VIRTUAL MARKOV CHAINS

STEVEN N. EVANS AND ADAM Q. JAFFE

Abstract. We introduce the space of virtual Markov chains (VMCs) as a projective limit of the spaces of all finite state space Markov chains (MCs), in the same way that the space of virtual permutations is the projective limit of the spaces of all permutations of finite sets. We introduce the notions of virtual initial distribution (VID) and a virtual transition matrix (VTM), and we show that the law of any VMC is uniquely characterized by a pair of a VID and VTM which have to satisfy a certain compatibility condition. Lastly, we study various properties of compact convex sets associated to the theory of VMCs, including that the Birkhoff-von Neumann theorem fails in the virtual setting.

Steve (the first author) will never forget showering with Vaughan one day in the Berkeley gym after their weekly game of squash while Vaughan sang an excerpt from the aria "I Know that my Redeemer Liveth" from Handel's "Messiah" in his wonderful bass voice while a group of startled undergraduates looked on. This episode captured Vaughan perfectly: his love of sport as well as the life of the mind, his cultured but frequently profane irreverence, and his tremendous zest for living.

Mathematicians visiting Steve's house would always be impressed when he answered the phone and it was Vaughan on the line—They were less impressed when they found out that Vaughan invariably wanted Steve for his body and not his mind and was arranging a squash match or a game of tennis. Nonetheless, Vaughan and Steve did find a lot of time to discuss mathematics and Steve misses their friendship deeply.

As part of those conversations, Vaughan would sometimes ask Steve what he was working on. Steve would answer with some topic in probability and Vaughan would think for a while. If he was wearing his operator algebras hat that day, he would reply with, "Oh, that's the commutative case," and if he was wearing his statistical mechanics hat, he would reply with, "Oh, that's the high temperature case." Steve wonders what he would say about the following ...

1. Introduction

For each $N \in \mathbb{N} := \{1, 2, \ldots\}$, write $\mathcal{S}_N$ for the group of all permutations of the finite set $\{1, 2, \ldots, N\}$. A *virtual permutation* is a sequence $\boldsymbol{\sigma} = \{\sigma_N\}_{N \in \mathbb{N}} \in \prod_{N \in \mathbb{N}} \mathcal{S}_N$ with the extra projectivity condition that for each $N \in \mathbb{N}$ the cycle structure of σ_N is exactly the cycle structure of σ_{N+1} but with the element $N + 1$

removed. The space of virtual permutations $\mathcal{S}$ was introduced in the landmark paper [**10**] as a method for studying the representation theory of the infinite symmetric group $\mathcal{S}_{<\infty} := \bigcup_{N=1}^{\infty} \mathcal{S}_N$, and subsequent research has focused both on furthering this direction [**11, 13, 15**] and on using similar ideas to study the representation theory of other "large" non-abelian groups [**1, 4**].

There are a few natural ways to enrich this picture by introducing an element of randomness. To motivate this, we return to the case of permutations of a finite set, that is, $\mathcal{S}_N$ for a fixed $N \in \mathbb{N}$. One approach is to study a family of probability measures on $\mathcal{S}_N$, and then to ask any number of questions about the resulting distributions, typically in terms of the asymptotics of various statistics as $N \to \infty$. This study of *random permutations* is an extensive area of active research in probability, combinatorics, and mathematical physics, so we will not attempt the futile task of trying to summarize this research in just a few lines. In the case of virtual permutations, the natural generalization of this program is to define a family of probability measures directly on $\mathcal{S}$ and to ask analogous questions about the resulting distributions; this line of research has been very fruitful, and there are a number of interesting probabilistic questions that have been proposed and answered [**2, 3, 9, 14**].

There is another simple way to introduce an aspect of randomness to this picture. To see what this is, we return again to the case of $\mathcal{S}_N$ for a fixed $N \in \mathbb{N}$; the approach is to view each fixed element of $\mathcal{S}_N$ as a deterministic transition rule for some dynamics on $\{1, 2, \ldots, N\}$, and to then expand our scope to include probabilistic transition rules. This study of random dynamics is, of course, nothing other than the study of Markov chains (MCs) on finite state spaces, which has become its own subfield of modern probability theory. In the case of virtual permutations, the natural generalization of this program is to define a family of probabilistic transition rules on $\{1, 2, \ldots, N\}$ simultaneously for all $N \in \mathbb{N}$ in such a way that the extra projectivity condition is, in a suitable sense, always maintained. In this paper, we the take this second perspective as a starting point and hence initiate the study of the so-called *virtual Markov chains (VMCs)*.

Roughly speaking, a VMC on some probability space is a sequence $\boldsymbol{X} = \{X_N\}_{N \in \mathbb{N}}$ of MCs with the extra condition that almost surely for each $N \in \mathbb{N}$ the sample path X_N is exactly the sample path of X_{N+1} but with all instances of the element $N+1$ removed. In particular, a VMC is nothing more than a certain (rather strong) notion of coupling for a sequence of MCs onto the same underlying probability space. We will later give several classes of examples (Subsection 2.3) which show that VMCs are naturally encoded in the enumeration of the hitting locations of various continuous-time Markov processes.

Subsequently, much of the paper is dedicated to generalizing existing results for MCs to the setting of VMCs, with an emphasis on the elements which have behavior that is different from or more interesting than the analogous elements of the classical case. For example, one of our main results (Theorem 3.22), which generalizes the classical result stating that the law of a MC is uniquely described by a pair of an initial distribution (ID) and a transition matrix (TM), shows that the law of a VMC is uniquely characterized by a pair of a *virtual initial distribution (VID)* and *virtual transition matrix (VTM)*, but that these two data must satisfy an additional *compatibility* condition.
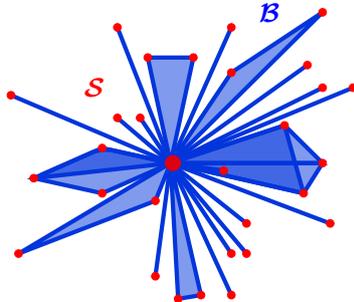
FIGURE 1. A cartoon of the virtual Birkhoff polytope $\mathcal{B}$ (blue) and the virtual permutation matrices $\mathcal{S}$ within it (red).

