



Markov processes on the path space of the Gelfand–Tsetlin graph and on its boundary

Alexei Borodin^{a,b,c,*}, Grigori Olshanski^{d,e}

^a *California Institute of Technology, United States*

^b *Massachusetts Institute of Technology, United States*

^c *Institute for Information Transmission Problems, Russian Academy of Sciences, Russia*

^d *Institute for Information Transmission Problems, Bolshoy Karetny 19, Moscow 127994, Russia*

^e *Independent University of Moscow, Russia*

Received 30 November 2011; accepted 20 March 2012

Available online 12 April 2012

Communicated by D. Voiculescu

Abstract

We construct a four-parameter family of Markov processes on infinite Gelfand–Tsetlin schemes that preserve the class of central (Gibbs) measures. Any process in the family induces a Feller Markov process on the infinite-dimensional boundary of the Gelfand–Tsetlin graph or, equivalently, the space of extreme characters of the infinite-dimensional unitary group $U(\infty)$. The process has a unique invariant distribution which arises as the decomposing measure in a natural problem of harmonic analysis on $U(\infty)$ posed in Olshanski (2003) [44]. As was shown in Borodin and Olshanski (2005) [11], this measure can also be described as a determinantal point process with a correlation kernel expressed through the Gauss hypergeometric function.

© 2012 Elsevier Inc. All rights reserved.

Keywords: Markov processes; Gelfand–Tsetlin schemes

Contents

1. Introduction	249
1.1. Gelfand–Tsetlin graph and its boundary	250

* Corresponding author at: California Institute of Technology, United States.
E-mail addresses: borodin@caltech.edu (A. Borodin), olsh2007@gmail.com (G. Olshanski).

1.2.	zw -measures	252
1.3.	Doob's transforms of N -fold products of birth and death processes	252
1.4.	Main result	253
1.5.	Markov process on Gelfand–Tsetlin schemes	253
1.6.	Further work	254
1.7.	Organization of the paper	254
2.	Abstract construction	255
2.1.	Markov kernels	255
2.2.	Feller kernels	255
2.3.	Feller semigroups	256
2.4.	Feller semigroups and Markov processes	256
2.5.	Boundary	256
2.6.	Feller boundary	257
2.7.	Extension of semigroups to the boundary	257
2.8.	Invariant measures	258
3.	Specialization. Gelfand–Tsetlin graph	259
3.1.	Spaces and links	259
3.2.	Boundary	260
3.3.	The boundary is Feller	261
4.	Generalities on Markov chains on countable spaces	262
4.1.	Regularity	262
4.2.	Collapsibility	264
4.3.	Infinitesimal generator	266
5.	Semigroups on \mathbb{GT}_N	267
5.1.	Case $N = 1$. Birth and death process on \mathbb{Z}	267
5.2.	The case of general N	271
6.	Commutativity	275
6.1.	Parameterization	275
6.2.	Infinitesimal commutativity	276
6.3.	From matrices of transition rates to semigroups	280
7.	Invariant measures	282
7.1.	zw -measures	282
7.2.	Invariance	284
8.	Stochastic dynamics on paths. General formalism	286
8.1.	Overview	286
8.2.	Construction of bivariate Markov chains	287
8.3.	Construction of multivariate Markov chains	292
9.	Stochastic dynamics on paths. Gelfand–Tsetlin graph	296
9.1.	Central measures on paths and the boundary	296
9.2.	Matrices of transition rates on $\mathbb{GT}^{(N)}$	297
9.3.	Regularity	298
9.4.	Exclusion process	299
	Acknowledgments	301
	References	301

1. Introduction

This work is a result of interaction of two circles of ideas. The first one deals with a certain class of random growth models in two space dimensions [54,41,6,7,9,10,3], while the second

one addresses constructing and analyzing stochastic dynamics on spaces of point configurations with distinguished invariant measures that are often given by, or closely related to, determinantal point processes [14–16,45,46].

Our main result is a construction of a Feller Markov process that preserves the so-called zw -measure on the (infinite-dimensional) space Ω of extreme characters of the infinite-dimensional unitary group $U(\infty)$. The four-parameter family of zw -measures arises naturally in a problem of harmonic analysis on $U(\infty)$ as the decomposing measures for a distinguished family of characters [44]. A zw -measure gives rise to a determinantal point process on the real line with two punctures, and the corresponding correlation kernel is given in terms of the Gauss hypergeometric function [11]. Such point processes degenerate, via suitable limits and/or specializations, to essentially all known one-dimensional determinantal processes with correlation kernels expressible through classical special functions.

The problem of constructing a Markov process that preserves a given determinantal point process with infinite point configurations has been addressed in [50,29,47] for the sine process, in [28] for the Airy process, and in [46] for the Whittaker process describing the z -measures from the harmonic analysis on the infinite symmetric group.

Our approach to constructing the infinite-dimensional stochastic dynamics differs from the ones used in previous papers. We employ the fact (of representation theoretic origin) that the probability measures on Ω are in one-to-one correspondence with *central* or *Gibbs* measures on infinite Gelfand–Tsetlin schemes that can also be viewed as stepped surfaces or lozenge tilings of a half-plane. The projections of a zw -measure to suitably defined slices of the infinite schemes yield *orthogonal polynomial ensembles* with weight functions corresponding to hypergeometric Askey–Lesky orthogonal polynomials.

These orthogonal polynomials are eigenfunctions for a birth and death process on \mathbb{Z} with quadratic jump rates; a standard argument then shows that the N -dimensional Askey–Lesky orthogonal polynomial ensemble is preserved by a Doob’s h -transform of N independent birth and death processes.

We further show that the Markov processes on the slices are *consistent* with respect to stochastic projections of the N th slice to the $(N - 1)$ st one (these projections are uniquely determined by the Gibbs property). This consistency is in no way obvious, and we do not have a conceptual explanation for it. However, it turns out to be essentially sufficient for defining the corresponding Markov process on Ω .

We do a bit more — using a continuous time analog of the general formalism of [6] (which in turn was based on an idea from [18]), we construct a Markov process on Gelfand–Tsetlin schemes that preserves the class of central (= Gibbs) measures and that induces the same Markov process on Ω .

The ideas of the present work were applied in a different situation, when the boundary turns out to be a discrete space; see the recent paper [8].

We now proceed to a more detailed description of our work.

1.1. Gelfand–Tsetlin graph and its boundary

Following [55], for $N \geq 1$ define a *signature* of length N as an N -tuple of nonincreasing integers $\lambda = (\lambda_1 \geq \dots \geq \lambda_N)$, and denote by \mathbb{GT}_N the set of all such signatures. Elements of \mathbb{GT}_N parameterize irreducible representations of $U(N)$ or $GL(N, \mathbb{C})$ (the signatures serve as the *highest weights* of the corresponding representations).

For $\lambda \in \mathbb{GT}_N$ and $\nu \in \mathbb{GT}_{N+1}$, we say that $\lambda < \nu$ if $\nu_{j+1} \leq \lambda_j \leq \nu_j$ for all meaningful values of indices. These inequalities are well known to be equivalent to the condition that the restriction of the ν -representation of $U(N+1)$ to $U(N)$ contains a λ -component; see e.g. [23, Chapter 8], [57, Section 66, Theorem 2].

Set $\mathbb{GT} = \bigsqcup_{N \geq 1} \mathbb{GT}_N$, and equip \mathbb{GT} with edges by joining λ and ν iff $\lambda < \nu$ or $\nu < \lambda$. This turns \mathbb{GT} into a graph that we call the *Gelfand–Tsetlin graph*. A path of length $M \in \{1, 2, \dots\} \cup \{\infty\}$ in \mathbb{GT} is a length M sequence

$$\lambda^{(1)} < \lambda^{(2)} < \dots, \quad \lambda^{(j)} \in \mathbb{GT}_j.$$

Equivalently, such a path can be viewed as an array of numbers $\{\lambda_i^{(j)}\}$ satisfying the inequalities $\lambda_{i+1}^{(j+1)} \leq \lambda_i^{(j)} \leq \lambda_i^{(j+1)}$; it is also called a *Gelfand–Tsetlin scheme*. An interpretation of paths in \mathbb{GT} in terms of lozenge tilings or stepped surfaces can be found in the introduction of [6].

The Gelfand–Tsetlin schemes of length N parameterize the vectors in the *Gelfand–Tsetlin basis* of the irreducible representation of $U(N)$ corresponding to $\lambda^{(N)}$; see [57]. Denote by $\text{Dim}_N \lambda$ the number of such schemes with $\lambda^{(N)} = \lambda$; this is also the dimension of the irreducible representation of $U(N)$ corresponding to λ . It is essentially equal to the Vandermonde determinant in shifted coordinates of λ :

$$\text{Dim}_N(\lambda) = \text{const}_N \prod_{1 \leq i < j \leq N} (\lambda_i - i - \lambda_j + j).$$

A probability measure on infinite paths in \mathbb{GT} is called *central* (or *Gibbs*) if any two finite paths with the same top end are equiprobable, cf. [31]. Let μ_N be the projection of such a measure to \mathbb{GT}_N . Centrality is easily seen to be equivalent to the relation $\mu_N = \mu_{N+1} \Lambda_N^{N+1}$, $N \geq 1$, where μ_N and μ_{N+1} are viewed as row-vectors with coordinates $\{\mu_N(\lambda)\}_{\lambda \in \mathbb{GT}_N}$ and $\{\mu_{N+1}(\nu)\}_{\nu \in \mathbb{GT}_{N+1}}$, and

$$\Lambda_N^{N+1}(\nu, \lambda) = \frac{\text{Dim}_N(\lambda)}{\text{Dim}_N(\nu)} \mathbf{1}_{\lambda < \nu}, \quad \lambda \in \mathbb{GT}_N, \quad \nu \in \mathbb{GT}_{N+1}, \quad (1.1)$$

is the stochastic matrix of *cotransition probabilities*. There is a one-to-one correspondence between central measures on \mathbb{GT} and characters of $U(\infty)$ (equivalently, equivalence classes of unitary spherical representations of the Gelfand pair $(U(\infty) \times U(\infty), \text{diag } U(\infty))$), see [44].

As shown in [44, Section 9] (see also [53, 52, 43, 17]), the space of all central probability measures is isomorphic to the space of all probability measures on the set $\Omega \subset \mathbb{R}_+^{4\infty+2}$ consisting of the sextuples $\omega = (\alpha^+, \beta^+, \alpha^-, \beta^-, \delta^+, \delta^-) \in \mathbb{R}_+^{4\infty+2}$ satisfying the conditions

$$\begin{aligned} \alpha^\pm &= (\alpha_1^\pm \geq \alpha_2^\pm \geq \dots \geq 0), & \beta^\pm &= (\beta_1^\pm \geq \beta_2^\pm \geq \dots \geq 0), & \delta^\pm &\geq 0, \\ & \sum_{i=1}^{\infty} (\alpha_i^\pm + \beta_i^\pm) \leq \delta^\pm, & \beta_1^+ + \beta_1^- &\leq 1. \end{aligned}$$

The set Ω is called the *boundary* of \mathbb{GT} ; its points parameterize the *extreme* characters of $U(\infty)$. The map from central measures on \mathbb{GT} to measures on Ω amounts to certain asymptotic relations described in Section 9.1 below.

1.2. zw -measures

Let z, z', w, w' be four complex parameters such that

$$(z + k)(z' + k) > 0 \quad \text{and} \quad (w + k)(w' + k) > 0 \quad \text{for any } k \in \mathbb{Z} \quad (1.2)$$

and

$$z + z' + w + w' > -1 \quad (1.3)$$

(note that (1.2) implies that $z + z'$ and $w + w'$ are real). For $N \geq 1$, define a probability measure on \mathbb{GT}_N by (below $l_i = \lambda_i + N - i$)

$$M_{z,z',w,w'|N}(\lambda) = \text{const}_N \prod_{1 \leq i < j \leq N} (l_i - l_j)^2 \prod_{i=1}^N W_{z,z',w,w'}(l_i), \quad (1.4)$$

where

$$W_{z,z',w,w'}(x) = \frac{1}{\Gamma(z + N - x)\Gamma(z' + N - x)\Gamma(w + 1 + x)\Gamma(w' + 1 + x)}.$$

We call it the N th level zw -measure. It is the N -point orthogonal polynomial ensemble with weight $W_{z,z',w,w'}(\cdot)$, see e.g. [33] and the references therein for general information on such ensembles.

One can show that the finite level zw -measures are consistent: For any $N \geq 1$, $M_{z,z',w,w'|N} = M_{z,z',w,w'|N+1} A_N^{N+1}$. Therefore, the collection $\{M_{z,z',w,w'|N}\}_{N \geq 1}$ defines a central measure on the paths in \mathbb{GT} and a character of $U(\infty)$. For $z' = \bar{z}$, $w' = \bar{w}$, this character corresponds to a remarkable substitute for the nonexistent regular representation of $U(\infty)$, see [44] for details.

The corresponding measure $M_{z,z',w,w'}$ on Ω is called the *spectral* zw -measure. If $\omega = (\alpha^\pm, \beta^\pm, \delta^\pm) \in \Omega$ is distributed according to $M_{z,z',w,w'}$ then the random point process generated by the coordinates

$$\left\{ \frac{1}{2} + \alpha_i^+, \frac{1}{2} - \beta_i^+, -\frac{1}{2} + \beta_i^-, -\frac{1}{2} - \alpha_i^- \right\}_{i=1}^\infty$$

is determinantal, see [11,12] for details.

1.3. Doob's transforms of N -fold products of birth and death processes

It is not hard to show that the first level zw -measure $M_{z,z',w,w'|1}$ on $\mathbb{GT}_1 = \mathbb{Z}$ is the symmetrizing measure for the bilateral birth and death process that from a point $x \in \mathbb{Z}$ jumps to the right with intensity $(x - u)(x - u')$ and jumps to the left with intensity $(x + v)(x + v')$, where $(u, u', v, v') = (z, z', w, w')$. Denote by $\mathcal{D} = \mathcal{D}_{u,u',v,v'}$ the corresponding matrix of transition rates.

More generally, we show that the N th level zw -measure (1.4) is the symmetrizing measure for a continuous time Markov chain on \mathbb{GT}_N with transition rates

$$\mathcal{D}^{(N)}(\lambda, v) = \frac{\text{Dim}_N(v)}{\text{Dim}_N(\lambda)} \left(\mathcal{D}(l_1, n_1) \mathbf{1}_{\{l_i=n_i, \forall i \neq 1\}} + \mathcal{D}(l_2, n_2) \mathbf{1}_{\{l_i=n_i, \forall i \neq 2\}} + \cdots \right. \\ \left. + \mathcal{D}(l_N, n_N) \mathbf{1}_{\{l_i=n_i, \forall i \neq N\}} \right) - d_N \cdot \mathbf{1}_{\lambda=v} \quad (1.5)$$

where $l_j = \lambda_j + N - j$, $n_j = v_j + N - j$, $1 \leq j \leq N$, d_N is a suitable constant, and we take $(u, u', v, v') = (z + N - 1, z' + N - 1, w, w')$ in the definition of \mathcal{D} .

Observe that $\mathcal{D}^{(N)}$ can be viewed as a version of Doob's h -transform of N copies of the Markov chain defined by \mathcal{D} with $h(\cdot) = \text{Dim}_N(\cdot)$. Note that in our case, $\text{Dim}_N(\cdot)$ is an eigenfunction of the corresponding matrix of transition rates with a *nonzero* eigenvalue.

For any $N \geq 1$, let $(P_N(t))_{t \geq 0}$ be the Markov semigroup corresponding to the matrix $\mathcal{D}^{(N)}$ of transition rates on \mathbb{GT}_N (we show that $(P_N(t))_{t \geq 0}$ is uniquely defined and it possesses the Feller property). The key fact that we prove is the consistency (or commutativity) relation

$$P_{N+1}(t) \Lambda_N^{N+1} = \Lambda_N^{N+1} P_N(t), \quad t \geq 0, N \geq 1.$$

Although this relation looks natural, we have no *a priori* reason to expect it to hold, and we verify it by a brute force computational argument.

1.4. Main result

We prove that for any $(z, z', w, w') \in \mathbb{C}^4$ subject to conditions (1.2) and (1.3), there exists a unique Markov semigroup $(P(t))_{t \geq 0}$ on Ω that preserves the spectral zw -measure $M_{z,z',w,w'}$, and whose trace on \mathbb{GT}_N coincides with Doob's transforms $(P_N(t))_{t \geq 0}$ introduced above. Moreover, the semigroup $(P(t))_{t \geq 0}$ is Feller (it preserves $C_0(\Omega)$, the Banach space of continuous functions vanishing at infinity; note that the space Ω is locally compact).

By general theory, see e.g. [19, Ch. 4, Theorem 2.7], this means that for any probability measure μ on Ω , there exists a Markov process on Ω corresponding to $(P(t))_{t \geq 0}$ with initial distribution μ and càdlàg sample paths. We also show that $M_{z,z',w,w'}$ is the unique invariant measure for this Markov process.

1.5. Markov process on Gelfand–Tsetlin schemes

Via the correspondence between the probability measures on Ω and central measures on paths in \mathbb{GT} , the semigroup $(P(t))_{t \geq 0}$ defines a Markov evolution of central measures. It is natural to ask if there exists a Markov process on *all* probability measures on the path space in \mathbb{GT} that agrees with the one we have when restricted to the central measures. We construct one such process; let us describe its transition rates.

Let $\{\lambda_i^{(j)}\}$ be a starting Gelfand–Tsetlin scheme. Then

- Each coordinate $\lambda_i^{(k)}$ tries to jump to the right by 1 with rate

$$(\lambda_i^{(k)} - i - z + 1)(\lambda_i^{(k)} - i - z' + 1)$$

and to the left by 1 with rate

$$(\lambda_i^{(k)} + k - i + w)(\lambda_i^{(k)} + k - i + w'),$$

independently of other coordinates.

- If the $\lambda_i^{(k)}$ -clock of the right jump rings but $\lambda_i^{(k)} = \lambda_{i-1}^{(k-1)}$, the jump is blocked. If its left clock rings but $\lambda_i^{(k)} = \lambda_i^{(k-1)}$, the jump is also blocked. (If any of the two jumps were allowed then the resulting set of coordinates would not have corresponded to a path in \mathbb{GT} .)
- If the right $\lambda_i^{(k)}$ -clock rings and there is no blocking, we find the greatest number $l \geq k$ such that $\lambda_i^{(j)} = \lambda_i^{(k)}$ for $j = k, k+1, \dots, l$, and move all the coordinates $\{\lambda_i^{(j)}\}_{j=k}^l$ to the right by one. Given the change $\lambda_i^{(k)} \mapsto \lambda_i^{(k)} + 1$, this is the minimal modification of the initial Gelfand–Tsetlin scheme that preserves interlacing.
- If the left $\lambda_i^{(k)}$ -clock rings and there is no blocking, we find the greatest number $l \geq k$ such that $\lambda_{i+j-k}^{(j)} = \lambda_i^{(k)}$ for $j = k, k+1, \dots, l$, and move all the coordinates $\{\lambda_{i+j-k}^{(j)}\}_{j=k}^l$ to the left by one. Again, given the change $\lambda_i^{(k)} \mapsto \lambda_i^{(k)} - 1$, this is the minimal modification of the set of coordinates that preserves interlacing.

Since the update rule for each coordinate $\lambda_i^{(k)}$ typically depends only on a few surrounding coordinates, one can argue that we have a model of *local* random growth. It should be compared to the models treated in [6,10], where a similar block-push mechanism was considered with constant jumps rates, and in [3], where the jump rates were also dependent on the location and numbering of the coordinates.

The key new feature of the Markov process above is the *absence of the limit shape phenomenon*. Often taken for granted in local growth models, it is simply nonexistent here.

This fact becomes more apparent if we restrict ourselves to coordinates $\{\lambda_1^{(j)}\}_{j \geq 1}$ only. The evolution of this set of coordinates is also Markov, and it represents a kind of an exclusion process. Our results imply that this process has a unique equilibrium measure. Moreover, with respect to this measure, the asymptotic density $(\lim_{j \rightarrow \infty} \lambda_1^{(j)} / j)^{-1}$ is well-defined and *random*. It changes over time, and its distribution is given by a solution to the classical Painlevé VI (second order nonlinear) differential equation, cf. [4].

1.6. Further work

While our main goal in this paper is to present a rigorous construction of the infinite-dimensional Markov dynamics we are interested in, in a subsequent publication we will analyze the dynamics in some detail. More exactly, we will show how to write the generator of the semigroup $(P(t))$ as a second order differential operator in countably many “coordinates” in Ω (those *are not* the (α, β, δ) coordinates from Section 1.1 above). We will also prove that under appropriate embeddings of the slices \mathbb{GT}_N into the boundary, the Markov semigroups $(P_N(t))_{t \geq 0}$ converge, in the sense of finite-dimensional distributions, to $(P(t))_{t \geq 0}$ as $N \rightarrow \infty$.

This paves the way to proving that the equilibrium Markov process on the boundary can be described as a time-dependent determinantal point process. The fact that the dynamical correlation functions are determinantal on each \mathbb{GT}_N easily follows from known techniques, although deriving useful formulas for the correlation kernel is a separate task.

1.7. Organization of the paper

In Section 2 we present an abstract scheme of constructing a Markov semigroup on the boundary out of a consistent family of semigroups on the slices. In Section 3 we describe how the Gelfand–Tsetlin graph fits into this abstract scheme. Section 4 is a brief collection of general

facts about continuous time Markov chains on countable spaces. Section 5 provides the construction of the Markov chains on \mathbb{GT}_N 's. In Section 6 we verify the consistency of these Markov chains. Section 7 contains a brief description of the zw -measures. In Section 8 we develop a general formalism of building continuous time Markov chains on paths out of a consistent family of those on the slices. In Section 9 we apply this formalism to our specific example and discuss the exclusion type processes.

2. Abstract construction

2.1. Markov kernels

For a more detailed exposition, see e.g. [40, Ch. IX].

Let E and E' be measurable spaces. A *Markov kernel* $K : E \rightarrow E'$ is a function $K(x, A)$, where $x \in E$ and $A \subset E'$ is a measurable subset, such that $K(x, \cdot)$ is a probability measure on E' and $K(\cdot, A)$ is a measurable function on E .

Let $\mathcal{B}(E)$ and $\mathcal{B}(E')$ denote the Banach spaces of real-valued bounded measurable functions with the sup-norm on E and E' , respectively. A Markov kernel $K : E \rightarrow E'$ induces a linear operator $\mathcal{B}(E') \rightarrow \mathcal{B}(E)$ of norm 1 via $(Kf)(x) = \int_{E'} K(x, dy) f(y)$.

For two Markov kernels $K_1 : E \rightarrow E'$ and $K_2 : E' \rightarrow E''$, their composition $K_1 \circ K_2 : E \rightarrow E''$ is also a Markov kernel.

Denote by $\mathcal{M}(\cdot)$ the Banach space of signed measures of bounded variation with the norm given by the total variation. Let $\mathcal{M}_+(\cdot)$ be the cone of finite positive measures, and let $\mathcal{M}_p(\cdot)$ be the simplex of the probability measures.

