, H.-T. Universality of random matrices and local relaxation flow. **Invent. Math.**, v. 185, n. 1, p. 75–119, 2011. ISSN 0020-9910. Available: <https://doi.org/10.1007/s00222-010-0302-7>. Citation on page 19.

ERDőS, L.; YAU, H.-T. Universality of local spectral statistics of random matrices. **Bull. Amer. Math. Soc. (N.S.)**, v. 49, n. 3, p. 377–414, 2012. ISSN 0273-0979. Available: <https://doi.org/10.1090/S0273-0979-2012-01372-1>. Citation on page 20.

FERRARI, P. L. **Introduction to Integrable Probability**. 2019. Unpublished. Available: <https://wt.iam.uni-bonn.de/ferrari/teaching/lectures-homepages/advtopicsws19>. Citations on pages 19, 63, and 123.

FISCHER, H. **A history of the central limit theorem: from classical to modern probability theory**. [S.l.]: Springer, 2011. Citation on page 19.

FOLLAND, G. B. **Real analysis: modern techniques and their applications**. [S.l.]: John Wiley & Sons, 1999. Citation on page 90.

GORIN, V.; SHKOLNIKOV, M. Limits of multilevel TASEP and similar processes. **Ann. Inst. Henri Poincaré Probab. Stat.**, v. 51, n. 1, p. 18–27, 2015. ISSN 0246-0203. Available: <https://doi.org/10.1214/13-AIHP555>. Citation on page 22.

GWA, L.-H.; SPOHN, H. Bethe solution for the dynamical-scaling exponent of the noisy burgers equation. **Physical Review A**, APS, v. 46, n. 2, p. 844, 1992. Citation on page 25.

JACOBI, C. G. J. De functionibus alternantibus earumque divisione per productum e differentiis elementorum conflatum. **Journal für die reine und angewandte Mathematik (Crelles Journal)**, De Gruyter, v. 1841, n. 22, p. 360–371, 1841. Citation on page 49.

JOHANSSON, K. Shape fluctuations and random matrices. **Comm. Math. Phys.**, v. 209, n. 2, p. 437–476, 2000. ISSN 0010-3616. Available: <https://doi.org/10.1007/s002200050027>. Citations on pages 20, 21, 134, and 137.

_____. Discrete orthogonal polynomial ensembles and the Plancherel measure. **Ann. of Math. (2)**, v. 153, n. 1, p. 259–296, 2001. ISSN 0003-486X. Available: <https://doi.org/10.2307/2661375>. Citations on pages 20 and 134.

_____. Random growth and random matrices. In: **European Congress of Mathematics, Vol. I (Barcelona, 2000)**. [S.l.]: Birkhäuser, Basel, 2001, (Progr. Math., v. 201). p. 445–456. Citation on page 110.

_____. **Random matrices and determinantal processes**. arXiv, 2005. Available: <https://arxiv.org/abs/math-ph/0510038>. Citation on page 85.

KEROV, S.; OKOUNKOV, A.; OLSHANSKI, G. The boundary of the Young graph with Jack edge multiplicities. **Internat. Math. Res. Notices**, n. 4, p. 173–199, 1998. ISSN 1073-7928. Available: <https://doi.org/10.1155/S1073792898000154>. Citation on page 60.

KEROV, S. V. **Asymptotic representation theory of the symmetric group and its applications in analysis**. American Mathematical Society, Providence, RI, 2003. xvi+201 p. (Translations of Mathematical Monographs, v. 219). Translated from the Russian manuscript by N. V. Tsilevich, With a foreword by A. Vershik and comments by G. Olshanski. ISBN 0-8218-3440-1. Available: <https://doi.org/10.1090/mmono/219>. Citation on page 60.

LENARD, A. Correlation functions and the uniqueness of the state in classical statistical mechanics. **Communications in Mathematical Physics**, Springer, v. 30, n. 1, p. 35–44, 1973. Citations on pages 88, 90, and 91.

_____. States of classical statistical mechanical systems of infinitely many particles. i. **Archive for Rational Mechanics and Analysis**, Springer, v. 59, n. 3, p. 219–239, 1975. Citations on pages 88, 90, and 91.

_____. States of classical statistical mechanical systems of infinitely many particles. ii. characterization of correlation measures. **Archive for Rational Mechanics and Analysis**, Springer, v. 59, n. 3, p. 241–256, 1975. Citations on pages 88, 90, and 91.

LINDSTRÖM, B. On the vector representations of induced matroids. **Bulletin of the London Mathematical Society**, Oxford University Press, v. 5, n. 1, p. 85–90, 1973. Citation on page 53.

LITTLEWOOD, D. E.; RICHARDSON, A. R. Group characters and algebra. **Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character**, The Royal Society London, v. 233, n. 721-730, p. 99–141, 1934. Citations on pages 55 and 59.

MACDONALD, I. G. **Symmetric functions and Hall polynomials**. [S.l.]: Oxford university press, 1998. Citations on pages 27, 46, 53, 54, 55, 58, and 59.

MATETSKI, K.; QUASTEL, J.; REMENIK, D. The KPZ fixed point. **Acta Math.**, v. 227, n. 1, p. 115–203, 2021. ISSN 0001-5962. Available: <https://doi.org/10.4310/acta.2021.v227.n1.a3>. Citation on page 22.