Another collection of results is related to various compact convex spaces associated to the study of VMCs. In the case of MCs, the TM is usually regarded as containing "more interesting" information than the ID; this intuition extends to the case of VMCs as well, hence we later focus our attention on VTMs and we let VIDs play a secondary role. From this perspective it makes sense to fix a VTM and to consider the space of all VIDs which are compatible with it; it turns out (Subsection 4.1) that this is a compact convex space of measures whose extreme points have many nice properties that we explore. In particular, this study of compatibility is equivalent to characterizing the solution set of an infinite system of balayage inverse problems, which generalize the one-step balayage inverse problems originally studied in [**7, 8**]. We also define the notions of equilibrium distribution and stationary distribution, and show (Subsection 4.2) that they are distinct, unlike in the case of classical MCs where they coincide.

Finally, we give (Subsection 4.4) a collection of results which connects the theory of VMCs back to theory of virtual permutations from whence it came. We know in the classical setting that the space of doubly-stochastic TMs, called the *Birkhoff polytope*, coincides with (closed) convex hull of the permutation matrices; in the virtual setting, we show that the space of doubly-stochastic VTMs, which we analogously call the *virtual Birkhoff polytope* $\mathcal{B}$, is not even convex. However, not all hope is lost in understading the convexity structure of $\mathcal{B}$: We show that it exhibits an "all-or-nothing-convexity" property, namely, that for any two points in $\mathcal{B}$, the relative interior of the line segment joining them is either contained in or disjoint from $\mathcal{B}$. Moreover, we show that there is a unique point in $\mathcal{B}$ with the property that the line segment joining it to any other point of $\mathcal{B}$ is contained in $\mathcal{B}$; this point is none other than the VTM of the virtual permutation corresponding to the identity; in Figure 1 we give a cartoon depiction of the virtual Birkhoff polytope.

The remainder of this paper is structured as follows. In the last part of this section, we outline some important notational conventions, some topological preliminaries, and some remarks about the categorical aspects of the constructions herein. In Section 2 we give the basic definitions and properties of our theory, and we outline several classes of examples of VMCs which can be built from sufficiently regular continuous-time Markov processes. In Section 3 we state and prove the fundamental representation theorem which gives the correspondence between VMCs

and compatible pairs of VIDs and VTMs. In Section 4, we study a few problems of convexity in infinite-dimensional vector spaces that naturally arise in the theory we have developed. Many examples are given in every section.

Notation We use calligraphic characters to denote spaces of objects, and we use boldface characters to denote virtual objects; in LaTeX, these correspond to the `mathcal` and `mathbf` environments, respectively. Correspondingly we use calligraphic boldface characters to denote spaces of virtual objects. For example, $\mathcal{S}$ denotes the space of all permutations of $\mathbb{N}$, a generic element of which is denoted σ; likewise, $\boldsymbol{\mathcal{S}}$ denotes the space of all virtual permutations, a generic element of which is denoted $\boldsymbol{\sigma}$.

We also overload notation and use $\{P_N\}_{N \in \mathbb{N}}$ to denote several different families of projections operations and ι to denote several different natural injections. This overloading should cause no confusion since the argument always disambiguates which map is which.

Topology Unless otherwise stated, all subsets of topological spaces are endowed with the relative topology, all cartesian products of topological spaces are endowed with the product topology, and all spaces of probability measures on a topological space are endowed with the topology of weak convergence.

Category While we will not make the categorical aspects of our work rigorous in this paper, we make a few remarks for the interested reader.

In the case of [**10**], one works in the category of topological groups and considers the sequence $\{\mathcal{S}_N\}_{N \in \mathbb{N}}$ of permutation groups. Then, the infinite symmetric group $S_{<\infty} = \bigcup_{N \in \mathbb{N}} S_N$ is an inductive limit of this sequence when endowed with the natural collection of injective homomorphisms $\iota_{NM} : \mathcal{S}_N \to \mathcal{S}_M$ which, for each $N, M \in \mathbb{N}$ with $N \leq M$, send each permutation on $\{1, 2, \ldots N\}$ to the unique permutation on $\{1, 2, \ldots M\}$ which has the same cycle structure but with each element of $\{N + 1, N + 2, \ldots M - 1, M\}$ placed into its own cycle. Dually, the space of virtual permutations $\boldsymbol{\mathcal{S}}$ is a projective limit of this sequence when endowed with the natural collection of surjective homomorphisms $P_{NM} : \mathcal{S}_M \to \mathcal{S}_N$ which, for each $N, M \in \mathbb{N}$ with $N \leq M$, send each permutation on $\{1, 2, \ldots M\}$ to the permutation on $\{1, 2, \ldots N\}$ which has the same cycle structure but with each element of $\{N + 1, N + 2, \ldots M - 1, M\}$ removed and with the gaps "stitched up".

In the present paper, one works in a certain "category of Markov chains", which, to the best of our knowledge, has not been studied. Roughly speaking, the objects in this category are the Markov chains defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and the morphisms in this category are sample-path transformations that preserve the Markov property. For each $N \in \mathbb{N}$, write $\mathcal{M}_N$ for the collection of all Markov chains on $\{1, 2, \ldots, N\}$ defined on some fixed $(\Omega, \mathcal{F}, \mathbb{P})$, and consider the sequence $\{\mathcal{M}_N\}_{N \in \mathbb{N}}$. Then, one can form an inductive limit of this sequence when it is endowed with the natural collection of injective morphisms $\iota_{NM} : \mathcal{M}_N \to \mathcal{M}_M$ which, for each $N, M \in \mathbb{N}$ with $N \leq M$, send a sample path on $\{1, 2, \ldots N\}$ to the same sample path but viewed as a path on $\{1, 2, \ldots M\}$; the resulting objects is the collection all the finite state space Markov chains on $(\Omega, \mathcal{F}, \mathbb{P})$. Dually, one can form a projective limit when endowing this sequence with the natural collection of surjective morphisms $P_{NM} : \mathcal{M}_M \to \mathcal{M}_N$ which, for each $N, M \in \mathbb{N}$ with $N \leq M$, send each sample path on $\{1, 2, \ldots M\}$ to the sample path on $\{1, 2, \ldots N\}$ which

results from removing all instances of $\{N + 1, N + 2, \ldots M - 1, M\}$ and with the resulting gaps "stitched up". (Note that there is some care to be taken in the case that the sample path eventually leaves $\{1, 2, \ldots, N\}$ and never returns.) The resulting object is the collection of all virtual Markov chains.