A Markov kernel $K : E \rightarrow E'$ also induces a linear operator $\mathcal{M}(E) \rightarrow \mathcal{M}(E')$ of norm 1 via $(\mu K)(dy) = \int_E \mu(dx) K(x, dy)$. This operator maps $\mathcal{M}_+(E)$ to $\mathcal{M}_+(E')$ and $\mathcal{M}_p(E)$ to $\mathcal{M}_p(E')$. Note that $\delta_x K = K(x, \cdot)$, where δ_x is the Dirac delta-measure at $x \in E$.

The space $\mathcal{M}(E)$ (and hence $\mathcal{M}_+(E)$ and $\mathcal{M}_p(E)$) is equipped with a σ -algebra of measurable sets: Any preimage of a Borel set under the map $\mu \mapsto \mu(A)$ from $\mathcal{M}(E)$ to \mathbb{R} for any measurable A is measurable.

2.2. Feller kernels

Let E and E' be locally compact topological spaces with countable bases. Let us take Borel σ -algebra as the σ -algebra of measurable sets for both of them.

Let $C(\cdot) \subset \mathcal{B}(\cdot)$ be the Banach space of bounded continuous functions, and let $C_0(\cdot) \subset C(\cdot)$ be its subspace of functions that tend to 0 at infinity.

Definition 2.1. A Markov kernel $K : E \rightarrow E'$ is called *Feller* if the induced map $\mathcal{B}(E') \rightarrow \mathcal{B}(E)$ maps $C_0(E')$ to $C_0(E)$.

Note that different authors may use different (nonequivalent) definitions for the Feller property.

The convenience of the space $C_0(\cdot)$ is based on the fact that this space is separable (as opposed to $C(\cdot)$ which is not separable, except in the case when the initial topological space is compact), and $\mathcal{M}(\cdot)$ is its Banach dual.

2.3. Feller semigroups

A *Markov semigroup* is a family of Markov kernels $P(t) : E \rightarrow E$, where $t \geq 0$, $P(0) = 1$ (in the obvious sense), and $P(s)P(t) = P(s+t)$. Such a semigroup induces a semigroup of linear operators in $\mathcal{B}(E)$ as well as a semigroup of linear operators in $\mathcal{M}(E)$, see above.

We say that a Markov semigroup $(P(t))_{t \geq 0}$ is *Feller* if

- E is a locally compact topological space with countable base;
- the corresponding operator semigroup in $\mathcal{B}(E)$ preserves $C_0(E)$;
- the function $t \mapsto P(t)$ is strongly continuous, i.e. $t \mapsto P(t)f$ is a continuous map from $[0, +\infty)$ to $C_0(E)$ for any $f \in C_0(E)$ (an equivalent condition is the continuity at $t = 0$).

2.4. Feller semigroups and Markov processes

For more details, see e.g. [19, Ch. 4, Theorem 2.7].

Let E be a locally compact separable metric space, and let $(P(t))_{t \geq 0}$ be a Feller semigroup on E . Then for each $\mu \in \mathcal{M}_p(E)$, there exists a Markov process corresponding to $(P(t))_{t \geq 0}$ with initial distribution μ and càdlàg sample paths. Moreover, this process is strong Markov.

2.5. Boundary

Let E_1, E_2, \dots be a sequence of measurable spaces linked by Markov kernels

$$\Lambda_N^{N+1} : E_{N+1} \rightarrow E_N, \quad N = 1, 2, \dots$$

Assume that we have another measurable space E_∞ and Markov kernels

$$\Lambda_N^\infty : E_\infty \rightarrow E_N, \quad N = 1, 2, \dots,$$

such that the natural commutativity relations hold:

$$\Lambda_{N+1}^\infty \circ \Lambda_N^{N+1} = \Lambda_N^\infty, \quad N = 1, 2, \dots \quad (2.1)$$

The kernels Λ_N^{N+1} induce the chain of maps, cf. Section 2.1,

$$\dots \rightarrow \mathcal{M}_p(E_{N+1}) \rightarrow \mathcal{M}_p(E_N) \rightarrow \dots \rightarrow \mathcal{M}_p(E_2) \rightarrow \mathcal{M}_p(E_1), \quad (2.2)$$

and we can define the projective limit $\varprojlim \mathcal{M}_p(E_N)$ with respect to these maps. By definition, it consists of sequences of measures $(\mu_N)_{N \geq 1}$, $\mu_N \in \mathcal{M}_p(E_N)$, that are linked by the maps from (2.2). The space $\varprojlim \mathcal{M}_p(E_N)$ is measurable; the σ -algebra of measurable sets is generated by the cylinder sets in which μ_N must lie inside a measurable subset of $\mathcal{M}_p(E_N)$, and all other coordinates $(\mu_k)_{k \neq N}$, are unrestricted.

Observe that to any $\mu_\infty \in \mathcal{M}_p(E_\infty)$ one can assign an element of $\varprojlim \mathcal{M}_p(E_N)$ by setting μ_N equal to the image of μ_∞ under the map $\mathcal{M}_p(E_\infty) \rightarrow \mathcal{M}_p(E_N)$ induced by Λ_N^∞ . The commutativity relations (2.1) ensure that the resulting sequence $(\mu_N)_{N \geq 1}$ is consistent with (2.2).

Definition 2.2. We say that E_∞ is a *boundary* of the sequence $(E_N)_{N \geq 1}$ if the map $\mathcal{M}_p(E_\infty) \rightarrow \varprojlim \mathcal{M}_p(E_N)$ described in the previous paragraph is a bijection and also an isomorphism of measurable spaces.

2.6. Feller boundary

In the setting of the previous subsection, let us further assume that $(E_N)_{N \geq 1}$ and E_∞ are locally compact topological spaces with countable bases, and all the links $(\Lambda_N^{N+1})_{N \geq 1}$, $(\Lambda_N^\infty)_{N \geq 1}$ are Feller kernels, cf. Section 2.2. Then if E_∞ satisfies Definition 2.2, we shall call it the *Feller boundary* for $(E_N)_{N \geq 1}$.

According to Section 2.2, the links $(\Lambda_N^\infty)_{N \geq 1}$ induce linear operators $C_0(E_N) \rightarrow C_0(E_\infty)$.

Lemma 2.3. *The union of images of these maps over all $N \geq 1$ is dense in the Banach space $C_0(E_\infty)$.*

Proof. Since $\mathcal{M}(E_\infty)$ is the Banach dual to $C_0(E_\infty)$, it suffices to verify that if $\mu \in \mathcal{M}(E_\infty)$ kills all functions in our union then $\mu = 0$.

Assume μ is a signed measure on E_∞ that kills the image of $C_0(E_N)$, $N \geq 1$. This is equivalent to saying that $\mu K_N^\infty = 0$ for all $N \geq 1$. We can represent μ as difference of finite positive measures

$$M = \alpha\mu' - \beta\mu'', \quad \mu', \mu'' \in \mathcal{M}_p(E_\infty), \quad \alpha, \beta \geq 0.$$

Hence, $\alpha\mu' K_N^\infty = \beta\mu'' K_N^\infty$ for all $N \geq 1$. Since $\mu' K_N^\infty$ and $\mu'' K_N^\infty$ are in $\mathcal{M}_p(E_N)$, we must have $\alpha = \beta$, and $\mu' K_N^\infty = \mu'' K_N^\infty$. Definition 2.2 implies $\mu' = \mu''$, thus $\mu = 0$. \square

2.7. Extension of semigroups to the boundary

In the setting of Section 2.5, assume that for any $N \geq 1$, we have a Markov semigroup $(P_N(t))_{t \geq 0}$ on E_N , and these semigroups are compatible with the links:

$$P_{N+1}(t) \circ \Lambda_N^{N+1} = \Lambda_N^{N+1} \circ P_N(t), \quad t \geq 0, \quad N = 1, 2, \dots \quad (2.3)$$

Proposition 2.4. *In the above assumptions, there exists a unique Markov semigroup $P(t)$ on E_∞ such that*

$$P(t) \circ \Lambda_N^\infty = \Lambda_N^\infty \circ P_N(t), \quad t \geq 0, \quad N = 1, 2, \dots \quad (2.4)$$

If E_∞ is Feller (cf. Section 2.6) and $(P_N(t))_{t \geq 0}$ is a Feller semigroup for any $N \geq 1$, then $(P(t))_{t \geq 0}$ is also a Feller semigroup.

Proof. Denote by δ_x the delta-measure at a point $x \in E_\infty$. To construct the semigroup $(P(t))_{t \geq 0}$, we need to define, for any $t \geq 0$, a probability measure $P(t; x, \cdot)$ on E_∞ . This measure has to satisfy

$$P(t; x, \cdot) \Lambda_N^\infty = \delta_x (\Lambda_N^\infty \circ P_N(t)), \quad N \geq 1.$$

The right-hand side defines a sequence of probability measures on E_N 's, and (2.1), (2.3) immediately imply that these measures are compatible with maps (2.2). Hence, we obtain an element of $\varprojlim \mathcal{M}_p(E_N)$, which defines, by definition of the boundary, a probability measure on E_∞ . The dependence of this measure on x is measurable since this is true for any of its coordinates.

Thus, we have obtained a Markov kernel $P(t)$ which satisfies

$$\delta_x(P(t) \circ \Lambda_N^\infty) = \delta_x(\Lambda_N^\infty \circ P_N(t)), \quad N \geq 1,$$

which is equivalent to (2.4).

To verify the semigroup property (Chapman–Kolmogorov equation) for $(P(t))_{t \geq 0}$ it suffices to check that

$$(P(s) \circ P(t)) \circ \Lambda_N^\infty = P(s+t) \circ \Lambda_N^\infty, \quad s, t \geq 0, \quad N \geq 1,$$

and this immediately follows from (2.4) and the corresponding relation for $(P_N(t))_{t \geq 0}$.

The uniqueness is obvious since $P(t)$ is uniquely determined by $(P(t) \circ \Lambda_N^\infty)_{N \geq 1}$ that are given by (2.4).

Finally, let us prove the Feller property assuming that the boundary is Feller and all $(P_N(t))_{t \geq 0}$ are Feller.

We need to show that for $f \in C_0(E_\infty)$ we have $P(t)f \in C_0(E_\infty)$, and that $P(t)f$ is continuous in t in the topology of $C_0(E_\infty)$. Both properties can be verified on a dense subset. Lemma 2.3 then shows that it suffices to consider f of the form $f = \Lambda_N^\infty f_N$ with $f_N \in C_0(E_N)$. By (2.4)

$$P(t)f = P(t)(\Lambda_N^\infty f_N) = \Lambda_N^\infty(P_N(t)f_N),$$

which is in $C_0(E_\infty)$ because Λ_N^∞ and $P_N(t)$ are Feller. The continuity in t is obvious as $P_N(t)f_N$ is continuous in t , and $\Lambda_N^\infty : C_0(E_N) \rightarrow C_0(E_\infty)$ is a contraction. \square

It is worth noting that our definition of the semigroup $P(t)$ is nonconstructive: We are not able to describe $P(t; x, A)$ explicitly, and we have to appeal to the isomorphism in Definition 2.2 instead. Thus, the difficulty in making $P(t)$ explicit is hidden in the implicit nature of that isomorphism.

2.8. Invariant measures

In the setting of Section 2.5, assume that for any $N \geq 1$, there exists $\mu_N \in \mathcal{M}_p(E_N)$ such that $\mu_N P_N(t) = \mu_N$ (i.e., μ_N is an invariant measure for $(P_N(t))_{t \geq 0}$). If we assume that μ_N 's are compatible with the links,

$$\mu_{N+1} \Lambda_N^{N+1} = \mu_N, \quad N \geq 1,$$

then, via Definition 2.2, they yield a measure $\mu \in \mathcal{M}_p(E_\infty)$ such that $\mu \Lambda_N^\infty = \mu_N$ for any $N \geq 1$. Note that μ is uniquely determined by the sequence $\{\mu_N\}$.

One easily sees that μ is invariant with respect to $(P(t))_{t \geq 0}$. Indeed,

$$(\mu P(t)) \Lambda_N^\infty = (\mu \Lambda_N^\infty) P_N(t) = \mu_N P_N(t) = \mu_N = \mu \Lambda_N^\infty.$$

Moreover, if μ_N is a unique invariant measure for $(P_N(t))_{t \geq 0}$ for any $N \geq 1$ then the invariant measure for $(P(t))_{t \geq 0}$ is unique too as its convolution with Λ_N^∞ must coincide with μ_N .

3. Specialization. Gelfand–Tsetlin graph

3.1. Spaces and links

Let N be a positive integer. A *signature* λ of length N is an N -tuple of weakly decreasing integers: $\lambda = (\lambda_1 \geq \dots \geq \lambda_N) \in \mathbb{Z}^N$. Denote by \mathbb{GT}_N the set of all signatures of length N (the notation \mathbb{GT} is explained below). This countable set will serve as our space E_N from the previous section.

Signatures of length N parameterize irreducible representations of the unitary group $U(N)$ and are often referred to as *highest weights*, cf. [55,57]. For $\lambda \in \mathbb{GT}_N$ denote the corresponding representation by π_λ , and denote by $\text{Dim}_N \lambda$ the dimension of the corresponding linear space. It is well known that

$$\text{Dim}_N \lambda = \frac{\prod_{1 \leq i < j \leq N} (\lambda_i - i - \lambda_j + j)}{\prod_{i=1}^{N-1} i!}, \quad \lambda \in \mathbb{GT}_N.$$

Define a matrix $[A_N^{N+1}(\lambda, \nu)]_{\lambda \in \mathbb{GT}_{N+1}, \nu \in \mathbb{GT}_N}$ with rows parameterized by \mathbb{GT}_{N+1} and columns parameterized by \mathbb{GT}_N via

$$A_N^{N+1}(\lambda, \nu) = \begin{cases} N! \cdot \frac{\prod_{1 \leq i < j \leq N} (\nu_i - i - \nu_j + j)}{\prod_{1 \leq i < j \leq N+1} (\lambda_i - i - \lambda_j + j)}, & \text{if } \nu \prec \lambda, \\ 0, & \text{otherwise,} \end{cases}$$

where the notation $\nu \prec \lambda$ stands for interlacing:

$$\nu \prec \lambda \iff \lambda_1 \geq \nu_1 \geq \lambda_2 \geq \nu_2 \geq \dots \geq \nu_N \geq \lambda_{N+1}.$$

Note that the nonzero entries of A_N^{N+1} can also be written in the form

$$A_N^{N+1}(\lambda, \nu) = \frac{\text{Dim}_N \nu}{\text{Dim}_{N+1} \lambda}. \quad (3.1)$$

It is not hard to show that A_N^{N+1} is a stochastic matrix: $\sum_{\nu \in \mathbb{GT}_N} A_N^{N+1}(\lambda, \nu) = 1$ for any $\lambda \in \mathbb{GT}_{N+1}$. Indeed, $\text{Dim}_{N+1} \lambda$ is equal to the number of the sequences (known as *Gelfand–Tsetlin schemes*, thus the notation \mathbb{GT})

$$\lambda^{(1)} \prec \lambda^{(2)} \prec \dots \prec \lambda^{(N+1)} = \lambda, \quad \lambda^{(j)} \in \mathbb{GT}_j,$$

and $A_N^{N+1}(\lambda, \nu)$ is the fraction of the sequences with $\lambda^{(N)} = \nu$. The stochasticity also follows from the branching rule for the representations of unitary groups: For any $\lambda \in \mathbb{GT}_{N+1}$,

$$\pi_\lambda|_{U(N)} \sim \bigoplus_{\nu \in \mathbb{GT}_N: \nu \prec \lambda} \pi_\nu.$$

The matrices Λ_N^{N+1} viewed as Markov kernels $\Lambda_N^{N+1} : \mathbb{GT}_{N+1} \rightarrow \mathbb{GT}_N$ are our links, cf. Section 2.5. Set $\mathbb{GT} = \bigsqcup_{N \geq 1} \mathbb{GT}_N$. We endow \mathbb{GT} with the structure of a graph: Two vertices λ and ν are joined by an edge if and only if $\nu < \lambda$ or $\lambda < \nu$. This graph is called the *Gelfand–Tsetlin graph*, and the matrix elements of the links are often called *cotransition probabilities* for this graph, cf. [31].

3.2. Boundary

Let $\mathbb{R}_+ \subset \mathbb{R}$ be the set of nonnegative real numbers and \mathbb{R}_+^∞ be the product of countably many copies of \mathbb{R}_+ . Consider the space

$$\mathbb{R}_+^{4\infty+2} := \mathbb{R}_+^\infty \times \mathbb{R}_+^\infty \times \mathbb{R}_+^\infty \times \mathbb{R}_+^\infty \times \mathbb{R}_+ \times \mathbb{R}_+$$

and equip it with the product topology. We choose E_∞ to be the closed subset $\Omega \subset \mathbb{R}_+^{4\infty+2}$ consisting of the sextuples

$$\omega = (\alpha^+, \beta^+, \alpha^-, \beta^-, \delta^+, \delta^-) \in \mathbb{R}_+^{4\infty+2}$$

satisfying the conditions

$$\begin{aligned} \alpha^\pm &= (\alpha_1^\pm \geq \alpha_2^\pm \geq \dots), \quad \beta^\pm = (\beta_1^\pm \geq \beta_2^\pm \geq \dots), \quad \delta^\pm \geq 0, \\ \sum_{i=1}^\infty (\alpha_i^\pm + \beta_i^\pm) &\leq \delta^\pm, \quad \beta_1^+ + \beta_1^- \leq 1. \end{aligned}$$

One easily sees that Ω is a locally compact metrizable topological space with a countable base. We endow Ω with the corresponding Borel structure which makes Ω a measurable space.

It will be convenient to use the notation

$$\gamma^\pm = \delta^\pm - \sum_{i=1}^\infty (\alpha_i^\pm + \beta_i^\pm) \geq 0.$$

Define the projections/links $\Lambda_N^\infty : \Omega \rightarrow \mathbb{GT}_N$, $N \geq 1$, by

$$\Lambda_N^\infty(\omega, \lambda) = \text{Dim}_N \lambda \cdot \det[\varphi_{\lambda_i - i + j}]_{i,j=1}^N, \quad \omega \in \Omega, \lambda \in \mathbb{GT}_N, \quad (3.2)$$

where $\{\varphi_n\}_{n=-\infty}^{+\infty}$ are the Laurent coefficients of the following function in variable u , $|u| = 1$:

$$\Phi(u; \omega) := e^{\gamma^+(u-1) + \gamma^-(u^{-1}-1)} \prod_{i=1}^\infty \frac{1 + \beta_i^+(u-1)}{1 - \alpha_i^+(u-1)} \frac{1 + \beta_i^-(u^{-1}-1)}{1 - \alpha_i^-(u^{-1}-1)} = \sum_{n=-\infty}^{+\infty} \varphi_n u^n. \quad (3.3)$$

Theorem 3.1. *The space $E_\infty = \Omega$ is the boundary of the chain of spaces $(E_N = \mathbb{GT}_N)_{N \geq 1}$ with links as above in the sense of Definition 2.2.*

Proof. It is readily verified that the families of links $\{\Lambda_N^\infty\}$ and $\{\Lambda_N^{N+1}\}$ satisfy the commutativity relations (2.1), see [17, Proposition 2.9(ii)]. Therefore, we get a map

$$\mathcal{M}_p(\Omega) \rightarrow \varprojlim \mathcal{M}_p(\mathbb{GT}_N). \quad (3.4)$$

Next, we claim that (3.4) is a bijection. Indeed, this (nontrivial) result is precisely the assertion of [17, Theorem 2.15]. Alternatively, it is essentially proved in [44, Section 9].

Finally, we have to prove that (3.4) is a Borel isomorphism. As shown in the proof of Theorem 8.1 of [44], the map $\omega \mapsto \Lambda_N^\infty(\omega, \lambda)$ is continuous for every $N = 1, 2, \dots$ and every $\lambda \in \mathbb{GT}_N$. This implies that the map (3.4) is Borel. To show that the inverse map is also Borel one can apply an abstract result (Theorem 3.2 in [39]), which asserts that a Borel one-to-one map of a standard Borel space onto a subset of a countably generated Borel space is a Borel isomorphism. This result is applicable in our situation, since the Borel structure of Ω is standard, so that the induced Borel structure on $\mathcal{M}_p(\Omega)$ is standard, too. \square

Remark 3.2. Observe that the maps on $(\mathbb{GT}_N)_{N \geq 1}$ consisting of shifts of all coordinates of signatures by 1,

$$\lambda = (\lambda_1, \dots, \lambda_N) \mapsto \tilde{\lambda} = (\tilde{\lambda}_1 = \lambda_1 + 1, \dots, \tilde{\lambda}_N = \lambda_N + 1),$$

leave the links intact: $\Lambda_N^{N+1}(\lambda, \nu) = \Lambda_N^{N+1}(\tilde{\lambda}, \tilde{\nu})$. There is also a corresponding homeomorphism of Ω , which amounts to the multiplication of the function $\Phi(u; \omega)$ by u : For $\omega = (\alpha^\pm, \beta^\pm, \delta^\pm) \in \Omega$ define $\tilde{\omega} = (\tilde{\alpha}^\pm, \tilde{\beta}^\pm, \tilde{\delta}^\pm) \in \Omega$ by

$$\begin{aligned} \tilde{\alpha}^\pm &= \alpha^\pm, \\ \tilde{\beta}_1^+ &= 1 - \beta_1^-, \quad (\tilde{\beta}_2^+, \tilde{\beta}_3^+, \dots) = (\beta_1^+, \beta_2^+, \dots), \quad (\tilde{\beta}_1^-, \tilde{\beta}_2^-, \dots) = (\beta_2^-, \beta_3^-, \dots) \\ \tilde{\delta}^+ &= \delta^+ + (1 - \beta_1^-), \quad \tilde{\delta}^- = \delta^- - \beta_1^- \end{aligned}$$

(note that $\tilde{\beta}_1^+ \geq \tilde{\beta}_2^+$ because $\beta_1^+ + \beta_1^- \leq 1$; the transformation of δ^\pm is such that γ^\pm remains intact). Then (3.2) and the relation

$$u(1 + \beta_1^-(u^{-1} - 1)) = 1 + (1 - \beta_1^-)(u - 1)$$

show that $\Lambda_N^\infty(\omega, \lambda) = \Lambda_N^\infty(\tilde{\omega}, \tilde{\lambda})$ for any $\lambda \in \mathbb{GT}_N$ and $N \geq 1$.

This automorphism of the Gelfand–Tsetlin graph and its boundary has a representation theoretic origin, cf. Remark 1.5 in [44] and Remark 3.7 in [11].

3.3. The boundary is Feller

Following definitions of Section 2.6, in order to show that $E_\infty = \Omega$ is a Feller boundary of the chain $(E_N = \mathbb{GT}_N)_{N \geq 1}$ we need to verify two statements:

- the spaces $(E_N)_{N \geq 1}$ and E_∞ are locally compact topological spaces with countable bases;
- the links $(\Lambda_N^{N+1})_{N \geq 1}$ and $(\Lambda_N^\infty)_{N \geq 1}$ are Feller kernels.

The first statement is obvious from the definitions. The goal of this subsection is to prove the second one.

Proposition 3.3. *For any $N \geq 1$, the linear operator $\mathcal{B}(\mathbb{GT}_N) \rightarrow \mathcal{B}(\mathbb{GT}_{N+1})$ induced by the Markov kernel Λ_N^{N+1} maps $C_0(\mathbb{GT}_N)$ to $C_0(\mathbb{GT}_{N+1})$.*

Proof. As the norm of the linear operator in question is equal to 1 and $C_0(\cdot)$ is a closed subspace of $\mathcal{B}(\cdot)$, it suffices to check that the images of all delta-functions on \mathbb{GT}_N are in $C_0(\mathbb{GT}_{N+1})$.