OKOUNKOV, A. Infinite wedge and random partitions. **Selecta Math. (N.S.)**, v. 7, n. 1, p. 57–81, 2001. ISSN 1022-1824. Available: <https://doi.org/10.1007/PL00001398>. Citations on pages 66 and 110.

OKOUNKOV, A.; RESHETIKHIN, N. Random skew plane partitions and the Pearcey process. **Comm. Math. Phys.**, v. 269, n. 3, p. 571–609, 2007. ISSN 0010-3616. Available: <https://doi.org/10.1007/s00220-006-0128-8>. Citation on page 22.

OKOUNKOV, A. Y. Thoma's theorem and representations of an infinite bisymmetric group. **Funktsional. Anal. i Prilozhen.**, v. 28, n. 2, p. 31–40, 95, 1994. ISSN 0374-1990. Available: <https://doi.org/10.1007/BF01076496>. Citation on page 60.

PASTUR, L.; SHCHERBINA, M. Universality of the local eigenvalue statistics for a class of unitary invariant random matrix ensembles. **J. Statist. Phys.**, v. 86, n. 1-2, p. 109–147, 1997. ISSN 0022-4715. Available: <https://doi.org/10.1007/BF02180200>. Citation on page 19.

PRäHOFER, M.; SPOHN, H. Scale invariance of the PNG droplet and the Airy process. In: . [s.n.], 2002. v. 108, n. 5-6, p. 1071–1106. Dedicated to David Ruelle and Yasha Sinai on the occasion of their 65th birthdays. Available: <https://doi.org/10.1023/A:1019791415147>. Citations on pages 20 and 134.

ROBINSON, G. d. B. On the representations of the symmetric group. **American Journal of Mathematics**, JSTOR, p. 745–760, 1938. Citation on page 33.

SAGAN, B. **The symmetric group: representations, combinatorial algorithms, and symmetric functions**. [S.l.]: Springer Science & Business Media, 2001. Citations on pages 27, 36, 46, 52, and 53.

SCHENSTED, C. Longest increasing and decreasing subsequences. **Canadian Journal of mathematics**, Cambridge University Press, v. 13, p. 179–191, 1961. Citation on page 33.

SCHUR, I. **Über eine Klasse von Matrizen, die sich einer gegebenen Matrix zuordnen lassen**. [S.l.: s.n.], 1901. Citation on page 49.

SCHüTZ, G. M. Exact solution of the master equation for the asymmetric exclusion process. **J. Statist. Phys.**, v. 88, n. 1-2, p. 427–445, 1997. ISSN 0022-4715. Available: <https://doi.org/10.1007/BF02508478>. Citation on page 25.

SOSHNIKOV, A. Universality at the edge of the spectrum in Wigner random matrices. **Comm. Math. Phys.**, v. 207, n. 3, p. 697–733, 1999. ISSN 0010-3616. Available: <https://doi.org/10.1007/s002200050743>. Citation on page 19.

_____. Determinantal random point fields. **Russian Mathematical Surveys**, IOP Publishing, v. 55, n. 5, p. 923–975, oct 2000. Citations on pages 89, 90, 95, and 96.

SPITZER, F. Interaction of Markov processes. **Advances in Math.**, v. 5, p. 246–290 (1970), 1970. ISSN 0001-8708. Available: <https://doi.org/10.1016/0001-8708(70)90034-4>. Citation on page 20.

TAO, T.; VU, V. Random matrices: universality of ESDs and the circular law. **Ann. Probab.**, v. 38, n. 5, p. 2023–2065, 2010. ISSN 0091-1798. With an appendix by Manjunath Krishnapur. Available: <https://doi.org/10.1214/10-AOP534>. Citation on page 19.

_____. Random matrices: universality of local eigenvalue statistics. **Acta Math.**, v. 206, n. 1, p. 127–204, 2011. ISSN 0001-5962. Available: <https://doi.org/10.1007/s11511-011-0061-3>. Citation on page 19.

THOMA, E. Die unzerlegbaren, positiv-definiten Klassenfunktionen der abzählbar unendlichen, symmetrischen Gruppe. **Math. Z.**, v. 85, p. 40–61, 1964. ISSN 0025-5874. Available: <https://doi.org/10.1007/BF01114877>. Citation on page 60.

TRACY, C. A.; WIDOM, H. Level-spacing distributions and the Airy kernel. **Comm. Math. Phys.**, v. 159, n. 1, p. 151–174, 1994. ISSN 0010-3616. Available: <http://projecteuclid.org/euclid.cmp/1104254495>. Citations on pages 19, 20, 21, 25, and 134.

_____. Integral formulas for the asymmetric simple exclusion process. **Comm. Math. Phys.**, v. 279, n. 3, p. 815–844, 2008. ISSN 0010-3616. Available: <https://doi.org/10.1007/s00220-008-0443-3>. Citation on page 25.

VERSHIK, A. M.; KEROV, S. V. Asymptotic theory of the characters of a symmetric group. **Funktsional. Anal. i Prilozhen.**, v. 15, n. 4, p. 15–27, 96, 1981. ISSN 0374-1990. Citation on page 60.

WONG, R. **Asymptotic approximations of integrals**. [S.l.]: SIAM, 2001. Citation on page 118.


ICMC
USP
SÃO CARLOS