2. Basic Theory

In this section we develop the basic theory of the main objects of interest in the paper. In Subsection 2.1 we define the notion of *virtual path space*, and in Subsection 2.2 we define *virtual Markov chains (VMCs)* as nothing other than probability measures on virtual path space with some nice properties. Then in Subsection 2.3 we build various examples of VMCs, mostly from nice continuous-time Markov processes, and we explore some different properties.

2.1. Virtual path space. Set $\mathbb{N} = \{1, 2, \ldots\}$ and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. For $a, b \in \mathbb{N}_0$, write $[\![a, b]\!] = \{a, a + 1, \ldots, b - 1, b\}$, which may be empty. Also write $[\![a, \infty[\![= \{a, a + 1, \ldots\}$ for each $a \in \mathbb{N}_0$.

Definition 2.1. The set

$$\mathcal{C} := \left\{ x \in \mathbb{N}_0^{\mathbb{N}_0} : \begin{array}{l} \text{if } x(i) = 0 \text{ for some } i \in \mathbb{N}_0, \\ \text{then } x(j) = 0 \text{ for } j \in [\![i, \infty[\![\end{array} \right\}$$

is called the space of *paths*.

In the above, the elements of $\mathbb{N}$ serve as possible states of paths, and 0 serves as a distinguished state, called the *cemetery*, since any chain which visits 0 gets trapped at 0 forever after. (Customary symbols for a cemetery state are $\partial, \dagger$, or Δ, but we choose to use 0 since it fits in well with the ordering on $\mathbb{N}_0$.) For $N \in \mathbb{N}$, we also define an analogous space of paths on the state space $[\![0, N]\!]$, via

$$\mathcal{C}_N := \left\{ x \in [\![0, N]\!]^{\mathbb{N}_0} : \begin{array}{l} \text{if } x(i) = 0 \text{ for some } i \in \mathbb{N}_0, \\ \text{then } x(j) = 0 \text{ for } j \in [\![i, \infty[\![\end{array} \right\}.$$

Observe in particular that for $M, N \in \mathbb{N}$ with $M \geq N$ we have $\mathcal{C}_N \subseteq \mathcal{C}_M \subseteq \mathcal{C}$. We endow $\mathcal{C}$ with the relative topology induced by the product topology on $\mathbb{N}_0^{\mathbb{N}_0}$, which is metrizable; thus, compactness is equivalent to sequential compactness, continuity is equivalent to sequential continuity, etc.

Lemma 2.2. *The space $\mathcal{C}$ is a Polish space, and the space $\mathcal{C}_N$ is a compact Polish space for each $N \in \mathbb{N}$.*

Proof. Since $\mathbb{N}_0^{\mathbb{N}_0}$ is Polish, the first claim follows if we can show that $\mathcal{C}$ is closed in $\mathbb{N}_0^{\mathbb{N}_0}$. This is immediate by writing

$$\mathcal{C} = \bigcap_{i \in \mathbb{N}_0} \left(\left\{ x \in \mathbb{N}_0^{\mathbb{N}_0} : x(i) \neq 0 \right\} \cup \bigcap_{j=i}^{\infty} \left\{ x \in \mathbb{N}_0^{\mathbb{N}_0} : x(j) = 0 \right\} \right)$$

and noting that $\{x \in \mathbb{N}_0^{\mathbb{N}_0} : x(k) = 0\}$ is clopen in $\mathbb{N}_0^{\mathbb{N}_0}$ for all $k \in \mathbb{N}_0$.

Now let $N \in \mathbb{N}$ be arbitrary. Since $[\![0, N]\!]^{\mathbb{N}_0}$ is a compact Polish space, the second claim follows if we can show that $\mathcal{C}_N$ is closed in $[\![0, N]\!]^{\mathbb{N}_0}$, and this follows by the same argument as above. $\qquad\square$

For $x \in \mathcal{C}$ and $N \in \mathbb{N}$, set $I_{x,N}(0) = \inf\{i \in \mathbb{N}_0 : x(i) \in [\![0, N]\!]\}$ and recursively $I_{x,N}(j+1) = \inf\{i \in \mathbb{N}_0 : i > I_{x,N}(j), x(i) \in [\![0, N]\!]\}$ for $j \in \mathbb{N}$; we use the standard convention that $\inf \emptyset = \infty$. Observe that $\{I_{x,N}(j)\}_{j \in \mathbb{N}}$ enumerates the indices at which the path x visits states in the set $[\![0, N]\!]$. In particular, $\{I_{x,N}(j)\}_{j \in \mathbb{N}}$ is non-decreasing and is possibly eventually equal to infinity.

Then, for $x \in \mathcal{C}$ and $N \in \mathbb{N}$, we define $P_N(x) \in \mathcal{C}_N$ via

$$(P_N(x))(j) = \begin{cases} x(I_{x,N}(j)), & \text{if } I_{x,N}(j) < \infty, \\ 0, & \text{if } I_{x,N}(j) = \infty. \end{cases}$$

Since $\mathcal{C}_M \subseteq \mathcal{C}$ for all $M \in \mathbb{N}$, we can also think of the map $P_N : \mathcal{C}_M \to \mathcal{C}_N$ for any $M, N \in \mathbb{N}$. Intuitively, $x \mapsto P_N(x)$ is the operation which removes from x all of its excursions that visit states higher than N, including a possible final excursion of infinite length; in the case that x has an infinite excursion, we pad the path $P_N(x)$ with 0s. Also note that, for all $N \in \mathbb{N}$, we have the fundamental projectivity property $P_N \circ P_{N+1} = P_N$ on $\mathcal{C}$.