For an arbitrary $v \in \mathbb{GT}_N$, let δ_v be the delta-function on \mathbb{GT}_N concentrated at v . Then for $\lambda \in \mathbb{GT}_{N+1}$

$$(\Lambda_N^{N+1} \delta_v)(\lambda) = \begin{cases} N! \cdot \frac{\prod_{1 \leq i < j \leq N} (v_i - i - v_j + j)}{\prod_{1 \leq i < j \leq N+1} (\lambda_i - i - \lambda_j + j)}, & \text{if } v < \lambda, \\ 0, & \text{otherwise.} \end{cases}$$

If we assume that $(\Lambda_N^{N+1} \delta_v)(\lambda) \neq 0$ then $\lambda \rightarrow \infty$ is equivalent to either $\lambda_1 \rightarrow +\infty$, or $\lambda_{N+1} \rightarrow -\infty$, or both; all other coordinates must remain bounded because of the interlacement condition $v < \lambda$. But then it is immediate that at least one of the factors in the denominator in $(\Lambda_N^{N+1} \delta_v)(\lambda)$ tends to infinity. Thus, for any fixed $v \in \mathbb{GT}_N$, $(\Lambda_N^{N+1} \delta_v)(\lambda) \rightarrow 0$ as $\lambda \rightarrow \infty$ as needed. \square

Proposition 3.4. *For any $N \geq 1$, the linear operator $\mathcal{B}(\mathbb{GT}_N) \rightarrow \mathcal{B}(\Omega)$ induced by the Markov kernel Λ_N^∞ maps $C_0(\mathbb{GT}_N)$ to $C_0(\Omega)$.*

Proof. See [17, Proposition 2.10]. \square

4. Generalities on Markov chains on countable spaces

4.1. Regularity

Let E be a countable set, and let $(P(t))_{t \geq 0}$ be a Markov semigroup on E . Each $P(t)$ may be viewed as a matrix with rows and columns marked by elements of E ; its entries will be denoted by $P(t; a, b)$, $a, b \in E$. By definition, $P(t; a, b)$ is the probability that the process will be in the state b at the time moment t conditioned that it is in the state a at time 0. Thus, all matrix elements of $P(t)$ are nonnegative, and their sum is equal to 1 along any row meaning that the matrix $P(t)$ is stochastic. The transition matrices $P(t)$ also satisfy the Chapman–Kolmogorov equation $P(s)P(t) = P(s+t)$.

Assume that there exists an $E \times E$ matrix Q such that

$$P(t; a, b) = \mathbf{1}_{a=b} + Q(a, b)t + o(t), \quad t \downarrow 0. \quad (4.1)$$

This relation implies that $Q(a, b) \geq 0$ for $a \neq b$ and $Q(a, a) \leq 0$. Further, we will always assume that

$$\sum_{b: b \neq a} Q(a, b) = -Q(a, a) \quad \text{for any } a \in E.$$

This is the infinitesimal analog of the condition $\sum_{b \in E} P(t; a, b) = 1$.

It is well known that the Chapman–Kolmogorov equation implies that $P(t)$ satisfies *Kolmogorov’s backward equation*

$$\frac{d}{dt} P(t) = QP(t), \quad t > 0, \quad (4.2)$$

with the initial condition

$$P(0) = \text{Id}. \quad (4.3)$$

Under certain additional conditions, $P(t)$ will also satisfy *Kolmogorov’s forward equation*

$$\frac{d}{dt} P(t) = P(t)Q, \quad t > 0. \quad (4.4)$$

One says that Q is the *matrix of transition rates* for $(P(t))_{t \geq 0}$.

One often wants to define a Markov semigroup by giving the transition rates. However, it may happen that this does not specify the semigroup uniquely (then the backward equation has many solutions). Uniqueness always holds if E is finite or, more generally, if E is infinite but $\sup_{a \in E} |Q(a, a)| < \infty$. However, these simple conditions do not suit our purposes, and we need to go a little deeper into the general theory.

Let us write Q in the form $Q = -q + \tilde{Q}$, where $-q$ is the diagonal part of Q and \tilde{Q} is the off-diagonal part of Q . In other words,

$$q(a, b) = -Q(a, a)\mathbf{1}_{ab}, \quad \tilde{Q}(a, b) = \begin{cases} Q(a, b), & a \neq b, \\ 0, & a = b. \end{cases}$$

Define the matrices $P^{[n]}(t)$ recursively by

$$P^{[0]}(t) = e^{-tq}, \quad P^{[n]}(t) = \int_0^t e^{-\tau q} \tilde{Q} P^{[n-1]}(t - \tau) d\tau, \quad n \geq 1,$$

and set

$$\bar{P}(t) = \sum_{n=0}^{\infty} P^{[n]}(t), \quad t \geq 0.$$

Theorem 4.1. (See [20].) (i) The matrix $\bar{P}(t)$ is substochastic (i.e., its elements are nonnegative and $\sum_b P(t; a, b) \leq 1$). Its elements are continuous in $t \in [0, +\infty)$ and differentiable in $t \in (0, +\infty)$, and it provides a solution of Kolmogorov’s backward and forward equations (4.2), (4.4) with the initial condition (4.3).

(ii) $\bar{P}(t)$ also satisfies the Chapman–Kolmogorov equation.

(iii) $\bar{P}(t)$ is the minimal solution of (4.2) (or (4.4)) in the sense that for any other solution $P(t)$ of (4.2) (or (4.4)) with the initial condition (4.3) in the class of substochastic matrices, one has $P(t; a, b) \geq \bar{P}(t; a, b)$ for any $a, b \in E$.

Corollary 4.2. *If the minimal solution $\bar{P}(t)$ is stochastic (the sums of matrix elements along the rows are all equal to 1) then it is the unique solution of (4.2) (or (4.4)) with the initial condition (4.3) in the class of substochastic matrices.*

If the minimal solution $\bar{P}(t)$ is stochastic one says that the matrix of transition rates Q is *regular*, cf. Proposition 4.3.

Observe that the construction of $\bar{P}(t)$ is very natural: the summands $P^{[n]}(t; a, b)$ are the probabilities to go from a to b in n jumps. The condition of $\bar{P}(t)$ being stochastic exactly means that we cannot make infinitely many jumps in a finite amount of time.

A much more detailed account of Markov chains on countable sets can be found e.g. in [1].

Later on we will need the following sufficient condition for $\bar{P}(t)$ to be stochastic.

For any finite subset $X \subset E$ and an element $a \in X$, denote by $T_{a,X}$ the time of the first exit from X under the condition that the process is in a at time 0. Formally, we can modify E and Q by contracting all the states $b \in E \setminus X$ into one absorbing state \tilde{b} with $Q_{\tilde{b},c} = 0$ for any $c \in X \cup \{\tilde{b}\}$. We obtain a process with a finite number of states for which the solution $\tilde{P}(t)$ of the backward equation is unique. Then $T_{a,X}$ is a random variable with values in $(0, +\infty]$ defined by

$$\text{Prob}\{T_{a,X} \leq t\} = \tilde{P}(t; a, \tilde{b}).$$

Proposition 4.3. *Assume that for any $a \in E$ and any $t > 0$, $\varepsilon > 0$, there exists a finite set $X(\varepsilon) \subset E$ such that $\text{Prob}\{T_{a,X(\varepsilon)} \leq t\} \leq \varepsilon$. Then the minimal solution $\bar{P}(t)$ provided by Theorem 4.1 is stochastic.*

Proof. Consider the modified process on the finite state space $X(\varepsilon) \cup \{\tilde{b}\}$ described above. Since its transition matrix $\tilde{P}(t)$ is stochastic,

$$\sum_{b \in X(\varepsilon)} \tilde{P}(t; a, b) = 1 - \tilde{P}(t; a, \tilde{b}) \geq 1 - \varepsilon.$$

The construction of the minimal solution as the sum of $P^{[n]}$'s, see above, immediately implies that $P(t; a, b) \geq \tilde{P}(t; a, b)$. Thus, $\sum_b P(t; a, b) \geq 1 - \varepsilon$ for any $\varepsilon > 0$. \square

4.2. Collapsibility

In what follows we will also need a result on *collapsibility* or *lumpability* of Markov chains on discrete spaces. Let us describe it.

Let $E = \bigsqcup_{i \in I} E_i$ be a partition of the countable set E on disjoint subsets. Assume we are given a matrix Q_E of transition rates on E and a matrix Q_I of transition rates on I such that

$$\sum_{b \in E_j} Q_E(a, b) = Q_I(i, j) \quad \text{for any } a \in E_i, \quad i, j \in I. \quad (4.5)$$

Denote

$$q_E(a) = -Q_E(a, a), \quad q_I(i) = -Q_I(i, i).$$

For any $i \in I$, let Q_i be a matrix of transition rates on E_i defined by

$$Q_i(a, b) = Q_E(a, b) \quad \text{if } a \neq b, \quad a, b \in E_i,$$

$$q_i(a) = -Q_i(a, a) = \sum_{b \in E_i, b \neq a} Q_i(a, b), \quad a \in E_i.$$

Observe that $q_E(a) = q_i(a) + q_I(i)$ for any $a \in E_i$.

Denote by $\bar{P}_E(t)$ and $\bar{P}_I(t)$ the minimal solutions of Kolmogorov's equations for Q_E and Q_I , respectively.

Proposition 4.4. Assume that for any $i \in I$, Q_i is regular. Then for any $t \geq 0$

$$\sum_{b \in E_j} \bar{P}_E(t; a, b) = \bar{P}_I(t; i, j) \quad (4.6)$$

for any $i, j \in I$ and any $a \in E_i$. In particular, if Q_I is regular then so is Q_E and vice versa.

Proof. Let us use notations q , \tilde{Q} , and $P^{[n]}$ for the diagonal and off-diagonal parts of the matrices of transition rates, and for the n th terms in the series representations of minimal solutions, respectively.

The hypothesis means that the identity

$$\sum_{b \in E_i} \sum_{n=0}^{\infty} P_i^{[n]}(t; a, b) = 1, \quad a \in E_i, \quad i \in I,$$

holds, where

$$P_i^{[n]}(t) = \int_{0 \leq t_1 \leq \dots \leq t_n \leq t} e^{-t_1 q_i} \tilde{Q}_i e^{-(t_2 - t_1) q_i} \tilde{Q}_i \dots e^{-(t_n - t_{n-1}) q_i} \tilde{Q}_i e^{-(t - t_n) q_i} dt_1 \dots dt_n$$

and $P_i^{(0)}(t) = e^{-t q_i}$. Using the fact that $q_i(\cdot) = q_E(\cdot) - q_I(i)$ on E_i , rewrite this identity as

$$\sum_{b \in E_i} \sum_{n=0}^{\infty} \int_{0 \leq t_1 \leq \dots \leq t_n \leq t} (e^{-t_1 q_E} \mathbf{1}_{E_i} \tilde{Q}_E \mathbf{1}_{E_i} e^{-(t_2 - t_1) q_E} \mathbf{1}_{E_i} \dots$$

$$\dots \mathbf{1}_{E_i} \tilde{Q}_E \mathbf{1}_{E_i} e^{-(t - t_n) q_E})(a, b) dt_1 \dots dt_n = e^{-t q_I(i)}. \quad (4.7)$$

The probabilistic meaning of this formula is that the time that the minimal solution $\bar{P}_E(t)$ started at $a \in E_i$ spends in E_i is exponentially distributed with rate $q_I(i)$, independent of a .

The minimal solution $\bar{P}_I(t)$ has the form

$$\bar{P}_I(t; i, j) = \sum_{n=0}^{\infty} \int_{0 \leq s_1 \leq \dots \leq s_n \leq t} \sum_{k_1, \dots, k_{n-1} \in I} e^{-s_1 q_I(i)} \tilde{Q}_I(i, k_1) e^{-(s_2 - s_1) q_I(k_1)} \dots \tilde{Q}_I(k_{n-1}, j)$$

$$\times e^{-(t - s_n) q_I(j)} ds_1 \dots ds_n. \quad (4.8)$$

By (4.5),

$$\tilde{Q}_I(k, l) = \sum_{d \in E_l} \tilde{Q}_E(c, d) \quad \text{for any } k, l \in I, k \neq l, c \in E_k. \quad (4.9)$$

Substituting the right-hand side of (4.9) for each $\tilde{Q}_I(\cdot, \cdot)$ and the left-hand side of (4.7) for each $e^{-sq_I(\cdot)}$, in the n th term we obtain the part of the series for $\bar{P}_E(t; a, b)$ with $a \in E_i$ that takes into account trajectories whose projections to I make exactly n jumps, and in addition to that there is a summation over $b \in E_j$. Clearly, the summation over n reproduces the complete series for $\bar{P}_E(t; a, b)$ thus proving (4.6).

The equivalence of stochasticity of $\bar{P}_E(t)$ and that of $\bar{P}_I(t)$ immediately follows from summation of (4.6) over $j \in I$. \square

4.3. Infinitesimal generator

The last part of the general theory that we need involves generators of Markov semigroups.

Assume that we have a regular matrix of transition rates Q . Let $(P(t))_{t \geq 0}$ be the corresponding Markov semigroup and assume in addition that it is Feller.

The generator A of the semigroup $(P(t))_{t \geq 0}$ is a linear operator in $C_0(E)$ defined by

$$Af = \lim_{t \rightarrow +0} \frac{P(t)f - f}{t}. \quad (4.10)$$

The set of functions $f \in C_0(E)$ for which this limit exists (in the norm topology of $C_0(E)$) is called the *domain* of the generator A and denoted by $D(A)$. It is well known that the operator A with $D(A)$ as above is closed and dissipative.

It turns out that the domain $D(A)$ can be characterized by an apparently weaker condition, which is easier to verify in practice:

Proposition 4.5. *If $f \in C_0(E)$ is such that the limit in the right-hand side of (4.10) exists pointwise and the limit function belongs to $C_0(E)$, then $f \in D(A)$, so that the limit actually holds in the norm topology.*

Proof. The idea is that the set of couples of vectors $(f, g) \in C_0(E) \times C_0(E)$, such that g is the pointwise limit of the right-hand side of (4.10), serves as the graph of a dissipative operator \tilde{A} extending A , whence $\tilde{A} = A$. A detailed argument can be found in [24, §4.8]. In fact, [24] considers the case of a compact state space E . However, the proof goes through word-for-word; the only property one needs is that for any $f \in C_0(E)$, f attains its minimum if it has negative values. \square

The following statement is probably well known but we were not able to locate it in the literature.

Proposition 4.6. *Assume that for any $a \in E$ the set of elements b such that $Q(a, b) \neq 0$ is finite. Then*

$$D(A) = \{f \in C_0(E) \mid Qf \in C_0(E)\}, \quad (4.11)$$

and for any $f \in D(A)$, $Af = Qf$.

Proof. First of all, due to the assumption on the matrix Q , Qf is well defined for any function f on E . We will show that for any $f \in C_0(E)$ and $a \in E$

$$\lim_{t \rightarrow +0} t^{-1} \sum_{b \in E} (P(t; a, b) - \mathbf{1}_{a=b}) f(b) = \sum_{b \in E} Q(a, b) f(b). \quad (4.12)$$

Then the claim of the proposition will follow from Proposition 4.5.

Set

$$X = \{a\} \cup \{b' \in E \mid Q(a, b') > 0\}.$$

By our hypothesis, this set is finite. We will show that

$$\sum_{b \in E \setminus X} P(t; a, b) = O(t^2), \quad t \rightarrow 0. \quad (4.13)$$

This would imply that we can keep only finitely many terms in (4.12), and then (4.12) would follow from (4.1).

Observe that the left-hand side of (4.13) is the probability of the event that the trajectory started at a is outside of X after time t . In order to exit X the trajectory started at a needs to make at least two jumps. Assume that the first two jumps are $a \rightarrow a' \rightarrow a''$ with $a' \in X$. Since X is finite, the rates of leaving a' (equal to $-Q(a', a')$) are bounded from above, and the probability of leaving X after time t can be estimated by

$$-Q(a, a) \max_{a' \in X} (-Q(a', a')) \cdot t^2 + o(t^2) = O(t^2), \quad t \rightarrow +0,$$

as required. \square

Corollary 4.7. *Under the hypothesis of Proposition 4.6 assume additionally that for any $b \in E$ the set of elements $a \in E$ with $Q(a, b) \neq 0$ is finite. Then any finitely supported function f on E belongs to $D(A)$.*

Proof. Indeed, this follows immediately from Proposition 4.6, since Qf is finitely supported and hence belongs to $C_0(E)$. \square

5. Semigroups on \mathbb{GT}_N

The goal of this section is to define Markov semigroups $(P_N(t))_{t \geq 0}$ on $E_N = \mathbb{GT}_N$ and prove that they are Feller.

5.1. Case $N = 1$. Birth and death process on \mathbb{Z}

Let (u, u') and (v, v') be two pairs of complex numbers such that $(u + k)(u' + k) > 0$ and $(v + k)(v' + k) > 0$ for any $k \in \mathbb{Z}$. The condition on (u, u') means that either $u' = \bar{u} \in \mathbb{C} \setminus \mathbb{R}$ or there exists $k \in \mathbb{Z}$ such that $k < u, u' < k + 1$; the condition on (v, v') is similar. Note that $u + u' \in \mathbb{R}$ and $v + v' \in \mathbb{R}$. Assume additionally that $u + u' + v + v' > -1$.

Define a matrix of transition rates $[\mathcal{D}(x, y)]_{x, y \in \mathbb{Z}}$ with rows and columns parameterized by elements of $E_1 = \mathbb{GT}_1 = \mathbb{Z}$ by

$$\mathcal{D}(x, y) = \begin{cases} (x - u)(x - u'), & \text{if } y = x + 1, \\ (x + v)(x + v'), & \text{if } y = x - 1, \\ -(x - u)(x - u') - (x + v)(x + v'), & \text{if } y = x, \\ 0, & \text{otherwise.} \end{cases} \quad (5.1)$$

In the corresponding Markov chain the particle would only be allowed to jump by one unit at a time; such processes on $\mathbb{Z}_{\geq 0}$ are usually referred to as *birth and death processes*, while our Markov chain is an example of so-called *bilateral birth and death processes* which were also considered in the literature, see e.g. [21, Section 17], [48, 56].

Note that $\mathcal{D}(x, x \pm 1) > 0$ for all $x \in \mathbb{Z}$, because of the conditions imposed on the parameters.

Theorem 5.1. *The matrix of transition rates \mathcal{D} is regular. Moreover, the corresponding Markov semigroup is Feller.*

In what follows we denote this semigroup by $(P_1(t))_{t \geq 0}$.

The proof of Theorem 5.1 is based on certain results from [22]; let us recall them first.

Consider a birth and death process on $\mathbb{Z}_{\geq 0}$ with transition rates given by

$$Q(x, y) = \begin{cases} \beta_x, & \text{if } y = x + 1, \\ \delta_x, & \text{if } y = x - 1, \\ -\beta_x - \delta_x, & \text{if } y = x, \\ 0, & \text{otherwise.} \end{cases}$$

Here $\{\beta_x\}_{x \geq 0}$, $\{\delta_x\}_{x \geq 1}$ are positive numbers, and we also set $\delta_0 = 0$.

The *natural scale* of the process is given by

$$x_0 = 0, \quad x_k = \sum_{l=0}^{k-1} \frac{\delta_1 \cdots \delta_l}{\beta_0 \cdots \beta_l}, \quad k = 1, 2, \dots, \quad x_\infty = \lim_{k \rightarrow \infty} x_k. \quad (5.2)$$

Note that x_∞ may be infinite. Denote by \mathcal{A} the operator on the space of functions on $\mathbb{A} = \{x_0, x_1, \dots\}$ defined by

$$(\mathcal{A}f)(x_i) = -(\delta_i + \beta_i)f(x_i) + \delta_i f(x_{i-1}) + \beta_i f(x_{i+1}), \quad i = 0, 1, \dots$$

Fix $n > 0$. Let $F_i(t)$ be the probability that the process started at i reaches n before time t . Let $G_n(t)$ be the probability that the process started at n reaches 0 before time t and before the process escapes to infinity.

Theorem 5.2. (See [22].) (i) *For any $a > 0$ there exists exactly one function u on \mathbb{A} such that $Au = au$, $u(x_0) = 1$. The function u is strictly increasing: $u(x_0) < u(x_1) < u(x_2) < \dots$ and satisfies*

$$u(x_n) = 1 + a \sum_{k=0}^{n-1} u(x_k)(x_n - x_k)\mu_k \quad (5.3)$$

with

$$\mu_k = \frac{\beta_0 \cdots \beta_{k-1}}{\delta_1 \cdots \delta_k}, \quad k = 1, 2, \dots, \quad \mu_0 = 1. \quad (5.4)$$

Furthermore,

$$\frac{u(x_i)}{u(x_n)} = \int_0^\infty e^{-at} dF_i(t), \quad 0 \leq i < n. \quad (5.5)$$

(ii) With $u(\cdot)$ as above, set

$$v(x_n) = u(x_n) \sum_{j=n}^\infty \frac{x_{j+1} - x_j}{u(x_j)u(x_{j+1})}, \quad n = 0, 1, \dots$$

This is a strictly decreasing function, and

$$\frac{v(x_n)}{v(x_0)} = \int_0^\infty e^{-at} dG_n(t), \quad n = 1, 2, \dots \quad (5.6)$$

Furthermore, $\lim_{n \rightarrow \infty} v(x_n) = 0$ if $x_\infty = \infty$ and $\sum_n x_n \mu_n$ diverges.

The following statement is contained in Feller's paper as well, but not explicitly; for that reason we formulate it separately.

Corollary 5.3. *If $x_\infty = \infty$ then $\lim_{n \rightarrow \infty} u(x_n) = \infty$.*

Proof. Let us estimate the sum in the right-hand side of (5.3):

$$\begin{aligned} \sum_{k=0}^{n-1} u(x_k)(x_n - x_k)\mu_k &\geq \sum_{k=0}^{n-1} (x_n - x_k)\mu_k = \sum_{k=0}^{n-1} \sum_{l=k+1}^n (x_l - x_{l-1})\mu_k \\ &= \sum_{l=1}^n \sum_{k=0}^{l-1} (x_l - x_{l-1})\mu_k = \sum_{l=0}^{n-1} (x_{l+1} - x_l) \sum_{k=0}^l \mu_k \\ &\geq \sum_{l=0}^{n-1} (x_{l+1} - x_l) = x_n. \end{aligned} \quad (5.7)$$

Hence, $u(x_n) \geq 1 + ax_n$, and the statement follows. \square

Let us now apply Feller's results to our situation.

Proof of Theorem 5.1. By Proposition 4.3, in order to show that the minimal solution is stochastic it suffices to prove that the probability that the first passage time from 0 to n is below a fixed number, converges to zero as $n \rightarrow +\infty$. Indeed, as shifts $x \rightarrow x + \text{const}$ and sign change $x \mapsto -x$ keep our class of processes intact, similar convergence would automatically hold for passage times to the left, and also for passage times from any initial position. Denote the first passage time from 0 to n by T_n .