Lemma 2.3. *For each $N \in \mathbb{N}$, the map $P_N : \mathcal{C} \to \mathcal{C}_N$ is continuous.*

Proof. An arbitrary cylinder set $C \subseteq \mathcal{C}_N$ is of the form

$$C = \{x \in \mathcal{C}_N : x(i_1) \in B_1, \ldots, x(i_\ell) \in B_\ell\}$$

for some $\ell \in \mathbb{N}_0$, distinct $i_1, \ldots, i_\ell \in \mathbb{N}_0$, and $B_1, \ldots, B_\ell \subseteq [\![0, N]\!]$. So, simply set $I = \max\{i_1, \ldots, i_\ell\}$ and define

$$C' = \left\{ x \in \mathcal{C} : \begin{array}{l} x(i) \in B_i \text{ for all } i \in \{i_1, \ldots, i_\ell\}, \text{ and} \\ x(i) \in [\![0, N]\!] \text{ for all } i \in [\![0, I]\!] \setminus \{i_1, \ldots, i_\ell\} \end{array} \right\}.$$

Notice that C' is a cylinder set in $\mathcal{C}$ and that $x \in C'$ clearly implies $P_N(x) \in C$, hence P_N is continuous. □

Lemma 2.4. *For each $x \in \mathcal{C}$, we have $P_N(x) \to x$ as $N \to \infty$.*

Proof. An arbitrary cylinder set $C \subseteq \mathcal{C}$ is of the form

$$C = \{x \in \mathcal{C} : x(i_1) \in B_1, \ldots, x(i_\ell) \in B_\ell\}$$

for some $\ell \in \mathbb{N}_0$, distinct $i_1, \ldots, i_\ell \in \mathbb{N}_0$, and $B_1, \ldots, B_\ell \subseteq \mathbb{N}_0$. Then set $I = \max\{i_1, \ldots, i_\ell\}$ and $M = \max\{x(k) : 0 \le k \le I\}$. Thus, if $x \in C$, then $N \ge M$ implies that the restrictions of the infinite sequences $P_N(x)$ and x to the interval of indices $[\![0, I]\!]$ must agree, hence $P_N(x) \in C$, as needed. □

Definition 2.5. The set

$$\mathcal{C} := \left\{ \{x_N\}_{N \in \mathbb{N}} \in \prod_{N \in \mathbb{N}} \mathcal{C}_N : P_N(x_{N+1}) = x_N \text{ for all } N \in \mathbb{N} \right\}$$

is called the space of *virtual paths*.

Lemma 2.6. *The space $\mathcal{C}$ is a compact Polish space.*

Proof. Since each $\mathcal{C}_N$ is a compact Polish space, the countable product $\prod_{N \in \mathbb{N}} \mathcal{C}_N$ is a compact Polish space. Thus, it suffices to show that $\boldsymbol{\mathcal{C}}$ is closed in $\prod_{N \in \mathbb{N}} \mathcal{C}_N$. To see this, just write

$$\boldsymbol{\mathcal{C}} = \bigcap_{N \in \mathbb{N}} \left\{ \{x_N\}_{N \in \mathbb{N}} \in \prod_{N \in \mathbb{N}} \mathcal{C}_N : P_N(x_{N+1}) = x_N \right\}$$

and note that each of these sets in the countable intersection is a closed subset of $\prod_{N \in \mathbb{N}} \mathcal{C}_N$ by virtue of Lemma 2.3. □

Lemma 2.7. *The map $\iota : \mathcal{C} \to \boldsymbol{\mathcal{C}}$ defined in the natural way via $\iota(x) = \{P_N(x)\}_{N \in \mathbb{N}}$ is well-defined, continuous, and injective.*

Proof. That this map is well-defined follows from the projectivity property of the projections $\{P_N\}_{N \in \mathbb{N}}$, and that it is continuous follows from Lemma 2.3. To see that it is injective, suppose that $x, y \in \mathcal{C}$ have $\iota(x) = \iota(y)$. Then for each $I \in \mathbb{N}_0$, set $N = \max\{\max\{x(k), y(k) : 0 \leq k \leq I\}\}$. The restrictions of $P_N(x)$ and x to $[\![0, I]\!]$ must agree, and the restrictions of $P_N(y)$ and y to the indices $[\![0, I]\!]$ must also agree. Thus $P_N(x) = P_N(y)$ implies that the restrictions of x and y to $[\![0, I]\!]$ must agree. As $I \in \mathbb{N}_0$ was arbitrary, this implies $x = y$. □

Lemma 2.8. *The space $\iota(\mathcal{C})$ is dense in $\boldsymbol{\mathcal{C}}$.*

Proof. Note that a basis for the topology of $\boldsymbol{\mathcal{C}}$ is given by all sets of the form

$$\{\{x_N\}_{N \in \mathbb{N}} \in \boldsymbol{\mathcal{C}} : x_{N_1} \in B_1, \ldots, x_{N_\ell} \in B_\ell\}$$

for some $\ell \in \mathbb{N}_0$, an increasing sequence $N_1, \ldots N_\ell \in \mathbb{N}$, and some non-empty open sets $B_i \subseteq \mathcal{C}_{N_i}$ for each $i = 1, \ldots \ell$. For any such set U which is non-empty, there is some $\{x_N\}_{N \in \mathbb{N}} \in U$, and it follows that the element $x_{N_\ell} \in \mathcal{C}_{N_\ell} \subseteq \mathcal{C}$ has $\iota(x_{N_\ell}) \in U$. Thus, $\iota(\mathcal{C})$ intersects every non-empty open set in $\boldsymbol{\mathcal{C}}$, so the result follows. □

In the previous two results we showed that $\boldsymbol{\mathcal{C}}$ contains a "copy" of $\mathcal{C}$, and that this copy is dense in $\boldsymbol{\mathcal{C}}$. Since we also showed that $\boldsymbol{\mathcal{C}}$ is compact, it is useful to regard $\boldsymbol{\mathcal{C}}$ as a certain kind of compactification of the space $\mathcal{C}$. Our final result completes the picture by giving a simple characterization of the elements of $\boldsymbol{\mathcal{C}} \setminus \iota(\mathcal{C})$.

Lemma 2.9. *A virtual path $\boldsymbol{x} = \{x_N\}_{N \in \mathbb{N}} \in \boldsymbol{\mathcal{C}}$ lies in $\iota(\mathcal{C})$ if and only if $\lim_{N \to \infty} x_N$ exists in $\mathcal{C}$ and $\boldsymbol{x} = \iota(\lim_{N \to \infty} x_N)$.*

Proof. If $\lim_{N \to \infty} x_N$ exists and $\boldsymbol{x} = \iota(\lim_{N \to \infty} x_N)$, then $\boldsymbol{x} \in \iota(\mathcal{C})$ trivially. Conversely, if we have $\boldsymbol{x} = \iota(x)$ for some $x \in \mathcal{C}$, then the result follows from Lemma 2.4. □

The preceding result states that $\lim_{N \to \infty} x_N$ is the only candidate for the preimage of $\boldsymbol{x} = \{x_N\}_{N \in \mathbb{N}}$ by ι. Thus, a virtual path fails to be identified with an actual path whenever either this candidate x does not exist, or it exists but it does not satisfy the identity $\boldsymbol{x} = \iota(x)$.

Now, let's see some simple examples of virtual paths. Our first three examples illustrate the ways in the hypotheses of Lemma 2.9 come into effect, and hence highlight the relationship between $\boldsymbol{\mathcal{C}}$ and $\mathcal{C}$.

Example 2.10. For each $N \in \mathbb{N}$, define the path $x_N \in \mathcal{C}_N$ via

$$x_N(i) = \begin{cases} i+1, & \text{if } i < N, \\ 0, & \text{if } i \geq N, \end{cases}$$

for $i \in \mathbb{N}_0$; see Figure 2 (top) for a depiction. Then, $\boldsymbol{x} = \{x_N\}_{N \in \mathbb{N}}$ is a virtual path. Also define the path $x \in \mathcal{C}$ via $x(i) = i + 1$ for $i \in \mathbb{N}_0$, and note that we have $\boldsymbol{x} = \iota(x)$.

Example 2.11. For each $N \in \mathbb{N}$, define the path $x_N \in \mathcal{C}_N$ via

$$x_N(i) = \begin{cases} i+1, & \text{if } i < N, \\ 2N - i - 1, & \text{if } N \leq i < 2N - 1, \\ 1, & \text{if } i \geq 2N - 1, \end{cases}$$

for $i \in \mathbb{N}_0$, and see Figure 2 (middle) for an illustration. Then, $\boldsymbol{x} = \{x_N\}_{N \in \mathbb{N}}$ is a virtual path. Intuitively, $\boldsymbol{x}$ is the virtual path which "goes to infinity and back" and then stays forever at the state 1. Observe that we have $\lim_{N \to \infty} x_N = x$, where $x \in \mathcal{C}$ is defined in the previous example, but that $\iota(x) \neq \boldsymbol{x}$. Thus, by Lemma 2.9, there is no path which corresponds to this virtual path.

Example 2.12. For each $N \in \mathbb{N}$, define the path $x_N \in \mathcal{C}_N$ via

$$x_N(i) = \begin{cases} N - i, & \text{if } i < N, \\ 1, & \text{if } i \geq N, \end{cases}$$

for $i \in \mathbb{N}_0$, and see Figure 2 (middle). Then, $\boldsymbol{x} = \{x_N\}_{N \in \mathbb{N}}$ is a virtual path. Intuitively, $\boldsymbol{x}$ is the virtual path which "comes down from infinity" and then stays forever at the state 1. Observe that $x_N(0) = N$, so $\lim_{N \to \infty} x_N$ cannot exist in $\mathcal{C}$. Thus, by Lemma 2.9, we see that there is no path which corresponds to this virtual path.

For our last class of examples, we connect the ideas herein to the concept of virtual permutations, as constructed in [**10**]. Recall that a *virtual permutation* is a collection $\boldsymbol{\sigma} = \{\sigma_N\}_{N \in \mathbb{N}} \in \boldsymbol{S}$ where σ_N is a permutation of $[\![1, N]\!]$ for each $N \in \mathbb{N}$, such that the cycle structure of σ_N is given by removing the element $N + 1$ from the cycle structure of σ_N.

Example 2.13. Let $\boldsymbol{\sigma} = \{\sigma_N\}_{N \in \mathbb{N}}$ be a virtual permutation, and let $a \in \mathbb{N}$ be arbitrary. Then define $\boldsymbol{x}_{\boldsymbol{\sigma},a} := \{x_N\}_{N \in \mathbb{N}}$ in $\boldsymbol{C}$ as follows: For $N \geq a$, set $x_N(0) := a$ and recursively $x_N(j + 1) := \sigma_N(x_N(j))$ for $j \in \mathbb{N}_0$. For $N < a$, recursively set $x_N := P_N(x_{N+1})$. Then $\boldsymbol{x}_{\boldsymbol{\sigma},a}$ represents the "cycle" of $\boldsymbol{\sigma}$ containing a. Observe by Lemma 2.9 that $\boldsymbol{x}_{\boldsymbol{\sigma},a} \in \iota(\mathcal{C})$ if and only if this cycle is finite.

2.2. Virtual Markov chains. Analogously to how a stochastic process can be identified with its law on a suitable space of paths, we propose that a probability measure on $\boldsymbol{C}$ should be regarded as a *virtual stochastic process*. Our main object of study is the collection of virtual stochastic processes which are analogous to the case of Markov chains (MCs) in the classical sense.