A simple coupling argument shows that T_n stochastically dominates the first passage time from 0 to n for the birth and death process on $\mathbb{Z}_{\geq 0}$ with the same transition rates (see (5.1)), except that the jump from 0 to -1 is forbidden. Let us denote this new first passage time by \tilde{T}_n . Thus,

$$\text{Prob}\{T_n \leq t\} \leq \text{Prob}\{\tilde{T}_n \leq t\} \quad \text{for any } n \geq 0 \text{ and } t > 0.$$

For the application of Theorem 5.1 we then set

$$\beta_x = (x - u)(x - u'), \quad x \geq 0, \quad \delta_x = (x + v)(x + v'), \quad x \geq 1, \quad \delta_0 = 0.$$

As

$$\frac{\delta_1 \cdots \delta_l}{\beta_0 \cdots \beta_l} = \text{const} \frac{\Gamma(v + l + 1)\Gamma(v' + l + 1)}{\Gamma(-u + l + 1)\Gamma(-u' + l + 1)} \sim \text{const} \cdot l^{u+u'+v+v'}, \quad l \rightarrow \infty,$$

our original assumption $u + u' + v + v' > -1$ implies

$$x_k \sim \text{const} \cdot k^{u+u'+v+v'+1}, \quad k \rightarrow \infty, \quad (5.8)$$

cf. (5.2), and $x_\infty = \lim_{k \rightarrow \infty} x_k = \infty$. Therefore, Corollary 5.3 yields

$$\lim_{n \rightarrow \infty} u(x_n) = \infty.$$

On the other hand, from (5.5) with $i = 0$ and any $a > 0$ we obtain

$$\frac{1}{u(x_n)} = \int_0^\infty e^{-a\tau} dF_0(\tau) \geq \int_0^t e^{-a\tau} dF_0(\tau) \geq e^{-at} \int_0^t dF_0(\tau) = e^{-at} \text{Prob}\{\tilde{T}_n \leq t\}$$

whence

$$\text{Prob}\{\tilde{T}_n \leq t\} \leq \frac{e^{at}}{u(x_n)} \rightarrow 0, \quad n \rightarrow +\infty.$$

Since \tilde{T}_n is dominated by T_n , we have shown that our Markov chain does not run away to infinity in finite time, and hence it is uniquely specified by the transition rates. Let $(P_1(t))_{t \geq 0}$ be the corresponding semigroup.

We now need to prove that $(P_1(t))_{t \geq 0}$ is Feller. This is equivalent to showing that $\lim_{n \rightarrow \pm\infty} P_1(t; n, i) = 0$ for any $i \in \mathbb{Z}$ and $t > 0$.

Shift and sign change invariance (see the beginning of the proof) imply that it suffices to consider $i = 0$ and $n \rightarrow +\infty$. Observe that $P_1(t; n, 0)$ cannot be greater than the probability that the first passage time from n to 0 is not more than t . Let us denote this first passage time by S_n ; we have $P_1(t; n, 0) \leq \text{Prob}\{S_n \leq t\}$.

This first passage time is the same for our birth and death process on \mathbb{Z} and for its modification on $\mathbb{Z}_{\geq 0}$ that was used in the first part of the proof. On the other hand, for the process on $\mathbb{Z}_{\geq 0}$ the Laplace transform of S_n is given by (5.6).

By (5.4) we have, as $k \rightarrow \infty$,

$$\mu_k = \frac{\beta_0 \cdots \beta_{k-1}}{\delta_1 \cdots \delta_k} = \text{const} \frac{\Gamma(-u+k)\Gamma(-u'+k)}{\Gamma(v+k+1)\Gamma(v'+k+1)} \sim \text{const} \cdot k^{-2-u-u'-v-v'}.$$

Hence, cf. (5.8)

$$x_k \mu_k \sim \text{const} \cdot k^{-1}, \quad k \rightarrow \infty,$$

with a nonzero constant, and $\sum_n x_n \mu_n$ diverges. Theorem 5.2(ii) then gives

$$\lim_{n \rightarrow \infty} v(x_n) = 0$$

and using (5.6) and estimating the Laplace transform as above we obtain

$$\text{Prob}\{S_n \leq t\} \leq \frac{e^{at} v(x_n)}{v(x_0)} \rightarrow 0, \quad n \rightarrow \infty.$$

As $P_1(t; n, 0) \leq \text{Prob}\{S_n \leq t\}$, the proof of Theorem 5.1 is complete. \square

5.2. The case of general N

Let $N > 1$ be a positive integer, and let (u, u') and (v, v') be as in Section 5.1.

Define a matrix $[\mathcal{D}^{(N)}(\lambda, \nu)]_{\lambda, \nu \in \mathbb{GT}_N}$ of transition rates with rows and columns parameterized by points of $E_N = \mathbb{GT}_N$ via

$$\begin{aligned} \mathcal{D}^{(N)}(\lambda, \nu) = & \frac{\text{Dim}_N(\nu)}{\text{Dim}_N(\lambda)} (\mathcal{D}(l_1, n_1) \mathbf{1}_{\{l_i=n_i, i \neq 1\}} + \mathcal{D}(l_2, n_2) \mathbf{1}_{\{l_i=n_i, i \neq 2\}} + \cdots \\ & + \mathcal{D}(l_N, n_N) \mathbf{1}_{\{l_i=n_i, i \neq N\}}) - d_N \cdot \mathbf{1}_{\lambda=\nu} \end{aligned} \quad (5.9)$$

with $l_j = \lambda_j + N - j$, $n_j = \nu_j + N - j$, $1 \leq j \leq N$, matrix $\mathcal{D}(\cdot, \cdot)$ as in (5.1), and

$$d_N = \frac{N(N-1)(N-2)}{3} - (u + u' + v + v') \frac{N(N-1)}{2}. \quad (5.10)$$

In other words, an off-diagonal element $\mathcal{D}^{(N)}(\lambda, \nu)$ can only be nonzero if there exists exactly one index i such that $\nu_i - \lambda_i = \pm 1$ while for all other indices j we have $\lambda_j = \nu_j$. Under this condition

$$\mathcal{D}^{(N)}(\lambda, \nu) = \begin{cases} (l_i - u)(l_i - u') \prod_{j \neq i} \frac{l_i + 1 - l_j}{l_i - l_j}, & \text{if } \nu_i - \lambda_i = 1, \\ (l_i + v)(l_i + v') \prod_{j \neq i} \frac{l_i - 1 - l_j}{l_i - l_j}, & \text{if } \nu_i - \lambda_i = -1. \end{cases}$$

With this explicit description, the diagonal entries of $\mathcal{D}^{(N)}$ have to be defined by

$$\mathcal{D}^{(N)}(\lambda, \lambda) = - \sum_{\nu \in \mathbb{GT}_N: \nu \neq \lambda} \mathcal{D}^{(N)}(\lambda, \nu), \quad \lambda \in \mathbb{GT}_N. \quad (5.11)$$

The fact that (5.11) holds for $\mathcal{D}^{(N)}$ defined by (5.9) will be proved in Step 1 of the proof of the following theorem.

Theorem 5.4. *The matrix of transition rates $\mathcal{D}^{(N)}$ is regular. The corresponding semigroup $(P_N(t))_{t \geq 0}$ has the form*

$$P_N(t; \lambda, \nu) = e^{-d_N t} \frac{\text{Dim}_N(\nu)}{\text{Dim}_N(\lambda)} \det[P_1(t; \lambda_i + N - i, \nu_j + N - j)]_{i,j=1}^N, \\ \lambda, \nu \in \mathbb{GT}_N, \quad (5.12)$$

with $(P_1(t))_{t \geq 0}$ as in Section 5.1. Moreover, this semigroup is Feller.

Proof. The proof of Theorem 5.4 will consist of several steps.

Step 1. Let us show that with definition (5.9), relation (5.11) holds. It is convenient to encode signatures of length N by N -tuples of strictly decreasing integers via

$$\lambda = (\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N) \longleftrightarrow (l_1 > l_2 > \dots > l_N), \quad l_j = \lambda_j + N - j, \quad 1 \leq j \leq N.$$

This establishes a bijection between \mathbb{GT}_N and the set

$$\mathfrak{X}_N = \{(x_1, \dots, x_N) \in \mathbb{Z}^N \mid x_1 > x_2 > \dots > x_N\}.$$

In \mathfrak{X}_N , matrix $\mathcal{D}^{(N)}$ from (5.9) takes the form

$$\mathcal{D}^{(N)}(X, Y) = \frac{V_N(Y)}{V_N(X)} (\mathcal{D}(x_1, y_1) \mathbf{1}_{\{x_i=y_i, i \neq 1\}} + \mathcal{D}(x_2, y_2) \mathbf{1}_{\{x_i=y_i, i \neq 2\}} + \dots \\ + \mathcal{D}(x_N, y_N) \mathbf{1}_{\{x_i=y_i, i \neq N\}}) - d_N \mathbf{1}_{X=Y} \quad (5.13)$$

with $X = (x_1, \dots, x_N) \in \mathfrak{X}_N$, $Y = (y_1, \dots, y_N) \in \mathfrak{X}_N$, and

$$V_N(z_1, \dots, z_N) = \prod_{1 \leq i < j \leq N} (z_i - z_j).$$

In this notation, (5.11) is equivalent to

$$(\mathcal{D}_1 + \dots + \mathcal{D}_N) V_N(X) = d_N V_N(X), \quad X \in \mathbb{Z}^N, \quad (5.14)$$

where \mathcal{D}_j denotes a linear operator on \mathbb{Z}^N with

$$\mathcal{D}_j(X, Y) = \mathcal{D}(x_j, y_j) \mathbf{1}_{\{x_i=y_i, i \neq j\}}.$$

Indeed, both sides of (5.14) are skew-symmetric, and restricting to \mathfrak{X}_N yields (5.11).

Let Δ and ∇ be the standard forward and backward difference operators on \mathbb{Z} :

$$\Delta f(x) = f(x+1) - f(x), \quad \nabla f(x) = f(x) - f(x-1)$$

for any function $f: \mathbb{Z} \rightarrow \mathbb{C}$. Note that $\Delta \nabla = \Delta - \nabla$.

One easily checks that the operator \mathcal{D} with matrix (5.1) has the form

$$\mathcal{D} = \sigma \Delta \nabla + \tau \Delta \tag{5.15}$$

with

$$\sigma = (x+v)(x+v'), \quad \tau = sx + (uu' - vv'), \quad s = -(u+u' + v+v').$$

Hence, for any $m = 0, 1, 2, \dots$

$$\mathcal{D}x^m = (m(m-1) + sm) \cdot x^m + \text{lower degree terms}, \tag{5.16}$$

in particular, \mathcal{D} preserves the degree of a polynomial. This implies that the left-hand side of (5.14) is a skew-symmetric polynomial of degree at most $N(N-1)/2$. It must be divisible by the Vandermonde determinant $V_N(X)$, and it remains to verify the constant prefactor. Following the highest in lexicographic order term $x_1^{N-1}x_2^{N-2}\dots x_N^0$ we see that upon the action of $(\mathcal{D}_1 + \dots + \mathcal{D}_N)$ it collects the coefficient

$$\sum_{j=0}^{N-1} (j(j-1) + sj),$$

which sums to (5.10). Thus, (5.11) is proved.

Step 2. Let us now prove that

$$\sum_{v \in \mathbb{GT}_N} P_N(t; \lambda, v) = 1, \quad \lambda \in \mathbb{GT}_N, \quad t \geq 0, \tag{5.17}$$

with P_N as in (5.12). In the space \mathfrak{X}_N , (5.12) reads

$$P_N(t; X, Y) = e^{-d_N t} \frac{V_N(Y)}{V_N(X)} \det[P_1(t; x_i, y_j)]_{i,j=1}^N, \quad X, Y \in \mathfrak{X}_N. \tag{5.18}$$

Since the action of \mathcal{D} in the space of polynomials $\mathbb{R}[x]$ is consistent with filtration by degree, see (5.16), the action of the corresponding semigroup $(P_1(t))_{t \geq 0}$ in $\mathbb{R}[x]$ is well-defined, and (5.16) implies

$$\sum_{y \in \mathbb{Z}} P_1(t; x, y) y^m = e^{(m(m-1)+sm)t} x^m + \text{lower degree terms}.$$

We obtain

$$\begin{aligned}
 & \sum_{Y \in \mathfrak{X}_N} P_N(t; X, Y) \\
 &= \frac{1}{N!} \sum_{Y \in \mathbb{Z}^N} P_N(t; X, Y) \\
 &= \frac{e^{-d_N t}}{V(X)} \sum_{\sigma \in S_N} \operatorname{sgn} \sigma \sum_{Y \in \mathbb{Z}^N} P_1(t; x_{\sigma(1)}, y_1) \cdots P_1(t; x_{\sigma(N)}, y_N) y_1^{N-1} y_2^{N-2} \cdots y_{N-1} \\
 &= \frac{e^{-d_N t}}{V(X)} \sum_{\sigma \in S_N} \operatorname{sgn} \sigma e^{t \sum_{j=0}^{N-1} (j(j-1) + s_j)} x_{\sigma(1)}^{N-1} x_{\sigma(2)}^{N-2} \cdots x_{\sigma(N-1)} = 1,
 \end{aligned} \tag{5.19}$$

where S_N denotes the group of permutations of $\{1, \dots, N\}$. Note that the first equality (change of the summation domain) holds because the expression for $P_N(t; X, Y)$ is symmetric in (y_j) , and it vanishes if $y_i = y_j$ for some $i \neq j$.

Step 3. Consider N independent copies of the bilateral birth and death process of Section 5.1, and denote by $\pi_n(t; X, Y)$, $X, Y \in \mathfrak{X}_N$, the probability that these processes started at x_1, \dots, x_N end up at y_1, \dots, y_N after time t having made a total of n jumps all together, and their trajectories had no common points at any time moment between 0 and t . We want to show that

$$P_N^{[n]}(t; X, Y) = e^{-d_N t} \frac{V(Y)}{V(X)} \pi_n(t; X, Y), \tag{5.20}$$

where $P_N^{[n]}$ is defined as in Section 4 using $\mathcal{D}^{(N)}$ as the matrix of transition rates.

Indeed, computing π_n 's boils down to recurrence relations

$$\begin{aligned}
 \pi_0(t; X, Y) &= e^{t \mathcal{D}_{ind}^{(N)}(X, Y)} \mathbf{1}_{X=Y}, \\
 \pi_n(t; X, Y) &= \int_0^t e^{\tau \mathcal{D}_{ind}^{(N)}(X, X)} \sum_{Z \in \mathfrak{X}_N, Z \neq X} \mathcal{D}_{ind}^{(N)}(X, Z) \pi_{n-1}(t - \tau; Z, Y) d\tau, \quad n \geq 1,
 \end{aligned}$$

where $\mathcal{D}_{ind}^{(N)} = \mathcal{D}_1 + \cdots + \mathcal{D}_N$ is the matrix of transition rates for the N independent birth and death processes.

For $n = 0$, (5.20) follows from (5.9). Assuming (5.20) holds for $n - 1$, we rewrite the recurrence relation for π_n 's as

$$\begin{aligned}
 \pi_n(t; X, Y) &= \int_0^t e^{\tau (\mathcal{D}^{(N)}(X, X) + d_N)} \sum_{Z \in \mathfrak{X}_N, Z \neq X} \frac{V(X)}{V(Z)} \mathcal{D}^{(N)}(X, Z) \\
 &\quad \cdot e^{d_N(t-\tau)} \frac{V(Z)}{V(Y)} P_N^{[n-1]}(t - \tau; Z, Y) d\tau.
 \end{aligned} \tag{5.21}$$

Comparing with the recurrence relation for $P^{[n]}$, cf. Section 4, yields (5.20).

Step 4. Following Section 4 and using (5.20), we see that the minimal solution for the backward equation with $\mathcal{D}^{(N)}$ as the matrix of transition rates, has the form

$$\bar{P}_N(t; X, Y) = \sum_{n=0}^{\infty} P_N^{[n]}(t; X, Y) = e^{-d_N t} \frac{V(Y)}{V(X)} \sum_{n=0}^{\infty} \pi_n(t; X, Y).$$

The last sum is clearly equal to the probability that N independent copies of the bilateral birth and death process of Section 5.1 started at x_1, \dots, x_N end up at y_1, \dots, y_N after time t without intermediate coincidences and without any restriction on the number of jumps. Note that we are using the fact that the birth and death process does not make infinitely many jumps in finite time (minimal solution is stochastic), cf. Theorem 5.1.

Such a probability of having nonintersecting paths is given by a celebrated formula of Karlin and McGregor [26]:

$$\sum_{n=0}^{\infty} \pi_n(t; X, Y) = \det[P_1(t; x_i, y_j)]_{i,j=1}^N, \quad X, Y, \in \mathfrak{X}_N.$$

Hence, the minimal solution $\bar{P}_N(t; X, Y)$ coincides with the right-hand side of (5.18), and by Step 2 it is stochastic. We have thus shown that the matrix $\mathcal{D}^{(N)}$ of transition rates on \mathbb{GT}_N is regular, and the semigroup has the form (5.12) (or (5.18)).

Step 5. To conclude the proof of Theorem 5.4 it remains to show that the Markov semigroup $(P_N(t))_{t \geq 0}$ is Feller. This is equivalent to proving that

$$\lim_{\lambda \rightarrow \infty} P_N(t; \lambda, v) = 0, \quad t \geq 0, \quad v \in \mathbb{GT}_N. \quad (5.22)$$

But this immediately follows from (5.12) because we already know that (5.22) holds for $N = 1$ (Theorem 5.1), and $\text{Dim}_N(\lambda)$ is always at least 1. \square

6. Commutativity

The goal of this section is to address the question of compatibility of the semigroups of Section 5 and links of Section 3, cf. (2.3).

6.1. Parameterization

As we shall see, in order for the commutativity relations (2.3) to be satisfied, the parameters (u, u', v, v') used to define semigroups $(P_N(t))_{t \geq 0}$ need to depend on N . For that reason, introduce two new pairs of parameters (z, z') and (w, w') that satisfy the same conditions as (u, u', v, v') before:

$$(z + k)(z' + k) > 0, \quad (w + k)(w' + k) > 0 \quad \forall k \in \mathbb{Z}; \quad z + z' + w + w' > -1. \quad (6.1)$$

Furthermore, for $N \geq 1$ define

$$u_N = z + N - 1, \quad u'_N = z' + N - 1, \quad v_N = w, \quad v'_N = w', \quad (6.2)$$

and let $(P_N(t))_{t \geq 0}$ be the Feller semigroup of the previous section with parameters $(u, u', v, v') = (u_N, u'_N, v_N, v'_N)$.

We are aiming to prove the following statement.

Theorem 6.1. *With links $\{\Lambda_N^{N+1}\}_{N \geq 1}$ as in Section 3.1 and semigroups $(P_N(t))_{t \geq 0}$ as above, the compatibility relations (2.3) hold.*

6.2. Infinitesimal commutativity

We first prove a version of (2.3) that involves matrices of transition rates.

Proposition 6.2. *For any $N \geq 1$, $u, u', v, v' \in \mathbb{C}$, and $\lambda \in \mathbb{GT}_{N+1}$, $\nu \in \mathbb{GT}_N$, we have*

$$\sum_{\kappa \in \mathbb{GT}_{N+1}} \tilde{\mathcal{D}}^{(N+1)}(\lambda, \kappa) \Lambda_N^{N+1}(\kappa, \nu) = \sum_{\rho \in \mathbb{GT}_N} \Lambda_N^{N+1}(\lambda, \rho) \mathcal{D}^{(N)}(\rho, \nu) \quad (6.3)$$

or, in matrix notation, $\tilde{\mathcal{D}}^{(N+1)} \Lambda_N^{N+1} = \Lambda_N^{N+1} \mathcal{D}^{(N)}$, where $\mathcal{D}^{(N)}$ is the operator defined by (5.9), and in $\tilde{\mathcal{D}}^{(N+1)}$ we replace N by $N+1$ and the parameters (u, u') by $(\tilde{u}, \tilde{u}') = (u+1, u'+1)$.

Proof. We start with the following simple lemma.

Lemma 6.3. *Let $[A(\lambda, \nu)]_{\lambda \in \mathbb{GT}_{N+1}, \nu \in \mathbb{GT}_N}$ be a matrix with rows parameterized by \mathbb{GT}_{N+1} and columns parameterized by \mathbb{GT}_N , and such that each row of A has finitely many nonzero entries. If for any symmetric polynomial F in N variables and any $\lambda \in \mathbb{GT}_{N+1}$ we have*

$$\sum_{\nu \in \mathbb{GT}_N} A(\lambda, \nu) F(\nu_1 + N - 1, \nu_2 + N - 2, \dots, \nu_N) = 0, \quad (6.4)$$

then $A(\lambda, \nu) \equiv 0$.

Proof. Assume $A(\hat{\lambda}, \hat{\nu}) \neq 0$ for some $\hat{\lambda}$ and $\hat{\nu}$. Let $\nu^{(1)}, \dots, \nu^{(l)} \in \mathbb{GT}_N$ be all signatures different from $\hat{\nu}$ and such that $A(\hat{\lambda}, \nu^{(j)}) \neq 0$.

Set $x = (\hat{\nu}_1 + N - 1, \dots, \hat{\nu}_N) \in \mathbb{Z}^N$ and

$$y^{(j)} = (\nu_1^{(j)} + N - 1, \dots, \nu_N^{(j)}) \in \mathbb{Z}^N, \quad j = 1, \dots, l.$$

Observe that the orbits of the vectors $x, y^{(1)}, \dots, y^{(l)}$ under the group of permutations of the coordinates do not intersect. It follows that there exists a polynomial f in N variables, which takes value 1 on the orbit of x and vanishes on the orbits of the vectors $y^{(1)}, \dots, y^{(l)}$. Then for the symmetrized polynomial $F(z_1, \dots, z_N) = \sum_{\sigma \in S_N} f(z_{\sigma(1)}, \dots, z_{\sigma(N)})$ the left-hand side of (6.4) is equal to $N!A(\hat{\lambda}, \hat{\nu}) \neq 0$. Contradiction. \square

Let us now introduce symmetric polynomials on which we will evaluate (in the sense of Lemma 6.3) both sides of (6.3). For a partition (= signature with nonnegative coordinates) $\mu \in \mathbb{GT}_N$ and $c \in \mathbb{C}$ set

$$F_{\mu,c}(x_1, \dots, x_n) = \frac{1}{(N)_\mu} \frac{\det[(x_i + c)^{\downarrow(\mu_j + N - j)}]_{i,j=1}^N}{\prod_{1 \leq i < j \leq N} (x_i - x_j)},$$

$$G_{\mu,c}(x_1, \dots, x_{n+1}) = \frac{1}{(N+1)_\mu} \frac{\det[(x_i + c)^{\downarrow(\mu_j + N + 1 - j)}]_{i,j=1}^{N+1}}{\prod_{1 \leq i < j \leq N+1} (x_i - x_j)},$$

where we assume $\mu_{N+1} = 0$ and use the notation ($a \in \mathbb{C}$, $k \in \mathbb{Z}_{\geq 0}$)

$$a^{\downarrow k} = a(a-1) \cdots (a-k+1), \quad (a)_k = a(a+1) \cdots (a+k-1), \quad a^{\downarrow 0} = (a)_0 = 1,$$

$$(a)_\mu = \prod_{j=1}^N (a-j+1)_{\mu_j}.$$

Clearly, $F_{\mu,c}$ and $G_{\mu,c}$ are symmetric polynomials in N and $N+1$ variables, respectively. Moreover, for any fixed $c \in \mathbb{C}$, the polynomials $\{F_{\mu,c}\}$ with μ ranging over all nonnegative signatures in \mathbb{GT}_N form a linear basis in the space of all symmetric polynomials in N variables. Indeed, this follows from the fact that the highest degree homogeneous component of $F_{\mu,c}$ coincides with the Schur polynomial $s_\mu(x_1, \dots, x_N)$, and those are well known to form a basis, see e.g. [38].