For a topological space S, write $\mathcal{B}(S)$ for the Borel σ-algebra on S; whenever we view a topological space as a measurable space, we assume that it is endowed with its Borel σ-algebra, unless stated otherwise. We also write $\mathcal{M}(S)$ for the space of
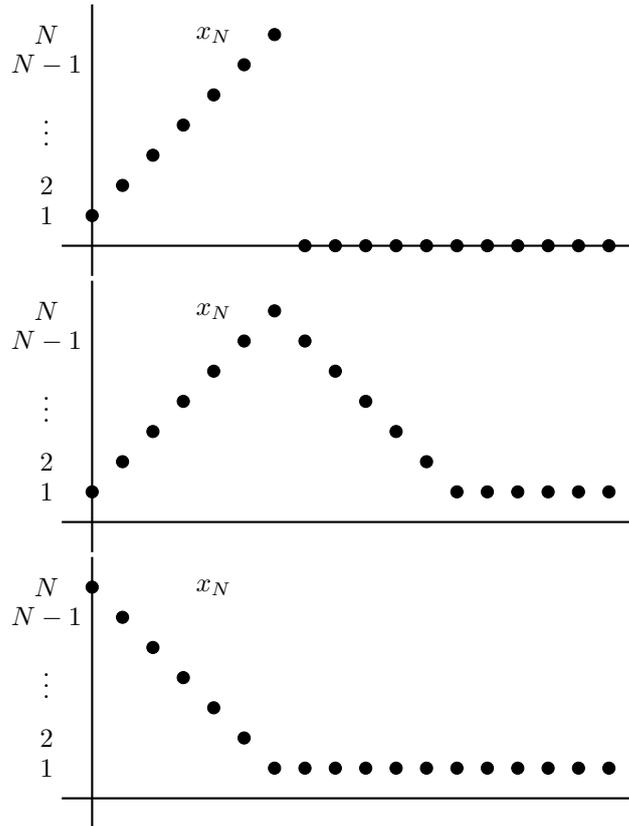
FIGURE 2. The virtual paths of Example 2.10 (top), Example 2.11 (middle), and Example 2.12 (bottom).

signed finite Borel measures on S, and $\mathcal{M}_1(S)$ for the space of non-negative Borel probability measures on S. Let us say that X *is a Markov chain in* $\mathbb{N}$ *with cemetery 0* if X is a Markov chain in $\mathbb{N}_0$ such that 0 is an absorbing state. Similarly, for any $N \in \mathbb{N}$, let us say that X *is a Markov chain in* $[\![1, N]\!]$ *with cemetery 0* if X is a Markov chain in $[\![0, N]\!]$ such that 0 is an absorbing state.

Definition 2.14. A $\mathcal{C}$-valued random variable $\boldsymbol{X} = \{X_N\}_{N \in \mathbb{N}}$ is called a *virtual Markov chain (VMC)* if, for each $N \in \mathbb{N}$, the random variable X_N is a Markov chain in $[\![1, N]\!]$ with cemetery 0.

Remark 2.15. It would be slightly more precise to call such objects *Markovian virtual paths*, but we stick with the slightly imprecise name because it sounds more natural.

Before we develop further properties of VMCs, we should see that they indeed generalize classical MCs. Towards this end, the following is an extension of Lemma 2.9 to the Markovian setting.

Proposition 2.16. *If X is a MC in $\mathbb{N}$ with cemetery 0, then $\iota(X)$ is a VMC. Conversely, if a VMC $\boldsymbol{X}$ satisfies $\boldsymbol{X} \in \iota(\mathcal{C})$ almost surely, then there is a Markov chain X in $\mathbb{N}$ with cemetery 0 such that $\boldsymbol{X} = \iota(X)$ almost surely.*

Proof. Suppose X is a Markov chain in $\mathbb{N}$ with cemetery 0 on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then write $\iota(X) = \{X_N\}_{N \in \mathbb{N}}$, so that $X_N = P_N(X)$ for all $N \in \mathbb{N}$, and note that we only need to check that X_N is a Markov chain in $[\![1, N]\!]$ with cemetery 0 for all $N \in \mathbb{N}$. Indeed, take any $N \in \mathbb{N}, \ell \in \mathbb{N}_0$ and $a_0, \ldots, a_{\ell+1} \in [\![0, N]\!]$, and let us show

$$\mathbb{P}(X_N(\ell+1) = a_{\ell+1} \mid X_N(0) = a_0, \ldots, X_N(\ell) = a_\ell)$$
$$= \mathbb{P}(X_N(\ell+1) = a_{\ell+1} \mid X_N(\ell) = a_\ell).$$

Note that, if $a_j = 0$ for some $j \in [\![0, \ell]\!]$, then this forces $a_\ell = 0$, hence both sides above are equal to $\mathbb{1}\{a_{\ell+1} = 0\}$. Therefore, we can assume $a_0, \ldots, a_\ell \in [\![1, N]\!]$, and we recall that $\{I_{X,N}(j)\}_{j \in \mathbb{N}_0}$ is a sequence of stopping times. So, if $a_{\ell+1} = 0$, then the strong Markov property implies

$$\mathbb{P}(X_N(\ell+1) = 0 \mid X_N(0) = a_0, \ldots, X_N(\ell) = a_\ell)$$
$$= \mathbb{P}(I_{X,N}(\ell+1) = \infty \mid X(I_{X,N}(0)) = a_0, \ldots, X(I_{X,N}(\ell)) = a_\ell)$$
$$= \mathbb{P}(I_{X,N}(\ell+1) = \infty \mid X(I_{X,N}(\ell)) = a_j)$$
$$= \mathbb{P}(X_N(\ell+1) = a_{\ell+1} \mid X_N(\ell) = a_\ell),$$

and, if $a_{\ell+1} \neq 0$, then the strong Markov property implies

$$\mathbb{P}(X_N(\ell+1) = a_{\ell+1} \mid X_N(0) = a_0, \ldots, X_N(\ell) = a_\ell)$$
$$= \mathbb{P}(X(I_{X,N}(\ell+1)) = a_{\ell+1} \mid X(I_{X,N}(0)) = a_0, \ldots, X(I_{X,N}(\ell)) = a_\ell)$$
$$= \mathbb{P}(X(I_{X,N}(\ell+1)) = a_{\ell+1} \mid X(I_{X,N}(\ell)) = a_j)$$
$$= \mathbb{P}(X_N(\ell+1) = a_{\ell+1} \mid X_N(\ell) = a_\ell).$$

This shows that X_N is a Markov chain in $[\![1, N]\!]$ with cemetery 0 for each $N \in \mathbb{N}$, as needed.