Hence, to prove Proposition 6.2 it suffices to verify that the two sides of (6.3) give the same results when applied to $F_{\mu,c}$ for a fixed c and μ varying over nonnegative signatures of length N .

Lemma 6.4. *For any $\lambda \in \mathbb{GT}_{N+1}$, any nonnegative signature $\mu \in \mathbb{GT}_N$, and $c \in \mathbb{C}$, we have*

$$\sum_{\nu \in \mathbb{GT}_N} \Lambda_N^{N+1}(\lambda, \nu) F_{\mu,c}(\nu_1 + N - 1, \dots, \nu_N) = G_{\mu,c}(\lambda_1 + N, \lambda_2 + N - 1, \dots, \lambda_{N+1}).$$

Proof. The argument is similar to that for relation (10.30) in [42]. Denote

$$(x_1, \dots, x_{N+1}) = (\lambda_1 + N, \dots, \lambda_{N+1}), \quad (y_1, \dots, y_N) = (\nu_1 + N - 1, \dots, \nu_N).$$

Then $\nu \prec \lambda$ means $x_{i+1} \leq y_i < x_i$ for all $i = 1, \dots, N$. Taking into account the definition of Λ_N^{N+1} , one sees that the relation in question is equivalent to the following one

$$\det[(x_i + c)^{\downarrow(\mu_j + N + 1 - j)}]_{i,j=1}^{N+1}$$

$$= \frac{(N+1)_\mu N!}{(N)_\mu} \sum_{\substack{y_1, \dots, y_N \in \mathbb{Z} \\ x_{i+1} \leq y_i < x_i \text{ for all } i}} \det[(y_i + c)^{\downarrow(\mu_j + N - j)}]_{i,j=1}^N. \quad (6.5)$$

The last column in the $(N+1) \times (N+1)$ matrix in the left-hand side of (6.5) consists of 1's. Subtracting from the i th row the $(i+1)$ st one for each $i = 1, \dots, N$, we see that the left-hand side is equal to the $N \times N$ determinant

$$\det[(x_i + c + 1)^{\downarrow(\mu_j + N + 1 - j)} - (x_{i+1} + c + 1)^{\downarrow(\mu_j + N + 1 - j)}]_{i,j=1}^N.$$

On the other hand, the summation in the right-hand side of (6.5) can be performed in each row separately using the relation

$$\sum_{y=a}^{b-1} (y+c)^{\downarrow m} = \frac{(b+c)^{\downarrow(m+1)} - (a+c)^{\downarrow(m+1)}}{m+1}.$$

Collecting constant prefactors completes the proof of Lemma 6.4:

$$\begin{aligned} & \frac{(N+1)_{\mu} N!}{(N)_{\mu} \prod_{j=1}^N (\mu_j + N - j + 1)} \\ &= N! \prod_{j=1}^N \frac{(\mu_j + N + 1 - j)! (N - j)!}{(N + 1 - j)! (\mu_j + N - j)! (\mu_j + N - j + 1)} = 1. \quad \square \end{aligned} \quad (6.6)$$

To conclude the proof of Proposition 6.2 we want to prove that, for a suitable fixed constant $c \in \mathbb{C}$, $\mathcal{D}^{(N)} F_{\mu,c}$ decompose on $\{F_{v,c}\}$ in exactly the same way as $\tilde{\mathcal{D}}^{(N+1)} G_{\mu,c}$ decompose on $\{G_{v,c}\}$.

It is actually convenient to take $c = v$, where v is one of the four parameters (u, u', v, v') . With this specialization we prove

Lemma 6.5. *For any $\lambda \in \mathbb{GT}_N$ and any nonnegative signature $\mu \in \mathbb{GT}_N$, with the notation $m_j = \mu_j + N - j$, $j = 1, \dots, N$, we have*

$$\begin{aligned} & \sum_{v \in \mathbb{GT}_N} \mathcal{D}^{(N)}(\lambda, v) F_{\mu,v}(v_1 + N - 1, \dots, v_N) \\ &= \left(\sum_{j=1}^N m_j(m_j - 1) + s \sum_{j=1}^N m_j - d_N \right) F_{\mu,v}(\lambda_1 + N - 1, \dots, \lambda_N) \\ &+ \sum_{j=1}^N ((m_j - 1)(v' - v + m_j - 1) + s(m_j - v - 1) + uu' - vv') \mathbf{1}_{\mu_j - 1 \geq \mu_{j+1}} \\ &\times F_{\mu - e_j, v}(\lambda_1 + N - 1, \dots, \lambda_N), \end{aligned} \quad (6.7)$$

where d_N is as in (5.10), $e_j = (0, \dots, 0, 1, 0, \dots, 0)$ with 1 at the j th place, and we assume $\mu_{N+1} = 0$.

Proof. We first compute, cf. (5.15),

$$\begin{aligned} \mathcal{D}(x+v)^{\downarrow m} &= (x+v)(x+v') \Delta \nabla (x+v)^{\downarrow m} + (sx + uu' - vv') \Delta (y+v)^{\downarrow m} \\ &= m(m-1)(x+v')(x+v)^{\downarrow(m-1)} + m(sx + uu' - vv')(x+v)^{\downarrow(m-1)}. \end{aligned} \quad (6.8)$$

This is the place where the choice of $c = v$ matters; for different values of c the expression for $\mathcal{D}(x+c)^{\downarrow m}$ would have been more complicated.

Substituting

$$\begin{aligned}x + v' &= (x + v - m + 1) + (v' - v + m - 1), \\sx + uu' - vv' &= s(x + v - m + 1) + (s(m - v - 1) + uu' - vv'),\end{aligned}$$

we obtain

$$\begin{aligned}\mathcal{D}(x + v)^{\downarrow m} &= (m(m - 1) + sm)(x + v)^{\downarrow m} + ((m - 1)(v' - v + m - 1) \\&\quad + s(m - v - 1) + uu' - vv')m(x + v)^{\downarrow(m-1)}.\end{aligned}\quad (6.9)$$

The statement now follows from (5.9) and the definition of $F_{\mu,c}$. \square

Let us complete the proof of Proposition 6.2.

Apply both sides of (6.3) to $F_{\mu,v}$ in the sense of Lemma 6.3. Using Lemma 6.4 we see that the left-hand side of (6.3) turns into

$$\sum_{\kappa \in \mathbb{G}\mathbb{T}_{N+1}} \tilde{\mathcal{D}}^{(N+1)}(\lambda, \kappa) G_{\mu,v}(\kappa_1 + N, \dots, \kappa_{n+1}),$$

and repeating the arguments of Lemma 6.5 we see that this is equal to

$$\begin{aligned}&\left(\sum_{j=1}^N \tilde{m}_j(\tilde{m}_j - 1) + \tilde{s} \sum_{j=1}^N \tilde{m}_j - \tilde{d}_{N+1} \right) G_{\mu,v}(\lambda_1 + N - 1, \dots, \lambda_N) \\&\quad + \sum_{j=1}^N ((\tilde{m}_j - 1)(v' - v + \tilde{m}_j - 1) + \tilde{s}(\tilde{m}_j - \tilde{v} - 1) + \tilde{u}\tilde{u}' - vv') \mathbf{1}_{\mu_j - 1 \geq \mu_{j+1}} \\&\quad \times G_{\mu - \delta_j, v}(\lambda_1 + N - 1, \dots, \lambda_N),\end{aligned}\quad (6.10)$$

where $\tilde{m}_j = \mu_j + N + 1 - 1 = m_j + 1$, and tildes over the other constants mean that in their definitions we replace (u, u') by $(\tilde{u}, \tilde{u}') = (u + 1, u' + 1)$.

On the other hand, by Lemmas 6.4 and 6.5 the right-hand side of (6.3) equals

$$\begin{aligned}&\left(\sum_{j=1}^N m_j(m_j - 1) + s \sum_{j=1}^N m_j - d_N \right) G_{\mu,v}(\lambda_1 + N - 1, \dots, \lambda_N) \\&\quad + \sum_{j=1}^N ((m_j - 1)(v' - v + m_j - 1) + s(m_j - v - 1) + uu' - vv') \\&\quad \times \mathbf{1}_{\mu_j - 1 \geq \mu_{j+1}} G_{\mu - \delta_j, v}(\lambda_1 + N - 1, \dots, \lambda_N).\end{aligned}\quad (6.11)$$

It is a straightforward computation to see that all the coefficients in (6.10) and (6.11) coincide. The proof of Proposition 6.2 is complete. \square

6.3. From matrices of transition rates to semigroups

In order to complete the proof of Theorem 6.1 we need the following lemma.

Lemma 6.6. *If f is a finitely supported function on \mathbb{GT}_N then $\Lambda_N^{N+1} f$ is in the domain of the generator A_{N+1} of the semigroup $(P_{N+1}(t))_{t \geq 0}$ (see Section 4 for the definition of the generator and its domain).*

Let us postpone the proof of Lemma 6.6 until the end of this subsection and proceed with the proof of Theorem 6.1.

In order to prove (2.3) it suffices to prove that the two sides are equal when applied to a function f on \mathbb{GT}_N with finite support (as such are dense in $C_0(\mathbb{GT}_N)$):

$$P_{N+1}(t) \Lambda_N^{N+1} f = \Lambda_N^{N+1} P_N(t) f, \quad t \geq 0, \quad N = 1, 2, \dots \quad (6.12)$$

Let us denote the left- and right-hand sides of (6.12) by $F_{\text{left}}(t)$ and $F_{\text{right}}(t)$. We will show that they solve the same Cauchy problem in the Banach space $C_0(\mathbb{GT}_{N+1})$. Then (6.12) will follow from an abstract uniqueness theorem for solutions of the Cauchy problem for vector functions with values in a Banach space,

$$\frac{d}{dt} F(t) = AF(t), \quad t > 0, \quad F(0) = \text{fixed vector},$$

which holds under the assumptions that (1) A is a closed dissipative operator, (2) $F(t)$ is continuous for $t \geq 0$ and strongly differentiable for $t > 0$, and (3) $F(t) \in D(A)$ for $t \geq 0$; see e.g. [27, IX.1.3].

In our situation, $A = A_{N+1}$ and the fixed vector is $\Lambda_N^{N+1} f$. Obviously, both $F_{\text{left}}(t)$ and $F_{\text{right}}(t)$ are continuous for $t \geq 0$ and they have the same initial value $\Lambda_N^{N+1} f$ at $t = 0$.

Let us check the differential equation for $F_{\text{left}}(t)$. By Lemma 6.6 we have $\Lambda_N^{N+1} f \in D(A_{N+1})$. Hence, $F_{\text{left}}(t) \in D(A_{N+1})$ (semigroups preserve the domains of the generators) and it satisfies

$$\frac{d}{dt} F_{\text{left}}(t) = A_{N+1} F_{\text{left}}(t), \quad t > 0.$$

Let us turn to F_{right} . By Corollary 4.7, f belongs to $D(A_N)$. It follows that the function $t \mapsto P_N(t)f$ is strongly differentiable and

$$\frac{d}{dt} P_N(t)f = A_N P_N(t)f = \mathcal{D}^{(N)} P_N(t), \quad t > 0.$$

Hence, $F_{\text{right}}(t)$ is also strongly differentiable and for $t > 0$

$$\frac{d}{dt} F_{\text{right}}(t) = \Lambda_N^{N+1} \frac{d}{dt} P_N(t)f = \Lambda_N^{N+1} \mathcal{D}^{(N)} P_N(t)f.$$

By definition, the last expression should be understood as $\Lambda_N^{N+1}(\mathcal{D}^{(N)}(P_N(t)f))$. However, since all rows of the matrices Λ_N^{N+1} and $\mathcal{D}^{(N)}$ have finitely many nonzero entries, we may write

$$\Lambda_N^{N+1}(\mathcal{D}^{(N)}(P_N(t)f)) = (\Lambda_N^{N+1}\mathcal{D}^{(N)})P_N(t)f.$$

By virtue of Proposition 6.2, this equals

$$\tilde{\mathcal{D}}^{(N+1)}\Lambda_N^{N+1}P_N(t)f = \tilde{\mathcal{D}}^{(N+1)}F_{right}(t),$$

so that

$$\frac{d}{dt}F_{right}(t) = \tilde{\mathcal{D}}^{(N+1)}F_{right}(t).$$

Next, as $\frac{d}{dt}F_{right}(t)$ is in $C_0(\mathbb{GT}_{N+1})$, so is $\tilde{\mathcal{D}}^{(N+1)}F_{right}(t)$. By Proposition 4.6, we may replace $\tilde{\mathcal{D}}^{(N+1)}$ by A_{N+1} , which gives the desired differential equation

$$\frac{d}{dt}F_{right}(t) = A_{N+1}F_{right}(t), \quad t > 0,$$

and we conclude that $F_{left} = F_{right}$.

Thus, we have proved Theorem 6.1 modulo Lemma 6.6.

Proof of Lemma 6.6. Let f be a finitely supported function on \mathbb{GT}_N , $g = \Lambda_N^{N+1}f$. Proposition 3.3 says that $g \in C_0(\mathbb{GT}_{N+1})$, and by Proposition 4.6 it suffices to check that $\mathcal{D}^{(N+1)}g \in C_0(\mathbb{GT}_{N+1})$. We have

$$(\mathcal{D}^{(N+1)}g)(\lambda) = \sum_{\varepsilon: \lambda + \varepsilon \in \mathbb{GT}_{N+1}} \mathcal{D}^{(N+1)}(\lambda, \lambda + \varepsilon)(g(\lambda + \varepsilon) - g(\lambda)),$$

where $\lambda \in \mathbb{GT}_{N+1}$, ε ranges over $\{\pm e_j\}_{j=1, \dots, N+1}$, with (e_j) being the standard basis in \mathbb{R}^{N+1} , and $\mathcal{D}^{(N+1)}(\lambda, \lambda + \varepsilon)$ are off-diagonal entries of the matrix $\mathcal{D}^{(N+1)}$.

Without loss of generality we may assume that f is the delta-function at some $\nu \in \mathbb{GT}_N$. We obtain

$$g(\lambda) = \frac{N! \prod_{1 \leq i < j \leq N} (v_i - i - v_j + j)}{\prod_{1 \leq i < j \leq N+1} (\lambda_i - i - \lambda_j + j)} \cdot \mathbf{1}_{\nu < \lambda},$$

and

$$(\mathcal{D}^{(N+1)}g)(\lambda) = \sum_{i=1}^{N+1} \sum_{\varepsilon_i = \pm 1} \frac{\text{Dim}_{N+1}(\lambda + \varepsilon_i e_i)}{\text{Dim}_{N+1}(\lambda)} \mathcal{D}(l_i, l_i + \varepsilon_i)(g(\lambda + \varepsilon_i e_i) - g(\lambda)) \quad (6.13)$$

where $l_j = \lambda_j + N + 1 - j$, $j = 1, \dots, N + 1$, and we assume $\text{Dim}_{N+1}(\lambda + \varepsilon_i e_i) = 0$ in case $\lambda + \varepsilon_i e_i \notin \mathbb{GT}_{N+1}$ (this is supported by the explicit formula for $\text{Dim}_{N+1}(\cdot)$).

Observe that for

$$\tilde{g}(\lambda) = \frac{N! \prod_{1 \leq i < j \leq N} (v_i - i - v_j + j)}{\prod_{1 \leq i < j \leq N+1} (\lambda_i - i - \lambda_j + j)} = \frac{\text{const}_1}{\text{Dim}_{N+1}(\lambda)}$$

(we removed the factor $\mathbf{1}_{v < \lambda}$ from $g(\lambda)$ above), we have

$$\begin{aligned} (\mathcal{D}^{(N+1)} \tilde{g})(\lambda) &= \frac{\text{const}_2}{\text{Dim}_{N+1}^2(\lambda)} \sum_{i=1}^{N+1} \sum_{\varepsilon_i = \pm 1} \mathcal{D}(l_i, l_i + \varepsilon_i) (\text{Dim}_{N+1}(\lambda + \varepsilon_i e_i) - \text{Dim}_{N+1}(\lambda)) \\ &= \frac{\text{const}_3}{\text{Dim}_{N+1}(\lambda)}, \end{aligned} \quad (6.14)$$

where we used (5.14).

Next, observe that the function

$$(\mathcal{D}^{(N+1)} \tilde{g})(\lambda) \mathbf{1}_{v < \lambda} = \frac{\text{const}_3}{\text{Dim}_{N+1}(\lambda)} \mathbf{1}_{v < \lambda}$$

belongs to $\mathbb{C}_0(\mathbb{GT}_{N+1})$. Indeed, if λ goes to infinity inside the subset $\{\lambda: v < \lambda\}$ then $\lambda_i - \lambda_j \rightarrow +\infty$ for at least one couple $i < j$ of indices, which entails $\text{Dim}_{N+1} \lambda \rightarrow +\infty$.

The discrepancy between $(\mathcal{D}^{(N+1)} \tilde{g})(\lambda) \mathbf{1}_{v < \lambda}$ and $(\mathcal{D}^{(N+1)} g)(\lambda)$ (or rather between the summations in (6.13) and (6.14)) comes from values of i and ε_i such that either $v < \lambda$ but $v \not\prec (\lambda + \varepsilon_i e_i)$, or $v < \lambda + \varepsilon_i e_i$ but $v \not\prec \lambda$. In both cases, for that value of i , the quantities λ_i, l_i , and $\mathcal{D}(l_i, l_i + \varepsilon_i)$ must remain bounded as v is fixed.

Note that $\lambda \rightarrow \infty$ inside the subset

$$\{\lambda \in \mathbb{GT}_{N+1}: v < \lambda \text{ or } v < \lambda + \varepsilon_i e_i \text{ for some } i\},$$

then either $\lambda_1 \rightarrow +\infty$ or $\lambda_{N+1} \rightarrow -\infty$, or both, while all other λ_j remain bounded from both sides. But then a direct inspection of the summands in (6.13) and (6.14) that contribute to the discrepancy shows that they converge to zero as $\lambda \rightarrow \infty$. Hence, $(\mathcal{D}^{(N+1)} g)(\lambda) \in C_0(\mathbb{GT}_{N+1})$. \square

7. Invariant measures

In three previous sections we defined a chain of countable sets $\{E_N = \mathbb{GT}_N\}_{N \geq 1}$, constructed links Λ_N^{N+1} between them, and identified the boundary $E_\infty = \Omega$. Furthermore, for any quadruple of complex parameters (z, z', w, w') satisfying (6.1) we constructed Feller semigroups $(P_N(t))_{t \geq 0}$ on \mathbb{GT}_N and showed that they are compatible with the links; by Proposition 2.4 this yields a Feller semigroup $(P(t))_{t \geq 0}$ on Ω .

The goal of this section is to exhibit an invariant measure for $(P(t))_{t \geq 0}$.

7.1. zw -measures

Let z, z', w, w' be complex parameters satisfying (6.1). As was pointed out in Section 5.1, this is equivalent to saying that each pair (z, z') and (w, w') belongs to one (or both) of the sets

$$\{(\zeta, \zeta') \in (\mathbb{C} \setminus \mathbb{Z})^2 \mid \zeta' = \bar{\zeta}\} \quad \text{and} \\ \{(\zeta, \zeta') \in (\mathbb{R} \setminus \mathbb{Z})^2 \mid m < \zeta, \zeta' < m + 1 \text{ for some } m \in \mathbb{Z}\},$$

and also $z + z' + w + w' > -1$.

For $\lambda \in \mathbb{GT}_N$ set

$$M_{z,z',w,w'|N}(\lambda) = (\text{const}_N)^{-1} \cdot M'_{z,z',w,w'|N}(\lambda)$$

where

$$M'_{z,z',w,w'|N}(\lambda) = \prod_{i=1}^N \left(\frac{1}{\Gamma(z - \lambda_i + i) \Gamma(z' - \lambda_i + i)} \right. \\ \left. \times \frac{1}{\Gamma(w + N + 1 + \lambda_i - i) \Gamma(w' + N + 1 + \lambda_i - i)} \right) \cdot (\text{Dim}_N(\lambda))^2, \quad (7.1)$$

and

$$\text{const}_N = \sum_{\lambda \in \mathbb{GT}_N} M'_{z,z',w,w'|N}(\lambda)$$

is the normalizing constant depending on z, z', w, w', N .

Theorem 7.1. (See [44].) Under our assumptions on the parameters, for any $N \geq 1$, $M_{z,z',w,w'|N}$ is a probability measure, we call it the N th zw -measure. Moreover, these measures are consistent with the links,

$$M_{z,z',w,w'|N} = M_{z,z',w,w'|N+1} \Lambda_N^{N+1}, \quad N \geq 1,$$

with Λ_N^{N+1} as in Section 3.1.

Theorem 7.1 implies that the system $(M_{z,z',w,w'|N})_{N \geq 1}$ defines a probability measure $M_{z,z',w,w'}$ on the boundary Ω that we call the *spectral zw -measure*, cf. Theorem 3.1, and a character of the infinite-dimensional unitary group $U(\infty)$, cf. [44]. For $z' = \bar{z}$ and $w' = \bar{w}$ one can find a geometric construction of the corresponding representations of $U(\infty)$ in [44]. There is also a fairly simple “coordinate-free” description of general zw -measures that we now give, cf. [13].

Let \mathbb{T} be the unit circle in \mathbb{C} and \mathbb{T}^N be the product of N copies of \mathbb{T} (the N -dimensional torus). For any $\lambda \in \mathbb{GT}_N$, the character χ^λ of the corresponding irreducible representation π_λ of $U(N)$ can be viewed as a symmetric function on \mathbb{T}^N , where coordinates are interpreted as eigenvalues of unitary matrices. Explicitly, the character is given by the (rational) Schur function

$$\chi^\lambda(u_1, \dots, u_N) = s_\lambda(u_1, \dots, u_N) = \frac{\det[u_i^{\lambda_j + N - j}]_{1 \leq i, j \leq N}}{\det[u_i^{N - j}]_{1 \leq i, j \leq N}}.$$

Consider the Hilbert space H_N of symmetric functions on \mathbb{T}^N , square integrable with respect to the measure

$$\frac{1}{N!} \prod_{1 \leq i < j \leq N} |u_i - u_j|^2 \prod_{i=1}^N du_i,$$

which is the push-forward of the normalized Haar measure on $U(N)$ under the correspondence $U \mapsto (u_1, \dots, u_N)$. Here du_i is the normalized invariant measure on the i th copy of \mathbb{T} .

Given two complex numbers z, w , we define a symmetric function on \mathbb{T}^N by

$$f_{z,w|N}(u) = \prod_{i=1}^N (1 + u_i)^z (1 + \bar{u}_i)^w.$$

If $\Re(z + w) > -\frac{1}{2}$ then $f_{z,w|N}$ belongs to the space H_N . Let (z', w') be another couple of complex numbers with $\Re(z' + w') > -\frac{1}{2}$. We set

$$M_{z,z',w,w'|N}(\lambda) = \frac{(f_{z,w|N}, \chi_\lambda)(\chi_\lambda, f_{\bar{w}', \bar{z}'|N})}{(f_{z,w|N}, f_{\bar{w}', \bar{z}'|N})}, \quad \lambda \in \mathbb{GT}_N,$$

where (\cdot, \cdot) is the inner product in H_N . It turns out that this definition leads us to the explicit formula given above.

The spectral zw -measures were the subject of an extensive investigation in [11] the upshot of which is the statement that with $\omega \in \Omega$ distributed according to $M_{z,z',w,w'}$, its coordinates

$$\left\{ \frac{1}{2} + \alpha_i^+, \frac{1}{2} - \beta_i^+, -\frac{1}{2} + \beta_i^-, -\frac{1}{2} - \alpha_i^- \right\}_{i=1}^\infty$$

(where possible zero values of α_i^\pm and β_i^\pm should be removed) form a *determinantal point process* on $\mathbb{R} \setminus \{\pm \frac{1}{2}\}$ with an explicit correlation kernel. See [11,12] for details.

7.2. Invariance

The main statement of this section is

Theorem 7.2. *For any quadruple (z, z', w, w') of parameters satisfying (6.1), the spectral zw -measure $M_{z,z',w,w'}$ is the unique invariant probability measure with respect to the semigroup $(P(t))_{t \geq 0}$.*

Proof. Let us prove the invariance first. By Section 2.8, it suffices to verify that for each $N \geq 1$, the N th level zw -measure is invariant with respect to $(P_N(t))_{t \geq 0}$. We will check this fact on the level of matrices of transition rates:

$$\sum_{\lambda \in \mathbb{GT}_N} M_{z,z',w,w'|N}(\lambda) \mathcal{D}^{(N)}(\lambda, \nu) = 0, \quad N \geq 1, \quad \nu \in \mathbb{GT}_N.$$

Since it is easy to check that $\mathcal{D}^{(N)}$ is reversible with respect to $M_{z,z',w,w'|N}(\lambda)$,

$$M_{z,z',w,w'|N}(\lambda)\mathcal{D}^{(N)}(\lambda,v) = \mathcal{D}^{(N)}(\lambda,v)M_{z,z',w,w'|N}(v), \quad \lambda, v \in \mathbb{GT}_N,$$

an argument in Section 3 of [30] shows that the invariance on the level of transition rates implies the invariance with respect to the corresponding semigroup.

As in the proof of Theorem 5.4, it is convenient to employ the bijection $\lambda \leftrightarrow (\lambda_j + N - j)_{1 \leq j \leq N}$ between \mathbb{GT}_N and \mathfrak{X}_N , see Section 5.2 for the notation and also recall that we are using parameterization (6.2). Under the bijection of \mathbb{GT}_N and \mathfrak{X}_N , the desired identity takes the form (removing irrelevant prefactors)

$$\sum_{X \in \mathfrak{X}_N} \left(\prod_{i=1}^N W(x_i) \right) V_N(X) \left((\mathcal{D}(x_1, y_1) \mathbf{1}_{\{x_i=y_i, i \neq 1\}} + \cdots + \mathcal{D}(x_N, y_N) \mathbf{1}_{\{x_i=y_i, i \neq N\}}) - d_N \mathbf{1}_{X=Y} \right) = 0, \quad (7.2)$$

where

$$W(x) = \frac{1}{\Gamma(z + N - x) \Gamma(z' + N - x) \Gamma(w + 1 + x) \Gamma(w' + 1 + x)}, \quad x \in \mathbb{Z}.$$

Let $p_0 = 1, p_1, p_2, \dots, \deg p_j = j$, be monic orthogonal polynomials on \mathbb{Z} corresponding to the weight function $W(x)$. As

$$W(x) = O(|x|^{-z-z'-w-w'-2N}), \quad x \rightarrow \infty,$$

the assumption $z + z' + w + w' > -1$ implies that $W(x)$ has at least $2N - 2$ finite moments, and polynomials p_j with $j = 0, 1, \dots, N - 1$ are well defined.

Polynomials $\{p_j\}$ can be written explicitly in terms of the hypergeometric function ${}_3F_2$ evaluated at 1. They were discovered by R. Askey [2], and independently by P. Lesky [34,35]; see also the recent book [32, §5.3, Theorem 5.2, Case IIIc]. We call them the *Askey–Lesky polynomials*.

The Askey–Lesky polynomials are eigenfunctions of the operator \mathcal{D} on \mathbb{Z} , see [11, §7]:

$$\sum_{y \in \mathbb{Z}} \mathcal{D}(x, y) p_j(y) = \gamma_j p_j(x) \quad \forall x \in \mathbb{Z}, \quad j = 0, 1, 2, \dots,$$

where

$$\gamma_j = j((j-1) - (u_N + u'_N + v_N + v'_N)).$$

Multiplying both sides by $W(x)$ and using the fact that $W(x)\mathcal{D}(x, y)$ is symmetric with respect to transposition $x \leftrightarrow y$ we obtain

$$\sum_{x \in \mathbb{Z}} p_j(x) W(x) \mathcal{D}(x, y) = \gamma_j p_j(y) \quad \forall y \in \mathbb{Z}, \quad j = 0, 1, 2, \dots \quad (7.3)$$

Let us rewrite the Vandermonde determinant in the left-hand side of (7.2) as

$$V_N(x) = \pm \det[p_{i-1}(x_j)]_{i,j=1}^N.$$

Applying operators $\mathcal{D}_1, \dots, \mathcal{D}_N$ to individual columns in this determinant multiplied by $W(x_1) \cdots W(x_N)$ according to (7.3), and recalling the definition of d_N , we obtain (7.2).

Let us now prove uniqueness. As explained in Section 2.8, it suffices to show that the N th level zw -measure is the unique invariant probability measure for $(P_N(t))_{t \geq 0}$ for any $N \geq 1$. But uniqueness of invariant measures holds in general for irreducible Markov chains on countable sets, see e.g. Theorem 1.6 in [1]. \square

8. Stochastic dynamics on paths. General formalism

8.1. Overview

Let us return to the general setting of Section 2 and assume that all E_N 's are discrete. For $N = 1, 2, \dots$ set

$$E^{(N)} = \left\{ (x_1, \dots, x_N) \in E_1 \times \cdots \times E_N \mid \prod_{k=1}^{N-1} \Lambda_k^{k+1}(x_{k+1}, x_k) \neq 0 \right\}. \quad (8.1)$$

There are natural projections $\Pi_N^{N+1} : E^{(N+1)} \rightarrow E^{(N)}$ consisting in forgetting the last coordinate; let $E^{(\infty)} = \varprojlim E^{(N)}$, where the projective limit is taken with respect to these projections. Obviously, $E^{(\infty)}$ is a closed subset of the infinite product space $\prod_{N=1}^{\infty} E_N$. Thus, elements of $E^{(\infty)}$ are some infinite sequences. Let $\Pi_N^{\infty} : E^{(\infty)} \rightarrow E^{(N)}$ be the map that extracts the first N members of such a sequence.

Definition 8.1. We say that a probability measure $\mu^{(N)}$ on $E^{(N)}$ is *central* if there exists a probability measure μ_N on E_N such that

$$\mu^{(N)}(x_1, \dots, x_N) = \mu_N(x_N) \Lambda_{N-1}^N(x_N, x_{N-1}) \cdots \Lambda_1^2(x_2, x_1) \quad (8.2)$$

for any $(x_1, \dots, x_N) \in E^{(N)}$. Relation (8.2) establishes a bijection between probability measures on E_N and central probability measures on $E^{(N)}$.

We say that $\mu^{(\infty)} \in \mathcal{M}_p(E^{(\infty)})$ is central if all its push-forwards under projections Π_N^{∞} are central. Relation (8.2) also establishes a bijection between central measures on $E^{(\infty)}$ and elements of $\varprojlim \mathcal{M}_p(E_N)$ of Section 2.5.

Finally, we say that a Markov semigroup $(P^{(N)}(t))_{t \geq 0}$ on $E^{(N)}$ is *central* if the associate linear operators in $\mathcal{M}(E^{(N)})$ map central measures to central measures.

Clearly, a central Markov semigroup $(P^{(N)}(t))_{t \geq 0}$ defines a Markov semigroup on E_N — in order to obtain $\mu_N P_N(t)$ for $\mu_N \in \mathcal{M}_p(E_N)$ one needs to define $\mu^{(N)}$ via (8.2), evaluate $\mu^{(N)} P^{(N)}(t)$, and read off a measure on E_N using Definition 8.1.

Proposition 8.2. Let $(P^{(N)}(t))_{t \geq 0}$, $N \geq 1$, be a sequence of central Markov semigroups on $E^{(N)}$'s that are compatible with the system of projections:

$$P^{(N+1)}(t) \circ \Pi_N^{N+1} = \Pi_N^{N+1} \circ P^{(N)}(t), \quad t \geq 0, \quad N \geq 1.$$

Then the corresponding Markov semigroups $(P_N(t))_{t \geq 0}$ on E_N , $N \geq 1$, are compatible with projections Λ_N^{N+1} as in (2.3).

Proof. Follows from the fact that if $\mu^{(N+1)}$ and μ_{N+1} are related as in Definition 8.1 then $\mu^{(N+1)} \prod_N^{N+1}$ and $\mu_{N+1} \Lambda_N^{N+1}$ are also related in the same way. \square

The goal of this section and the next one is to construct central Markov semigroups $(P^{(N)}(t))_{t \geq 0}$ that would yield, as in Proposition 8.2, semigroups $(P_N(t))_{t \geq 0}$ on $E_N = \mathbb{GT}_N$ that we dealt with in the previous sections. One reason for such a construction is the fact that for the Gelfand–Tsetlin graph, the isomorphism between central measures on $\mathbb{GT}^{(\infty)}$ and probability measures on the boundary Ω , cf. Definition 2.2, is somewhat explicit, see Section 9 below. Thus, $(P^{(N)}(t))_{t \geq 0}$ can be thought of as providing a more “hands-on” description of the corresponding semigroup $(P(t))_{t \geq 0}$ on Ω .

8.2. Construction of bivariate Markov chains

Let E and E^* be countable sets, and let Q and Q^* be matrices of transition rates on these sets. Let $\Lambda = [\Lambda(x^*, x)]_{x^* \in E^*, x \in E}$ be an additional stochastic matrix which we view as a stochastic link between E^* and E .

We will assume that for each of the three matrices Q , Q^* , and Λ , each row contains only finitely many nonzero entries. In addition, we assume the relation

$$\sum_{x \in E} \Lambda(x^*, x) Q(x, y) = \sum_{y^* \in E^*} Q^*(x^*, y^*) \Lambda(y^*, y), \quad x^* \in E^*, y \in E, \quad (8.3)$$

or $\Lambda Q = Q^* \Lambda$ in matrix notation.

Observe that in case $\Lambda(x^*, y) = 0$, the diagonal entries $Q(x, x)$ and $Q(x^*, x^*)$ give no contribution to (8.3), and the commutativity relation can be rewritten as

$$\sum_{x \in E, x \neq y} \Lambda(x^*, x) Q(x, y) = \sum_{y^* \in E^*, y^* \neq x^*} Q^*(x^*, y^*) \Lambda(y^*, y), \quad x^* \in E^*, y \in E. \quad (8.4)$$

We will denote the above expression by $\Delta(x^*, y)$; it is only defined if $\Lambda(x^*, y) = 0$.

In what follows we also use the notation

$$q_x = -Q(x, x), \quad x \in E; \quad q_{x^*}^* = -Q^*(x^*, x^*), \quad x^* \in E^*.$$

Consider the bivariate state space

$$E^{(2)} = \{(x^*, x) \in E^* \times E \mid \Lambda(x^*, x) \neq 0\}.$$

We want to construct a Markov chain on $E^{(2)}$ that would satisfy three conditions:

- The projection of this Markov chain to E gives the Markov chain defined by Q ;
- It preserves the class of measures on $E^{(2)}$ satisfying $\text{Prob}(x|x^*) = \Lambda(x^*, x)$;
- In this class of measures, the projection of this Markov chain to E^* gives the Markov chain defined by Q^* .

To this end, define a matrix $Q^{(2)}$ of transition rates on $E^{(2)}$ with off-diagonal entries given by

$$Q^{(2)}((x^*, x), (y^*, y)) = \begin{cases} Q(x, y), & x^* = y^*, \\ Q^*(x^*, y^*) \frac{\Lambda(y^*, x)}{\Lambda(x^*, x)}, & x = y, \\ Q(x, y) \frac{Q^*(x^*, y^*) \Lambda(y^*, y)}{\Delta(x^*, y)}, & \Lambda(x^*, y) = 0, \Delta(x^*, y) \neq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Note that $\Lambda(x^*, y) = 0$ implies $x^* \neq y^*$ and $x \neq y$ (provided that $(x^*, x), (y^*, y)$ are in $E^{(2)}$) so all the cases in the above definition are mutually exclusive.

The diagonal entries $Q^{(2)}((x^*, x), (x^*, x))$ with $(x^*, x) \in E^{(2)}$ are defined by

$$-Q^{(2)}((x^*, x), (x^*, x)) = q_{(x^*, x)}^{(2)} := \sum_{(y^*, y) \neq (x^*, x)} Q^{(2)}((x^*, x), (y^*, y)).$$

Clearly, any row of $Q^{(2)}$ also has only finitely many nonzero entries. One immediately verifies that for any $(x^*, x) \in E^{(2)}$ and $y \in E$ with $x \neq y$,

$$\sum_{y^*: (y^*, y) \in E^{(2)}} Q^{(2)}((x^*, x), (y^*, y)) = Q(x, y). \quad (8.5)$$

Indeed, one needs to consider two cases $\Lambda(x^*, y) = 0$ and $\neq 0$, and in both cases the statement follows from the definitions. As the row sums of $Q^{(2)}$ and Q are all zero, we obtain (8.5) for $x = y$ as well.

For any $x \in E$, let us also introduce a matrix of transition rates Q_x on the fiber $E_x = \{x^* \in E^* \mid \Lambda(x^*, x) \neq 0\}$ via

$$Q_x(x^*, y^*) = Q^{(2)}((x^*, x), (y^*, x)) = Q^*(x^*, y^*) \frac{\Lambda(y^*, x)}{\Lambda(x^*, x)}, \quad x^* \neq y^*,$$

and

$$Q_x(x^*, x^*) = - \sum_{y^* \in E_x, y^* \neq x^*} Q_x(x^*, y^*).$$

The following statement is similar to Lemma 2.1 of [6] proved in the discrete time setting. As we will see, the proof of the continuous time statement is significantly more difficult.

Proposition 8.3. Assume that the matrices of transition rates Q , Q^* , and Q_x for any $x \in E$ are regular. Then $Q^{(2)}$ is also regular, and denoting by $P(t)$, $P^*(t)$, and $P^{(2)}(t)$ the transition matrices corresponding to Q , Q^* , and $Q^{(2)}$, we have

$$\sum_{y^*: (y^*, y) \in E^{(2)}} P^{(2)}(t; (x^*, x), (y^*, y)) = P(t; x, y), \quad (8.6)$$

$$\sum_{x: (x^*, x) \in E^{(2)}} \Lambda(x^*, x) P^{(2)}(t; (x^*, x), (y^*, y)) = P^*(t; x^*, y^*) \Lambda(y^*, y), \quad (8.7)$$

where in the first relation $(x^*, x) \in E^{(2)}$, $y \in E$ are arbitrary, while in the second relation $x^* \in E^*$, $(y^*, y) \in E^{(2)}$ are arbitrary.

Proof. The regularity of $Q^{(2)}$ and collapsibility relation (8.6) follow from Proposition 4.4 with (4.5) specializing to (8.5).

Proving (8.7) is more difficult, and we will follow the following path. First, we will show that both sides of (8.7) satisfy the same differential equation (essentially the Kolmogorov backward equation for $P^*(t; x^*, y^*)$) with a certain initial condition. Then we will see that the right-hand side of (8.7) represents the minimal of all nonnegative solutions of this equation. Since for a fixed x^* , both sides of (8.7) represent probability measures on $E^{(2)}$, the equality will immediately follow.

For the first step, let us show that the left-hand side $f_t(x^*, y^*, y)$ of (8.7) satisfies

$$\frac{d}{dt} f_t(x^*, y^*, y) = \sum_{z^* \in E^*} Q^*(x^*, z^*) f_t(z^*, y^*, y) \quad (8.8)$$

with the initial condition

$$\lim_{t \rightarrow +0} f_t(x^*, y^*, y) = \mathbf{1}_{x^*=y^*} \Lambda(y^*, y). \quad (8.9)$$

The initial condition satisfied by $P^{(2)}(t)$ implies (8.9), so let us prove (8.8).

Using the Kolmogorov backward equation for $P^{(2)}(t)$, we obtain

$$\begin{aligned} \frac{d}{dt} f_t(x^*, y^*, y) = & \sum_{x: (x^*, x) \in E^{(2)}} \Lambda(x^*, x) \left(-q_{(x^*, x)}^{(2)} P^{(2)}(t; (x^*, x), (y^*, y)) \right. \\ & \left. + \sum_{(z^*, z) \neq (x^*, x)} Q^{(2)}((x^*, x), (z^*, z)) P^{(2)}(t; (z^*, z), (y^*, y)) \right). \end{aligned} \quad (8.10)$$

For the first term in the right-hand side, we use

$$q_{(x^*, x)}^{(2)} = q_x + \sum_{w^*: w^* \neq x^*} Q^*(x^*, w^*) \frac{\Lambda(w^*, x)}{\Lambda(x^*, x)},$$

which follows directly from the definition of $Q^{(2)}$. Thus, we can rewrite the first term in the right-hand side of (8.10) as

$$\begin{aligned} & - \sum_{x \in E} q_x \Lambda(x^*, x) P^{(2)}(t; (x^*, x), (y^*, y)) \\ & - \sum_{x: (x^*, x) \in E^{(2)}} \sum_{w^*: w^* \neq x^*} \Lambda(w^*, x) Q^*(x^*, w^*) P^{(2)}(t; (x^*, x), (y^*, y)). \end{aligned} \quad (8.11)$$

For the second term of the right-hand side of (8.10), according to the definition of $Q^{(2)}$, let us split the sum over (z^*, z) into three disjoint parts: (1) $x^* = z^*$, $x \neq z$; (2) $x^* \neq z^*$, $x = z$; (3) $\Lambda(x^*, z) = 0$ (hence, $x^* \neq z^*$, $x \neq z$).

Part (1) gives

$$(1) = \sum_{x: (x^*, x) \in E^{(2)}} \Lambda(x^*, x) \sum_{z: z \neq x, (x^*, z) \in E^{(2)}} Q(x, z) P^{(2)}(t; (x^*, z), (y^*, y)).$$

Interchanging the summations over x and z , we can employ the commutativity relation (8.3). This gives

$$\begin{aligned} (1) = & \sum_{z: (x^*, z) \in E^{(2)}} \sum_{v^* \in E^*} \Lambda(v^*, z) Q^*(x^*, v^*) P^{(2)}(t; (x^*, z), (y^*, y)) \\ & + \sum_{x \in E} q_x \Lambda(x^*, x) P^{(2)}(t; (x^*, x), (y^*, y)). \end{aligned} \quad (8.12)$$

Observe that the last term cancels out with the first term in (8.11), while the sum of the first term of (8.12) and the second term of (8.11), with identification $z = x$, $v^* = w^*$ of the summation variables, yields (only terms with $v^* = x^*$ survive)

$$-q_{x^*} \sum_{x \in E} \Lambda(x^*, x) P^{(2)}(t; (x^*, x), (y^*, y)). \quad (8.13)$$

Further, part (2) of the second term of (8.10) reads

$$(2) = \sum_{x: (x^*, x) \in E^{(2)}} \sum_{z^*: z^* \neq x^*} \Lambda(z^*, x) Q^*(x^*, z^*) P^{(2)}(t; (z^*, x), (y^*, y)). \quad (8.14)$$

Finally, part (3) gives

$$\begin{aligned} (3) = & \sum_{x: (x^*, x) \in E^{(2)}} \sum_{(z^*, z) \in E^{(2)}: \Lambda(x^*, z)=0} \Lambda(x^*, x) Q(x, z) \frac{Q^*(x^*, z^*) \Lambda(z^*, z)}{\Delta(x^*, z)} \\ & \times P^{(2)}(t; (z^*, z), (y^*, y)) \\ = & \sum_{z: (x^*, z) \notin E^{(2)}} \sum_{z^*: z^* \neq x^*} \Lambda(z^*, z) Q^*(x^*, z^*) P^{(2)}(t; (z^*, z), (y^*, y)), \end{aligned} \quad (8.15)$$

where we used the definition of Δ , see (8.4), to perform the summation over $x \neq z$. One readily sees that adding (8.13), (8.14), (8.15) yields the right-hand side of (8.8).

Assume now that we have a nonnegative solution $f_t(x^*, y^*, y)$ of (8.8) satisfying the initial condition (8.9). Multiplying both sides of (8.8) by $\exp(q_{x^*}^* t)$ we obtain

$$(\exp(q_{x^*}^* t) f_t(x^*, y^*, y))' = \exp(q_{x^*}^* t) \sum_{z^* \neq x^*} Q^*(x^*, z^*) f_t(z^*, y^*, y).$$

Integrating both sides over t and using (8.9) gives

$$f_t(x^*, y^*, y) = \mathbf{1}_{x^*=y^*} \Lambda(y^*, y) \exp(-q_{x^*}^* t) + \int_0^t \exp(-q_{x^*}^* s) \sum_{z^* \neq x^*} Q^*(x^*, z^*) f_{t-s}(z^*, y^*, y) ds. \quad (8.16)$$

Set $F_t^{(0)}(x^*, y^*) = \mathbf{1}_{x^*=y^*} \exp(-q_{x^*}^* t)$, and for $n = 1, 2, \dots$ define

$$F_t^{(n)}(x^*, y^*) = F_t^{(0)}(x^*, y^*) + \int_0^t \exp(-q_{x^*}^* s) \sum_{z^* \neq x^*} Q^*(x^*, z^*) F_{t-s}^{(n-1)}(z^*, y^*) ds.$$

Clearly, (8.16) implies $f_t(x^*, y^*, y) \geq F_t^{(0)}(x^*, y^*) \Lambda(y^*, y)$, and substituting such estimates into (8.16) recursively we see that

$$f_t(x^*, y^*, y) \geq F_t^{(n)}(x^*, y^*) \Lambda(y^*, y), \quad n = 0, 1, 2, \dots$$

On the other hand, we know that

$$\lim_{n \rightarrow \infty} F_t^{(n)}(x^*, y^*) = P^*(t; x^*, y^*),$$

see Section 4, [20,1]. Hence, any nonnegative solution of (8.8), (8.9) is bounded by $P^*(t; x^*, y^*) \Lambda(y^*, y)$ from below, and the proof of Proposition 8.3 is complete. \square

The following statement is the analog of Proposition 2.2 in [6].