Conversely, suppose that $\boldsymbol{X} = \{X_N\}_{N \in \mathbb{N}}$ is a VMC on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with $\mathbb{P}(\boldsymbol{X} \in \iota(\mathcal{C})) = 1$. Then by Lemma 2.9, the limit $X = \lim_{N \to \infty} X_N$ exists in $\mathcal{C}$ and we have $\boldsymbol{X} = \iota(X)$ almost surely. In particular, X is measurable since X_N is measurable for each $N \in \mathbb{N}$. Now we just need to show that X is a Markov chain in $\mathbb{N}$ with cemetery 0. To do this, take any $\ell \in \mathbb{N}_0$ and $a_0, \ldots, a_{\ell+1} \in \mathbb{N}_0$, and note that we have, by the Markov property of X_N for each $N \in \mathbb{N}$:

$$\mathbb{P}(X(\ell+1) = a_{\ell+1} \mid X(0) = a_0, \ldots, X(\ell) = a_\ell)$$
$$= \lim_{N \to \infty} \mathbb{P}(X_N(\ell+1) = a_{\ell+1} \mid X_N(0) = a_0, \ldots, X_N(\ell) = a_\ell)$$
$$= \lim_{N \to \infty} \mathbb{P}(X_N(\ell+1) = a_{\ell+1} \mid X_N(\ell) = a_\ell)$$
$$= \mathbb{P}(X(\ell+1) = a_{\ell+1} \mid X(\ell) = a_\ell).$$

This shows that X is a Markov chain in $\mathbb{N}_0$, as needed. To see that 0 is a cemetery for X, let $\ell \in \mathbb{N}_0$ and $a \in \mathbb{N}_0$ be arbitrary, and note that, since 0 is cemetery for

X_N for each $N \in \mathbb{N}$, we have:

$$\mathbb{P}(X(\ell + 1) = a \,|\, X(\ell) = 0)$$
$$= \lim_{N \to \infty} \mathbb{P}(X_N(\ell + 1) = a \,|\, X_N(\ell) = 0)$$
$$= \mathbb{1}\{a = 0\}.$$

This shows that X is a Markov chain in $\mathbb{N}$ with cemetery 0, as needed. $\quad\square$

In the result above it is established that the virtual Markov chains are a generalization of the finite state space Markov chains, but only set-theoretically. The next step is to show that the same result holds topologically, that is, that the virtual Markov chains can be seen as a compactification of the finite state space Markov chains.

Definition 2.17. Write

$$\mathcal{M} := \{\mu \in \mathcal{M}_1(\mathcal{C}) : \mu \text{ is the law of a MC in } \mathbb{N} \text{ with cemetery } 0\},$$

as well as $\boldsymbol{\mathcal{M}} := \{\mu \in \mathcal{M}_1(\boldsymbol{\mathcal{C}}) : \mu \text{ is the law of a VMC}\}$.

Lemma 2.18. *The space* $\boldsymbol{\mathcal{M}}$ *is compact.*

Proof. Since $\boldsymbol{\mathcal{C}}$ is compact by Lemma 2.6, it follows that $\mathcal{M}_1(\boldsymbol{\mathcal{C}})$ is compact. Thus, it suffices to show that $\boldsymbol{\mathcal{M}}$ is closed in $\mathcal{M}_1(\boldsymbol{\mathcal{C}})$. Indeed, suppose that $\{\mu_n\}_{n \in \mathbb{N}}$ in $\boldsymbol{\mathcal{M}}$ have $\mu_n \to \mu$ for some $\mu \in \mathcal{M}_1(\boldsymbol{\mathcal{C}})$. Then for any $N \in \mathbb{N}, \ell \in \mathbb{N}_0$, and $a_0, \dots a_\ell \in [\![0, N]\!]$, we have

$$\mu(X_N(\ell + 1) = a_{\ell+1} \,|\, X_N(0) = a_0, \dots, X_N(\ell) = a_\ell)$$
$$= \lim_{n \to \infty} \mu_n(X_N(\ell + 1) = a_{\ell+1} \,|\, X_N(0) = a_0, \dots, X_N(\ell) = a_\ell)$$
$$= \lim_{n \to \infty} \mu_n(X_N(\ell + 1) = a_{\ell+1} \,|\, X_N(\ell) = a_\ell)$$
$$= \mu(X_N(\ell + 1) = a_{\ell+1} \,|\, X_N(\ell) = a_\ell).$$

This proves $\mu \in \boldsymbol{\mathcal{M}}$, as needed. $\quad\square$

Now we define a map $\iota : \mathcal{M} \to \boldsymbol{\mathcal{M}}$ as follows: If $\mu \in \mathcal{M}$ is the law of a Markov chain, then write $(\Omega, \mathcal{F}, \mathbb{P})$ for a probability space on which is defined a Markov chain X, and set $\iota(\mu) := \mathbb{P} \circ (\iota(X))^{-1}$.

Lemma 2.19. *The map* $\iota : \mathcal{M} \to \boldsymbol{\mathcal{M}}$ *is well-defined, injective, and continuous.*

Proof. That ι is well-defined follows from Proposition 2.16, since $\iota(X)$ is a VMC on $(\Omega, \mathcal{F}, \mathbb{P})$. To see that ι is injective, suppose that $\mu, \mu' \in \mathcal{M}$ have $\iota(\mu) = \iota(\mu')$. Write $(\Omega, \mathcal{F}, \mathbb{P})$ and $(\Omega', \mathcal{F}', \mathbb{P}')$ for probability spaces on which are defined MCs X and X' with laws μ and μ', respectively, and write $\iota(X) = \boldsymbol{X} = \{X_N\}_{N \in \mathbb{N}}$ and $\iota(X') = \boldsymbol{X}' = \{X'_N\}_{N \in \mathbb{N}}$. In the proof of Proposition 2.16 it was shown that we have $\lim_{N \to \infty} X_N = X$ holding $\mathbb{P}$-almost surely and $\lim_{N \to \infty} X'_N = X'$ holding $\mathbb{P}'$-almost surely, hence $\mathbb{P} \circ \boldsymbol{X}^{-1} = \mathbb{P}' \circ (\boldsymbol{X}')^{-1}$ implies $\mu = \mathbb{P} \circ X^{-1} = \mathbb{P}' \circ (X')^{-1} = \mu'$.