Corollary 8.4. *Let $\mu^*(x^*)$ be a probability measure on E^* . For $t \geq 0$, let $(x^*(t), x(t))$ be an $E^{(2)}$ -valued random variable with*

$$\text{Prob}\{(x^*(t), x(t)) = (x^*, x)\} = \sum_{(y^*, y) \in E^{(2)}} \mu^*(y^*) \Lambda(y^*, y) P^{(2)}(t; (y^*, y), (x^*, x)).$$

Then for any time moments $0 \leq t_0 \leq t_1 \leq \dots \leq t_k \leq t_{k+1} \leq \dots \leq t_{k+l}$, the joint distribution of

$$(x^*(t_0), x^*(t_1), \dots, x^*(t_k), x(t_k), x(t_{k+1}), \dots, x(t_{k+l}))$$

coincides with the stochastic evolution of μ^* under transition matrices

$$(P^*(t_0), P^*(t_1 - t_0), \dots, P^*(t_k - t_{k-1}), \Lambda, P(t_{k+1} - t_k), \dots, P(t_{k+l} - t_{k+l-1})).$$

Proof. In the joint distribution

$$\begin{aligned} & \mu^*(y^*) \Lambda(y^*, y) P^*(t_0; (y^*, y), (x_0^*, x_0)) P^*(t_1 - t_0; (x_0^*, x_0), (x_1^*, x_1)) \cdots \\ & \times P^*(t_{k+l} - t_{k+l-1}; (x_{k+l-1}^*, x_{k+l-1}), (x_{k+l}^*, x_{k+l})) \end{aligned} \quad (8.17)$$

one uses (8.7) to sum over y, x_0, \dots, x_{k-1} and (8.6) to sum over $x_{k+1}^*, \dots, x_{k+l}^*$. \square

8.3. Construction of multivariate Markov chains

Let E_1, \dots, E_N be countable sets, Q_1, \dots, Q_N be matrices of transition rates on these sets, and $\Lambda_1^2, \dots, \Lambda_{N-1}^N$ be stochastic links:

$$\Lambda_{k-1}^k : E_k \times E_{k-1} \rightarrow [0, 1], \quad \sum_{y \in E_{k-1}} \Lambda_{k-1}^k(x, y) = 1, \quad x \in E_k, \quad k = 2, \dots, N.$$

It is also convenient to introduce a formal symbol Λ_0^1 with $\Lambda_0^1(\cdot, \cdot) \equiv 1$. It can be viewed as a stochastic link between E_1 and a singleton E_0 .

We assume that for each of the matrices Q_j , Λ_{j-1}^j , each row contains only finitely many nonzero entries, and that the following commutativity relations are satisfied:

$$\sum_{u \in E_{k-1}} \Lambda_{k-1}^k(x, u) Q_{k-1}(u, y) = \sum_{v \in E_k} Q_k(x, v) \Lambda_{k-1}^k(v, y), \quad k = 2, \dots, N,$$

or $\Lambda_{k-1}^k Q_{k-1} = Q_k \Lambda_{k-1}^k$ in matrix notation. If $\Lambda_{k-1}^k(x, y) = 0$, the terms with $u = y$ and $v = x$ give no contribution to the sums and thus can be excluded. In that case we define ($x \in E_k$, $y \in E_{k-1}$)

$$\Delta_{k-1}^k(x, y) := \sum_{u: u \neq y} \Lambda_{k-1}^k(x, u) Q_{k-1}(u, y) = \sum_{v: v \neq x} Q_k(x, v) \Lambda_{k-1}^k(v, y), \quad (8.18)$$

and also

$$\hat{Q}_k(x, v, y) = \begin{cases} \frac{Q_k(x, v) \Lambda_{k-1}^k(v, y)}{\Delta_{k-1}^k(x, y)}, & \text{if } \Delta_{k-1}^k(x, y) \neq 0, \\ 0, & \text{if } \Delta_{k-1}^k(x, y) = 0. \end{cases}$$

In case $\Delta_{k-1}^k(x, y) \neq 0$, $Q_k(x, v, y)$ is a probability distribution in $v \in E_k$ that depends on x and y .

In the application of this formalism that we consider in the next section, there is always exactly one v that contributes nontrivially to the right-hand side of (8.18), which means that the distribution $\hat{Q}_k(x, v, y)$ is supported by one point.

We define the state space $E^{(N)}$ for the multivariate Markov chain by (8.1) and then define the off-diagonal entries of the matrix $Q^{(N)}$ of transition rates on $E^{(N)}$ as (we use the notation $X_N = (x_1, \dots, x_N)$, $Y_N = (y_1, \dots, y_N)$)

$$Q^{(N)}(X_N, Y_N) = \begin{cases} Q_k(x_k, y_k) \frac{\Lambda_{k-1}^k(y_k, x_{k-1})}{\Lambda_{k-1}^k(x_k, x_{k-1})}, \\ Q_k(x_k, y_k) \frac{\Lambda_{k-1}^k(y_k, x_{k-1})}{\Lambda_{k-1}^k(x_k, x_{k-1})} \hat{Q}_{k+1}(x_{k+1}, y_{k+1}, y_k) \cdots \hat{Q}_l(x_l, y_l, y_{l-1}), \end{cases}$$

where for the first line we must have $x_j = y_j$ for all $j \neq k$ and some $k = 1, \dots, N$, while for the second line we must have $x_j = y_j$ iff $j < k$ or $j > l$ for some $1 \leq k < l \leq N$, and $\Lambda(x_j, y_{j-1}) = 0$ for $k + 1 \leq j \leq l$. If neither of the two sets of conditions is satisfied, we set $Q^{(N)}(X_N, Y_N) = 0$.

The diagonal entries $Q^{(N)}(X_N, X_N)$ are defined by

$$Q^{(N)}(X_N, X_N) = - \sum_{Y_N \neq X_N} Q^{(N)}(X_N, Y_N).$$

The definition of $Q^{(N)}$ can be interpreted as follows: Each of the coordinates $x_k, k = 1, \dots, N$, is attempting to jump to $y_k \in E_k$ with certain rates. Only y_k 's with $Q(x_k, y_k) \neq 0$ are eligible. Three situations are possible:

(1) The change of x_k to y_k does not move X_N out of the state space, that is $\Lambda_k^{k+1}(x_{k+1}, y_k) \Lambda_{k-1}^k(y_k, x_{k-1}) \neq 0$. Such jumps have rates $Q_k(x_k, y_k) \frac{\Lambda_{k-1}^k(y_k, x_{k-1})}{\Lambda_{k-1}^k(x_k, x_{k-1})}$. Note that for $k = 1$ the last factor is always 1.

(2) The change of x_k to y_k is in conflict with x_{k-1} , that is $\Lambda_{k-1}^k(y_k, x_{k-1}) = 0$. Such jumps are blocked.

(3) The change of x_k to y_k is in conflict with x_{k+1} , that is $\Lambda_k^{k+1}(x_{k+1}, y_k) = 0$. Then x_{k+1} has to be changed too, say to y_{k+1} . We must have $\Lambda_k^{k+1}(y_{k+1}, y_k) \neq 0$; relation (8.18) guarantees the existence of at least one such y_{k+1} . If the double jump $(x_k, x_{k+1}) \rightarrow (y_k, y_{k+1})$ keeps X_N in the state space, it is allowed, and its rate is $Q_k(x_k, y_k) \frac{\Lambda_{k-1}^k(y_k, x_{k-1})}{\Lambda_{k-1}^k(x_k, x_{k-1})} \hat{Q}_{k+1}(x_{k+1}, y_{k+1}, y_k)$. Otherwise, x_{k+2} has to be changed as well, and so on.

To say it differently, unless $\Lambda_{k-1}^k(y_k, x_{k-1}) = 0$, the move $x_k \rightarrow y_k$ always happens with rate $Q_k(x_k, y_k) \frac{\Lambda_{k-1}^k(y_k, x_{k-1})}{\Lambda_{k-1}^k(x_k, x_{k-1})}$, and it may cause a sequence of displacements of x_{k+1}, x_{k+2}, \dots , where each next x_j uses the distribution $\hat{Q}_j(x_j, \cdot, y_{j-1})$ to choose its new position. Displacements end once X_N is back in $E^{(N)}$. This description implies the following formula for the diagonal entries of $Q^{(N)}$:

$$Q^{(N)}(X_N, X_N) = - \sum_{k=1}^N \sum_{y_k \in E_k: y_k \neq x_k} Q_k(x_k, y_k) \frac{\Lambda_{k-1}^k(y_k, x_{k-1})}{\Lambda_{k-1}^k(x_k, x_{k-1})}. \quad (8.19)$$

The definition of $Q^{(N)}$ is explained by the following statement.

Proposition 8.5. Consider the matrix Λ with rows marked by elements of E_N , columns marked by $E^{(N-1)}$, and entries given by

$$\Lambda(x_N, (x_1, \dots, x_{N-1})) = \Lambda_{N-1}^N(x_N, x_{N-1}) \cdots \Lambda_1^2(x_2, x_1). \quad (8.20)$$

Then the commutativity relation $\Lambda Q^{(N-1)} = Q_N \Lambda$ holds.

Proof. We have

$$\Lambda Q^{(N-1)}(x_N, Y_{N-1}) = \sum_{X_{N-1} \in E^{(N-1)}} \Lambda(x_N, X_{N-1}) Q^{(N-1)}(X_{N-1}, Y_{N-1}). \quad (8.21)$$

By (8.19), the contribution of $X_{N-1} = Y_{N-1}$ to the right-hand side has the form

$$-\Lambda(x_n, Y_{N-1}) \sum_{k=1}^{N-1} \sum_{z_k \in E_k: z_k \neq y_k} Q_k(y_k, z_k) \frac{\Lambda_{k-1}^k(z_k, y_{k-1})}{\Lambda_{k-1}^k(z_k, y_{k-1})}. \quad (8.22)$$

For $X_{N-1} \neq Y_{N-1}$, the contribution of matrix elements of $Q^{(N-1)}(X_{N-1}, Y_{N-1})$ that correspond to jumps $(x_k, x_{k+1}, \dots, x_l) \rightarrow (y_k, y_{k+1}, \dots, y_l)$, $1 \leq k \leq l \leq N$, with all other $x_j = y_j$, has the form

$$\begin{aligned} & \sum \Lambda_{N-1}^N(x_N, y_{N-1}) \Lambda_{N-2}^{N-1}(y_{N-1}, y_{N-2}) \cdots \Lambda_{l+1}^{l+2}(y_{l+2}, y_{l+1}) \\ & \times \Lambda_l^{l+1}(y_{l+1}, x_l) \Lambda_{l-1}^l(x_l, x_{l-1}) \cdots \Lambda_k^{k+1}(x_{k+1}, x_k) \\ & \times \Lambda_{k-1}^k(x_k, y_{k-1}) \Lambda_{k-2}^{k-1}(y_{k-1}, y_{k-2}) \cdots \Lambda_1^2(y_2, y_1) \\ & \times Q_k(x_k, y_k) \frac{\Lambda_{k-1}^k(y_k, y_{k-1})}{\Lambda_{k-1}^k(x_k, y_{k-1})} \hat{Q}_{k+1}(x_{k+1}, y_{k+1}, y_k) \cdots \hat{Q}_l(x_l, y_l, y_{l-1}), \end{aligned} \quad (8.23)$$

where the summation is over x_k, \dots, x_l satisfying $x_i \neq y_i$ for all $k \leq i \leq l$ and

$$\Lambda_{k-1}^k(x_k, y_{k-1}) \neq 0, \quad \Lambda_{i-1}^i(x_i, y_{i-1}) = 0, \quad k < i \leq l. \quad (8.24)$$

Denote this expression by $A(k, l)$.

Observe that in (8.23), the factors $\Lambda_{k-1}^k(x_k, y_{k-1})$ cancel out. Let us denote by $B(k, l)$ the sum of same expressions (8.23) with canceled $\Lambda_{k-1}^k(x_k, y_{k-1})$, and with conditions (8.24) replaced by

$$\Lambda_{k-1}^k(x_k, y_{k-1}) = 0, \quad \Lambda_{i-1}^i(x_i, y_{i-1}) = 0, \quad k < i \leq l.$$

Thus, the sum $A(k, l) + B(k, l)$ has no restrictions on x_k other than $x_k \neq y_k$.

Using the definitions of Δ_k^{k+1} and \hat{Q}_{k+1} we see that

$$\begin{aligned} & \sum_{x_k: x_k \neq y_k} \Lambda_k^{k+1}(x_{k+1}, x_k) Q_k(x_k, y_k) \hat{Q}_{k+1}(x_{k+1}, y_{k+1}, y_k) \\ & = Q_{k+1}(x_{k+1}, y_{k+1}) \Lambda_k^{k+1}(y_{k+1}, y_k). \end{aligned} \quad (8.25)$$

Hence, $A(k, l) + B(k, l) = B(k+1, l)$. Noting that $B(1, l) = 0$, we obtain, for any $l = 1, \dots, N-1$,

$$\begin{aligned} & A(1, l) + A(2, l) + \cdots + A(l, l) \\ & = A(l, l) + B(l, l) \\ & = \frac{\Lambda(x_N, Y_{N-1})}{\Lambda_l^{l+1}(y_{l+1}, y_l)} \sum_{x_l: x_l \neq y_l} \Lambda_l^{l+1}(y_{l+1}, x_l) Q_l(x_l, y_l) \end{aligned}$$

$$\begin{aligned}
&= \frac{\Lambda(x_N, Y_{N-1})}{\Lambda_l^{l+1}(y_{l+1}, y_l)} \sum_{x_l \in E_l} \Lambda_l^{l+1}(y_{l+1}, x_l) Q_l(x_l, y_l) - \Lambda(x_N, Y_{N-1}) Q_l(y_l, y_l) \\
&= \Lambda(x_N, Y_{N-1}) \left(\sum_{z_{l+1} \in E_{l+1}} Q_{l+1}(y_{l+1}, z_{l+1}) \frac{\Lambda_l^{l+1}(z_{l+1}, y_l)}{\Lambda_l^{l+1}(y_{l+1}, y_l)} - Q_l(y_l, y_l) \right) \\
&= \Lambda(x_N, Y_{N-1}) \left(\sum_{z_{l+1} \neq y_{l+1}} Q_{l+1}(y_{l+1}, z_{l+1}) \frac{\Lambda_l^{l+1}(z_{l+1}, y_l)}{\Lambda_l^{l+1}(y_{l+1}, y_l)} \right. \\
&\quad \left. + Q_{l+1}(y_{l+1}, y_{l+1}) - Q_l(y_l, y_l) \right), \tag{8.26}
\end{aligned}$$

where we used the commutativity relation $\Lambda_l^{l+1} Q_l = Q_{l+1} \Lambda_l^{l+1}$ along the way. Hence, using (8.21) we obtain

$$\begin{aligned}
\Lambda Q^{(N-1)}(x_N, Y_{N-1}) &= \sum_{1 \leq k \leq l \leq N-1} A(k, l) + \Lambda(x_N, Y_{N-1}) Q^{(N-1)}(Y_{N-1}, Y_{N-1}) \\
&= \sum_{z_N \neq y_N} Q_N(x_N, z_N) \Lambda(z_N, Y_{N-1}) + Q_N(x_N, x_N) \Lambda(x_N, Y_{N-1}) \\
&= Q_N \Lambda(x_N, Y_{N-1}). \quad \square \tag{8.27}
\end{aligned}$$

For any $N \geq 2$ and $x_{N-1} \in E_{N-1}$ let us define a matrix $Q_{x_{N-1}}$ of transition rates on the fiber

$$E_{x_{N-1}} = \{x_N \in E_N \mid \Lambda_{N-1}^N(x_N, x_{N-1}) \neq 0\}$$

via

$$\begin{aligned}
Q_{x_{N-1}}(x_N, y_N) &= Q_N(x_N, y_N) \frac{\Lambda_{N-1}^N(y_N, x_{N-1})}{\Lambda_{N-1}^N(x_N, x_{N-1})}, \quad y_N \neq x_N, \\
Q_{x_{N-1}}(x_N, x_N) &= - \sum_{y_N \in E_{x_{N-1}}, y_N \neq x_N} Q_{x_{N-1}}(x_N, y_N). \tag{8.28}
\end{aligned}$$

The next statement is analogous to Proposition 2.5 in [6].

Proposition 8.6. Assume that the matrices of transition rates Q_1, \dots, Q_N and $Q_{x_1}, \dots, Q_{x_{N-1}}$ for any $x_j \in E_j$, $j = 1, \dots, N-1$ are regular. Then $Q^{(2)}, \dots, Q^{(N)}$ are also regular. Denote by $\{P_j(t)\}_{1 \leq j \leq N}$ and $P^{(N)}(t)$ the transition matrices for $\{Q_j(t)\}_{1 \leq j \leq N}$ and $Q^{(N)}(t)$.

Let μ_N be a probability measure on E_N , and for $t \geq 0$, let $(x_1(t), \dots, x_N(t))$ be an $E^{(N)}$ -valued random variable with

$$\begin{aligned}
&\text{Prob}\{(x_1(t), \dots, x_N(t)) = (x_1, \dots, x_N)\} \\
&= \sum_{Y_N \in E^{(N)}} \mu_N(y_N) \Lambda(y_N, Y_{N-1}) P^{(N)}(t; Y_N, X_N). \tag{8.29}
\end{aligned}$$

Then for any sequence of time moments

$$\begin{aligned} 0 \leq t_0^N \leq t_1^N \leq \dots \leq t_{k_N}^N = t_0^{N-1} \leq t_1^{N-1} \leq \dots \leq t_{k_{N-1}}^{N-1} = t_0^{N-2} \leq \dots \\ \leq t_{k_2}^2 = t_0^1 \leq t_1^1 \leq \dots \leq t_{k_1}^1 \end{aligned} \quad (8.30)$$

the joint distribution of $\{x_m(t_k^m)\}$ ordered as the time moments coincides with the stochastic evolution of μ_N under transition matrices

$$\begin{aligned} P_N(t_0^N), P_N(t_1^N - t_0^N), \dots, P_N(t_{k_N}^N - t_{k_{N-1}}^N), \Lambda_{N-1}^N, P_{N-1}(t_1^{N-1} - t_0^{N-1}), \dots, \\ P_{N-1}(t_{k_{N-1}}^{N-1} - t_{k_{N-2}}^{N-1}), \Lambda_{N-2}^{N-1}, \dots, P_1(t_1^1 - t_0^1), \dots, P_1(t_{k_1}^1 - t_{k_1-1}^1). \end{aligned} \quad (8.31)$$

Proof. It is a straightforward computation to see that the construction of the bivariate Markov chain from the previous section applied to $Q = Q^{(N-1)}$, $Q^* = Q_N$, and Λ given by (8.20) (the needed commutativity is proved in Proposition 8.5), yields exactly $Q^{(N)}$. We apply Corollary 8.4, and induction on N concludes the proof. \square

Corollary 8.7. In the assumptions of Proposition 8.6, $(P^{(N)}(t))_{t \geq 0}$ is central in the sense of Definition 8.1, and the induced semigroup on E_N is exactly $(P_N(t))_{t \geq 0}$. Furthermore, compatibility relations of Proposition 8.2 also hold.

Proof. The first two statements follow from Proposition 8.6 with

$$k_N = 1, \quad k_{N-1} = k_{N-2} = \dots = k_1 = 0.$$

The third statement is (8.7) with $Q = Q^{(N-1)}$, $Q^* = Q_N$, $Q^{(2)} = Q^{(N)}$, and Λ given by (8.20). \square

9. Stochastic dynamics on paths. Gelfand–Tsetlin graph

9.1. Central measures on paths and the boundary

Let us return to our concrete setup, cf. Section 3. We have $E_N = \mathbb{GT}_N$, the space of signatures of length N , and $E^{(N)}$ of (8.1) is the set of Gelfand–Tsetlin schemes of length N ; we denote it by $\mathbb{GT}^{(N)}$.

Due to (3.1), the notion of centrality for $\mu^{(N)} \in \mathcal{M}_p(\mathbb{GT}^{(N)})$ means the following, cf. (8.2): For any $\underline{\lambda} = (\lambda^{(1)} < \lambda^{(2)} < \dots < \lambda^{(N)}) \in \mathbb{GT}^{(N)}$, $\mu^{(N)}(\underline{\lambda})$ depends only on $\lambda^{(N)}$. For branching graphs, the notion of central measures was introduced in [51], see also [31].

In Section 8.1 we explained that central measures on the space $E^{(\infty)} =: \mathbb{GT}^{(\infty)}$ of infinite Gelfand–Tsetlin schemes are in bijection, thanks to Theorem 3.1, with $\mathcal{M}_p(\mathcal{Q})$. Let us make this bijection more explicit.

Given a signature $\lambda \in \mathbb{GT}_N$, denote by λ^+ and λ^- its positive and negative parts. These are two partitions (or Young diagrams) with $\ell(\lambda^+) + \ell(\lambda^-) \leq N$, where $\ell(\cdot)$ is the number of nonzero rows of a Young diagram. In other words,

$$\lambda = (\lambda_1^+, \dots, \lambda_k^+, 0, \dots, 0, -\lambda_l^-, \dots, -\lambda_1^-), \quad k = \ell(\lambda^+), \quad l = \ell(\lambda^-).$$

Given a Young diagram ν , denote by $d(\nu)$ the number of diagonal boxes in ν . Introduce *Frobenius coordinates* of ν via

$$p_i(\nu) = \nu_i - i, \quad q_i(\nu) = \nu'_i - i, \quad i = 1, \dots, d(\nu),$$

where ν' stands for the transposed diagram. We also set

$$p_i(\nu) = q_i(\nu) = 0, \quad i > d(\nu).$$

An element $\underline{\lambda} = (\lambda^{(1)} \prec \lambda^{(2)} \prec \dots) \in \mathbb{GT}^{(\infty)}$, which can be viewed as an infinite increasing path in the Gelfand–Tsetlin graph \mathbb{GT} , is called *regular* if there exist limits

$$\alpha_i^\pm = \lim_{N \rightarrow \infty} \frac{p_i(\lambda^{(N)})}{N}, \quad \beta_i^\pm = \lim_{N \rightarrow \infty} \frac{q_i(\lambda^{(N)})}{N}, \quad i = 1, 2, \dots, \quad \delta^\pm = \lim_{N \rightarrow \infty} \frac{|\lambda^\pm|}{N}.$$

The corresponding point $\omega = (\alpha^\pm, \beta^\pm, \delta^\pm) \in \Omega$ is called the *end* of this path.

Theorem 9.1. (See [44].) Any central measure on $\mathbb{GT}^{(\infty)}$ is supported by the Borel set of regular paths. Push-forward of such measures under the map that takes a regular path to its end, establishes an isomorphism between the space of central measures on $\mathbb{GT}^{(\infty)}$ and $\mathcal{M}_p(\Omega)$.

We refer the reader to Section 10 of [44] for details.

9.2. Matrices of transition rates on $\mathbb{GT}^{(N)}$

With $E_N = \mathbb{GT}_N$, $E^{(N)} = \mathbb{GT}^{(N)}$ and $Q_N = \mathcal{D}^{(N)}$, let us write out the specialization of the matrix $Q^{(N)}$ from Section 8.2. We will use the notation $\mathbf{D}^{(N)}$ for the resulting matrix of transition rates on $\mathbb{GT}^{(N)}$. As for the parameters, we will use (6.1) and (6.2) as before.

To any $\underline{\lambda} \in \mathbb{GT}^{(N)}$ we associate an array $\{l_i^j \mid 1 \leq i \leq j, 1 \leq j \leq N\}$ using $l_i^j = \lambda_i^{(j)} + j - i$. In these coordinates, the interlacing conditions $\lambda^{(j)} \prec \lambda^{(j+1)}$ take the form

$$l_i^{j+1} > l_i^j \geq l_{i+1}^{j+1}$$

for all meaningful values of i and j .¹

Similarly, assign $\mathbb{GT}^{(N)} \ni \underline{\nu} \leftrightarrow \{n_i^j = \nu_i^{(j)} + j - i\}_{1 \leq i \leq j, 1 \leq j \leq N}$. Gathering all the definitions together, we obtain that the off-diagonal entries of $\mathbf{D}^{(N)}$ have the form

$$\mathbf{D}^{(N)}(\underline{\lambda}, \underline{\nu}) = \begin{cases} (l_i^k - z - k + 1)(l_i^k - z' - k + 1), \\ (l_i^k + w)(l_i^k + w'), \end{cases}$$

where for the first line we must have i, k and l , $1 \leq i \leq k \leq l \leq N$, such that

¹ One could make the interlacing condition more symmetric (both inequalities being strict) by considering the coordinates $\tilde{l}_i^j = \lambda_i^{(j)} + (j+1)/2 - i$ instead. This would imply however that $\tilde{l}_i^j \in \mathbb{Z} + 1/2$ for even j while $\tilde{l}_i^j \in \mathbb{Z}$ for odd j .

$$l_i^j = l_i^k + j - k, \quad n_i^j = l_i^j + 1 \quad \text{for all } k \leq j \leq l,$$

and all other coordinates of $\underline{\lambda}$ and $\underline{\nu}$ are equal, while for the second line we must have i, k, l with $1 \leq i \leq k \leq l \leq N$ such that

$$l_{i+j-k}^j = l_i^k, \quad n_{i+j-k}^j = l_{i+j-k}^j - 1 \quad \text{for all } k \leq j \leq l,$$

and all other coordinates of $\underline{\lambda}$ and $\underline{\nu}$ are equal.

The Markov chain generated by $\mathbf{D}^{(N)}$ can be described as follows:

(1) Each coordinate l_i^k tries to jump to the right by 1 with rate $(l_i^k - z - k + 1)(l_i^k - z' - k + 1)$ and to the left by 1 with rate $(l_i^k + w)(l_i^k + w')$, independently of other coordinates.

(2) If the l_i^k -clock of the right jump rings but $l_i^k = l_{i-1}^{k-1}$, the jump is blocked. If its left clock rings but $l_i^k = l_i^{k-1} + 1$, the jump is also blocked. (If any of the two jumps were allowed then the resulting set of coordinates would not have corresponded to an element of $\mathbb{GT}^{(N)}$ as the interlacing conditions would have been violated.)

(3) If the right l_i^k -clock rings and there is no blocking, we find the greatest number $l \geq k$ such that $l_i^j = l_i^k + j - k$ for $j = k, k + 1, \dots, l$, and move all the coordinates $\{l_i^j\}_{j=k}^l$ to the right by one. Given the change $l_i^k \mapsto l_i^k + 1$, this is the minimal modification of the set of coordinates that preserves interlacing.

(4) If the left l_i^k -clock rings and there is no blocking, we find the greatest number $l \geq k$ such that $l_{i+j-k}^j = l_i^k$ for $j = k, k + 1, \dots, l$, and move all the coordinates $\{l_{i+j-k}^j\}_{j=k}^l$ to the left by one. Again, given the change $l_i^k \mapsto l_i^k - 1$, this is the minimal modification of the set of coordinates that preserves interlacing.

Certain Markov chain on interlacing arrays with a similar block-push mechanism have been studied in [6], see also [10]. In those examples the jump rates are constant though.

9.3. Regularity

In order to claim the benefits of Proposition 8.6 and Corollary 8.7, we need to verify the regularity of the fiber matrices of transition rates (8.28). In our concrete realization, they take the following form.

For any $N \geq 2$ and any $\kappa \in E_{N-1} = \mathbb{GT}_{N-1}$, the fiber $E_\kappa =: \mathbb{GT}_\kappa \subset \mathbb{GT}_N$ takes the form

$$\mathbb{GT}_\kappa = \{\lambda \in \mathbb{GT}_N \mid \kappa < \lambda\}.$$

Using the coordinates $\{l_i = N + \lambda_i - i\}_{i=1}^N$ for $\lambda \in \mathbb{GT}_N$ and $\{n_i = N + \nu_i - i\}_{i=1}^N$ for $\nu \in \mathbb{GT}_N$, the off-diagonal part of the matrix of transition rates $\mathcal{D}_\kappa := Q_\kappa$ on the fiber \mathbb{GT}_κ has the form

$$\mathcal{D}_\kappa(\lambda, \nu) = \begin{cases} (l_i - z - N + 1)(l_i - z' - N + 1), \\ (l_i + w)(l_i + w'), \end{cases}$$

where for the first line we must have $i, 1 \leq i \leq N$, such that

$$n_i = l_i + 1, \quad n_j = l_j \quad \text{for } j \neq i,$$

and for the second line we must have

$$n_i = l_i - 1, \quad n_j = l_j \quad \text{for } j \neq i.$$

Proposition 9.2. *For any $N \geq 2$ and any $\kappa \in \mathbb{GT}_{N-1}$, the matrix of transition rates \mathcal{D}_κ on \mathbb{GT}_κ is regular.*

Proof. The interlacing condition in the definition of \mathbb{GT}_κ implies that \mathcal{D}_κ is the matrix of transition rates for N independent birth and death processes conditioned to stay within N non-overlapping intervals inside \mathbb{Z} ; one interval per process. The results of Section 3.2 show that any such birth and death process is regular as such a process either lives on a finite set or it is a one-sided birth and death process of the type considered in the proof of Theorem 5.1. \square

Corollary 9.3. *For any $N \geq 1$, the matrix $\mathbf{D}^{(N)}$ of transition rates on $\mathbb{GT}^{(N)}$ is regular, and the corresponding semigroup $(P^{(N)}(t))_{t \geq 0}$ is central. The induced Markov semigroup on \mathbb{GT}_N coincides with that of Section 5.*

Proof. Follows from Proposition 8.6 and Corollary 8.7. \square

9.4. Exclusion process

Observe that the projection of the Markov chain generated by $\mathbf{D}^{(N)}$ to the coordinate l_1^1 is a bilateral birth and death process. Furthermore, the jumps of l_1^2 are only influenced by l_1^1 , the jumps of l_1^3 are only influenced by l_1^1 and l_1^2 , and so on. On the other side, the jumps of l_k^k are only influenced by $\{l_1^1, l_2^2, \dots, l_{k-1}^{k-1}\}$ for any $k \geq 2$.

Hence, the projection of the Markov chain defined by $\mathbf{D}^{(N)}$ to the coordinates $(l_N^N \leq l_{N-1}^{N-1} \leq \dots \leq l_1^1 < l_1^2 < \dots < l_1^N)$ is also a Markov chain.² The fibers of this projection are finite, hence, according to Proposition 4.4, our Markov chain on $\mathbb{GT}^{(N)}$ collapses to the smaller one, whose matrix of transition rates is also regular.

Let us project even further to $(l_1^1 < l_1^2 < \dots < l_1^N)$. Killing extra coordinates one-by-one and using the results of Section 5.1 to verify the regularity for the fiber chains, we see that the collapsibility of Proposition 4.4 holds. Let us give an independent description of the resulting Markov chain on $\{l_1^j\}_{j \geq 1}$.

Set

$$\begin{aligned} \mathfrak{Y}_N &= \{y_1 < y_2 < \dots < y_N \mid y_j \in \mathbb{Z}, 1 \leq j \leq N\}, \\ \mathfrak{Y}_\infty &= \{y_1 < y_2 < \dots \mid y_j \in \mathbb{Z}, j \geq 1\}. \end{aligned}$$

Define the matrix $\mathbf{D}_{top}^{(N)}$ of transition rates on \mathfrak{Y}_N by

$$\mathbf{D}^{(N)}(Y', Y'') = \begin{cases} (y_k - z - k + 1)(y_k - z' - k + 1), \\ (y_k + w)(y_k + w'), \end{cases}$$

² Once again, all the inequalities would be strict if we considered coordinates $\tilde{l}_i^j = \lambda_i^j + (j+1)/2 - i$.

where for the first line we must have k and l , $1 \leq k \leq l \leq N$, such that

$$y'_j = y'_k + j - k, \quad y''_j = y'_j + 1 \quad \text{for all } k \leq j \leq l,$$

and all other coordinates of Y' and Y'' are equal, while for the second line we must have

$$y''_k = y'_k - 1, \quad y''_m = y'_m, \quad \text{for } m \neq k.$$

In other words, each coordinate y_k tries to jump to the right by 1 with rate $(y_k - z - k + 1)(y_k - z' - k + 1)$, and it tries to jump to the left by 1 with rate $(y_k + w)(y_k + w')$, independently of other coordinates. If the left y_k -clock rings but $y_k = y_{k-1} + 1$ then the jump is blocked. If the right y_k -clock rings we find the greatest number $l \geq k$ such that $y_j = y_k + j - k$ for $j = k, k+1, \dots, l$, and move all the coordinates $\{y_k, \dots, y_l\}$ to the right by one. One could think of y_k “pushing” y_{k+1}, \dots, y_l . Alternatively, if one forgets about the labeling one could think of y_k jumping to the first available site on its right.

Clearly, these Markov chains are compatible with projections $\mathfrak{Y}_{N+1} \rightarrow \mathfrak{Y}_N$ that remove the last coordinate. Thus, we obtain a Markov semigroup on $\varprojlim \mathfrak{Y}_N = \mathfrak{Y}_\infty$.

This semigroup is a sort of an exclusion process — it is a one-dimensional interacting particle system with each site occupied by no more than one particle (exclusion constraint). A similar system, but with constant jump rates, was considered in [5] and called PushASEP. A system with one-sided jumps and blocking mechanism as above is usually referred to as *Totally Asymmetric Simple Exclusion Process* (TASEP), while a system with one-sided jumps and pushing mechanism as above is sometimes called *long range TASEP*. See [49,36] for more information on exclusion processes.

Proposition 9.4. *The exclusion process defined above has a unique invariant probability measure. With probability 1 with respect to this measure there exists a limit $r = \lim_{N \rightarrow \infty} y_N/N$, which is a random variable with values ≥ 1 . Under certain additional restrictions on parameters (z, z', w, w') , see below, the function*

$$\sigma(s) = s(s-1) \frac{d}{ds} \text{Prob}\{r \leq s\} - a_1^2 s + \frac{1}{2}(a_3 a_4 + a_1^2)$$

is the unique solution of the (second order nonlinear) differential equation

$$\begin{aligned} -\sigma'(s(s-1)\sigma'') &= \left(2\left(\left(s - \frac{1}{2}\right)\sigma' - \sigma\right)\sigma' - a_1 a_2 a_3 a_4\right)^2 \\ &\quad - (\sigma' + a_1^2)(\sigma' + a_2^2)(\sigma' + a_3^2)(\sigma' + a_4^2) \end{aligned} \quad (9.1)$$

with boundary condition

$$\sigma(s) = -a_1^2 s + \frac{1}{2}(a_3 a_4 + a_1^2) + \frac{\sin \pi z \sin \pi z'}{\pi^2} s^{-2a_1} + o(s^{-2a_1}), \quad s \rightarrow +\infty,$$

where the constants a_1, a_2, a_3, a_4 are given by

$$a_1 = a_2 = \frac{z + z' + w + w'}{2}, \quad a_3 = \frac{z - z' + w - w'}{2}, \quad a_4 = \frac{z - z' - w + w'}{2}.$$

Remarks. 1. The quantity $(\lim_{N \rightarrow \infty} y_N/N)^{-1}$ can be viewed as the asymptotic density of the system of particles (y_j) at infinity. Proposition 9.4 claims that for the invariant measure, this quantity is well-defined and random.

2. The restrictions on parameters come from Theorem 7.1 of [4]. They can be relaxed, see Remark 7.2 in [4] and the end of §3 in [37].

3. The differential equation above is the so-called σ -form of the Painlevé VI equation first appeared in [25].

Proof of Proposition 9.4. The invariant measure is simply the projection to

$$y_1 = \lambda^{(1)}, \quad y_2 = \lambda_1^{(2)} + 1, \quad y_3 = \lambda_1^{(3)} + 2, \quad \dots$$

of the central measure on $\mathbb{GT}^{(\infty)}$ corresponding to the spectral zw -measure. The uniqueness follows from the uniqueness of invariant measure on countable sets \mathfrak{Y}_N , cf. Theorem 1.6 of [1] (a similar argument was used in the proof of Theorem 7.2). The existence of $\lim_{N \rightarrow \infty} y_N/N$ follows from Theorem 9.1. Finally, the characterization of the distribution of this limit in terms of the Painlevé VI equation was proved in Theorem 7.1 of [4], see [37] for another proof. \square

Acknowledgments

A.B. was partially supported by NSF grants DMS-0707163 and DMS-1056390. G.O. was partially supported by the RFBR grant 08-01-00110, the RFBR-CNRS grant 10-01-93114, the project SFB 701 of Bielefeld University, and a grant from Simons Foundation (Simons–IUM Fellowship).

References

- [1] W.J. Anderson, Continuous Time Markov Chains: An Applications-Oriented Approach, Springer, 1991.
- [2] R. Askey, An integral of Ramanujan and orthogonal polynomials, J. Indian Math. Soc. 51 (1987) 27–36.
- [3] A. Borodin, Schur dynamics of the Schur processes, Adv. Math. 228 (2011) 2268–2291, arXiv:1001.3442.
- [4] A. Borodin, P. Deift, Fredholm determinants, Jimbo–Miwa–Ueno τ -functions, and representation theory, Comm. Pure Appl. Math. 55 (2002) 1160–1230, arXiv:math/0111007.
- [5] A. Borodin, P.L. Ferrari, Large time asymptotics of growth models on space-like paths I: PushASEP, Electron. J. Probab. 13 (2008) 1380–1418, arXiv:0707.2813.
- [6] A. Borodin, P.L. Ferrari, Anisotropic growth of random surfaces in $2 + 1$ dimensions, arXiv:0804.3035.
- [7] A. Borodin, V. Gorin, Shuffling algorithm for boxed plane partitions, Adv. Math. 220 (2009) 1739–1770, arXiv:0804.3071.
- [8] A. Borodin, V. Gorin, Markov processes of infinitely many nonintersecting random walks, preprint, arXiv:1106.1299.
- [9] A. Borodin, V. Gorin, E.M. Rains, q -Distributions on boxed plane partitions, Selecta Math. 16 (2010) 731–789, arXiv:0905.0679.
- [10] A. Borodin, J. Kuan, Random surface growth with a wall and Plancherel measures for $O(\infty)$, Comm. Pure Appl. Math. 63 (2010) 831–894, arXiv:0904.2607.
- [11] A. Borodin, G. Olshanski, Harmonic analysis on the infinite-dimensional unitary group and determinantal point processes, Ann. of Math. 161 (2005) 1–104, arXiv:math/0109194.
- [12] A. Borodin, G. Olshanski, Representation theory and random point processes, in: European Congress of Mathematics, Eur. Math. Soc., Zurich, 2005, pp. 73–94, arXiv:math/0409333.
- [13] A. Borodin, G. Olshanski, Random partitions and the Gamma kernel, Adv. Math. 194 (2005) 141–202, arXiv:math-ph/0305043.
- [14] A. Borodin, G. Olshanski, Markov processes on partitions, Probab. Theory Related Fields 135 (2006) 84–152, arXiv:math-ph/0409075.

- [15] A. Borodin, G. Olshanski, Stochastic dynamics related to Plancherel measure on partitions, in: V. Kaimanovich, A. Lodkin (Eds.), *Representation Theory, Dynamical Systems, and Asymptotic Combinatorics*, in: Amer. Math. Soc. Transl. Ser. 2: Adv. Math. Sci., vol. 217, 2006, pp. 9–21, arXiv:math-ph/0402064.
- [16] A. Borodin, G. Olshanski, Infinite-dimensional diffusions as limits of random walks on partitions, *Probab. Theory Related Fields* 144 (2009) 281–318, arXiv:0706.1034.
- [17] A. Borodin, G. Olshanski, The boundary of the Gelfand–Tsetlin graph: A new approach, preprint, arXiv:1109.1412, 2011.
- [18] P. Diaconis, J.A. Fill, Strong stationary times via a new form of duality, *Ann. Probab.* 18 (1990) 1483–1522.
- [19] S.N. Ethier, T.G. Kurtz, *Markov Processes — Characterization and Convergence*, Wiley–Interscience, New York, 1986.
- [20] W. Feller, On the integro-differential equations of purely discontinuous Markoff processes, *Trans. Amer. Math. Soc.* 48 (1940) 488–815; Errata: *Trans. Amer. Math. Soc.* 58 (1945) 474.
- [21] W. Feller, On boundaries and lateral conditions for the Kolmogorov differential equations, *Ann. of Math.* 65 (1957) 527–570.
- [22] W. Feller, The birth and death processes as diffusion processes, *J. Math. Pures Appl.* 38 (1959) 301–345.
- [23] R. Goodman, N.R. Wallach, *Symmetry, Representations, and Invariants*, Grad. Texts in Math., vol. 255, Springer, 2009.
- [24] K. Itô, *Essentials of Stochastic Processes*, Transl. Math. Monogr., vol. 231, Amer. Math. Soc., Providence, RI, 2006, translated from the 1957 Japanese original.
- [25] M. Jimbo, T. Miwa, Monodromy preserving deformation of linear ordinary differential equations with rational coefficients. II, *Phys. D* 2 (3) (1981) 407–448.
- [26] S. Karlin, J. McGregor, Coincidence probabilities, *Pacific J. Math.* 9 (1959) 1141–1164.
- [27] T. Kato, *Perturbation Theory of Linear Operators*, 2nd edition, Springer-Verlag, New York, 1980.
- [28] M. Katori, H. Tanemura, Zeros of Airy function and relaxation process, *J. Stat. Phys.* 136 (2009) 1177–1204, arXiv:0906.3666.
- [29] M. Katori, H. Tanemura, Non-equilibrium dynamics of Dyson’s model with an infinite number of particles, *Comm. Math. Phys.* 293 (2010) 469–497, arXiv:0812.4108.
- [30] F.P. Kelly, Invariant measures and the Q-matrix, in: *Probability, Statistics and Analysis*, in: London Math. Soc. Lecture Note Ser., vol. 79, Cambridge Univ. Press, Cambridge–New York, 1983, pp. 143–160s.
- [31] S.V. Kerov, *Asymptotic Representation Theory of the Symmetric Group and Its Applications in Analysis*, Amer. Math. Soc., Providence, RI, 2003.
- [32] R. Koekoek, P.A. Lesky, R.F. Swarttouw, *Hypergeometric Orthogonal Polynomials and Their q -Analogues*, Springer, 2010.
- [33] W. König, Orthogonal polynomial ensembles in probability theory, *Probab. Surv.* 2 (2005) 385–447.
- [34] P.A. Lesky, Unendliche und endliche Orthogonalsysteme von continuous Hahnpolynomen, *Results Math.* 31 (1997) 127–135.
- [35] P.A. Lesky, Eine Charakterisierung der kontinuierlichen und diskreten klassischen Orthogonalpolynome, *Mathematisches Institut A, Universität Stuttgart*, preprint 98–12.
- [36] T. Liggett, *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*, Grundlehren Math. Wiss., vol. 324, Springer-Verlag, Berlin, 1999.
- [37] O. Lisovyy, Dyson’s constant for the hypergeometric kernel, arXiv:0910.1914.
- [38] I.G. Macdonald, *Symmetric Functions and Hall Polynomials*, 2nd edition, Oxford Univ. Press, 1995.
- [39] G.W. Mackey, Borel structures on groups and their duals, *Trans. Amer. Math. Soc.* 85 (1957) 134–165.
- [40] P.-A. Meyer, *Probability and Potentials*, Blaisdell, 1966.
- [41] E. Nordenstam, On the shuffling algorithm for domino tilings, *Electron. J. Probab.* bf15 (3) (2010) 75–95, arXiv:0802.2592.
- [42] A. Okounkov, G. Olshanski, Shifted Schur functions, *Algebra i Analiz* 9 (2) (1997) 73–146 (in Russian); English translation: *St. Petersburg Math. J.* 9 (2) (1998) 239–300, arXiv:q-alg/9605042.
- [43] A. Okounkov, G. Olshanski, Asymptotics of Jack polynomials as the number of variables goes to infinity, *Int. Math. Res. Not. IMRN* 1998 (13) (1998) 641–682, arXiv:q-alg/9709011.
- [44] G. Olshanski, The problem of harmonic analysis on the infinite-dimensional unitary group, *J. Funct. Anal.* 205 (2003) 464–524, arXiv:math/0109193.
- [45] G. Olshanski, Anisotropic Young diagrams and infinite-dimensional diffusion processes with the Jack parameter, *Int. Math. Res. Not. IMRN* 2010 (6) (2010) 1102–1166, arXiv:0902.3395.
- [46] G. Olshanski, Laguerre and Meixner symmetric functions, and infinite-dimensional diffusion processes, *J. Math. Sci. (N. Y.)* 174 (1) (2011) 41–57, arXiv:1009.2037.

- [47] H. Osada, Interacting Brownian motions in infinite dimensions with logarithmic interaction potentials, arXiv: 0902.3561.
- [48] W.E. Pruitt, Bilateral birth and death processes, *Trans. Amer. Math. Soc.* 107 (1963) 508–525.
- [49] F. Spitzer, Interaction of Markov processes, *Adv. Math.* 5 (1970) 246–290.
- [50] H. Spohn, Interacting Brownian particles: A study of Dyson's model, in: G. Papanicolaou (Ed.), *Hydrodynamic Behavior and Interacting Particle Systems*, in: IMA Vol. Math. Appl., vol. 9, Springer-Verlag, Berlin, 1987, pp. 151–179.
- [51] A.M. Vershik, S.V. Kerov, Asymptotic theory of characters of the symmetric group, *Funct. Anal. Appl.* 15 (1981) 246–255.
- [52] A.M. Vershik, S.V. Kerov, Characters and factor-representations of the infinite unitary group, *Dokl. Akad. Nauk SSSR* 267 (2) (1982) 272–276 (in Russian); English translation: *Sov. Math. Dokl.* 26 (3) (1982) 570–574 (1983).
- [53] D. Voiculescu, Représentations factorielles de type II_1 de $U(\infty)$, *J. Math. Pures Appl.* 55 (1976) 1–20.
- [54] J. Warren, Dyson's Brownian motions, intertwining and interlacing, *Electron. J. Probab.* 12 (2007) 573–590, arXiv:math/0509720.
- [55] H. Weyl, *The Classical Groups. Their Invariants and Representations*, Princeton Univ. Press, 1939 (fifth edition, 1997).
- [56] Xiangqun Yang, *The Construction Theory of Denumerable Markov Processes*, Hunan Science and Technology Publ. House, 1990 (Wiley series in probability and mathematical statistics: Probability and mathematical statistics).
- [57] D.P. Zhelobenko, *Compact Lie Groups and Their Representations*, Nauka, Moscow, 1970 (in Russian); English translation *Transl. Math. Monogr.*, vol. 40, Amer. Math. Soc., Providence, RI, 1973.