To see that ι is continuous, suppose that $\{\mu_n\}_{n \in \mathbb{N}}$ and μ in $\mathcal{M}$ have $\mu_n \to \mu$. By [**6**, Theorem 4.30], we can get a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with $\{X_n\}_{n \in \mathbb{N}}$ and X Markov chains such that $X_n \to X$ holds $\mathbb{P}$-almost surely. By Lemma 2.7, this implies that $\iota(X_n) \to \iota(X)$ holds $\mathbb{P}$-almost surely, hence by the definitions that we have $\iota(\mu_n) \to \iota(\mu)$. $\quad\square$

Lemma 2.20. *The space $\iota(\mathcal{M})$ is dense in $\boldsymbol{\mathcal{M}}$.*

Proof. Suppose $\mu \in \boldsymbol{\mathcal{M}}$, and let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which is defined $\boldsymbol{X} = \{X_N\}_{N \in \mathbb{N}}$ a VMC with law μ. Then for each $n \in \mathbb{N}$, set $\boldsymbol{X}_n := \iota(X_n) = \{P_N(X_n)\}_{N \in \mathbb{N}}$. Now note that, for each $M \in \mathbb{N}$ and $A \in \mathcal{B}(C_M)$, we have $\mathbb{P}(P_M(X_n) \in A) \to \mathbb{P}(X_M \in A)$ as $n \to \infty$. Since $P_N(X_n)$ is a measurable function of $P_M(X_n)$ for all $N \in \mathbb{N}$ with $N \leq M$, this implies that for all $A_1 \in \mathcal{B}(C_1), \ldots, A_M \in \mathcal{B}(C_M)$ we have

$$\mathbb{P}(P_1(X_n) \in A_1, \ldots, P_M(X_n) \in A_M) \to \mathbb{P}(X_1 \in A_1, \ldots, X_M \in A_M)$$

as $n \to \infty$. In other words, if we define $\mu_n := \mathbb{P} \circ \boldsymbol{X}_n^{-1}$ for all $n \in \mathbb{N}$, then we have $\mu_n \to \mu$. Of course, we also have $\mu_n \in \iota(\mathcal{M})$, so the claim follows. $\square$

2.3. Examples. Before we develop any further theory of VMCs, it will be instructive to build a wide collection of examples. We have already seen in Example 2.13 that virtual permutations give rise to VMCs, although they provide no interesting behavior probabilistically. We have also seen in Proposition 2.16 that classical MCs give rise to VMCs, although they provide no interesting behavior "at infinity". Our goal is thus to find examples of VMCs that are non-trivial in both of these senses.

The main tool for doing this will be a general result which allows us to construct a large class of VMCs from various continuous-time stochastic processes. Before doing so, let us recall some definitions from [**12**]. When we say that

$$\mathcal{Y} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \{Y_t\}_{t \geq 0}, \{\theta_t\}_{t \geq 0}, \{\mathbb{P}_x\}_{x \in S})$$

is a *Markov process* in an abstract measurable space $(S, \mathcal{G})$, we mean the following:
- $(\Omega, \mathcal{F})$ is a measurable space,
- $\{\mathcal{F}_t\}_{t \geq 0}$ is a filtration of $\mathcal{F}$,
- $\{Y_t\}_{t \geq 0}$ is a collection of S-valued random variables adapted to $\{\mathcal{F}_t\}_{t \geq 0}$,
- $\{\theta_t\}_{t \geq 0}$ is a semigroup of measurable maps from Ω to itself satisfying $Y_s \circ \theta_t = Y_{t+s}$ for all $s, t \geq 0$,
- $\{\mathbb{P}_x\}_{x \in S}$ is a collection of probability measures on S such that the map $P(t, x, \Gamma) = \mathbb{P}_x(Y_t \in \Gamma)$ is a measurable function of $x \in S$ for each $t \geq 0$ and $\Gamma \in \mathcal{G}$, and
- for all $t, s \geq 0, \Gamma \in \mathcal{G}$, and $x \in S$, we have $\mathbb{P}_x(Y_{t+s} \in \Gamma \,|\, \mathcal{F}_t) = P(s, Y_t, \Gamma)$ holding $\mathbb{P}_x$-almost surely.

In this setting, for any $\mu \in \mathcal{M}_1(S)$, one defines the probability meaure $\mathbb{P}_\mu$ on S via $\mathbb{P}_\mu(A) = \int_{x \in S} \mathbb{P}_x(A) \, d\mu(x)$ for all $A \in \mathcal{B}(S)$. If S is a Radon topological space endowed with its Borel σ-algebra, then we say that $\mathcal{Y}$ is a *right process* if we additionally have the following:
- $\{\mathcal{F}_t\}_{t \geq 0}$ is right-continuous, and we have $\mathcal{F}_t = \bigcap_{\mu \in \mathcal{M}_1(S)} (\mathcal{F}_t \vee \mathcal{N}^\mu(\mathcal{F}))$ for all $t \geq 0$, where $\mathcal{N}^\mu(\mathcal{F})$ is the collection of all subsets of $\mathbb{P}_\mu$-null $\mathcal{F}$-measurable sets,
- the sample path $\{Y_t\}_{t \geq 0}$ is right-continuous $\mathbb{P}_x$-almost surely for all $x \in S$,
- $P(0, x, \{x\}) = 1$ holds for all $x \in S$, and
- two techinical conditions, usually denoted (HD1) and (HD2) which we will not describe, are satisfied.

See [**12**, Chapter 1, Section 8] for a precise description of (HD1) and (HD2), and for an authoritative account of the general theory of right processes. In particular, we note that right processes have the strong Markov property in that, if τ is any $\{\mathcal{F}_t\}_{t \geq 0}$-stopping time, then for any $s \geq 0, \Gamma \in \mathcal{B}(S)$, and $x \in S$, we have $\mathbb{P}_x(Y_{\tau+s} \in$

$\Gamma \mid \mathcal{F}_\tau) = P(s, Y_\tau, \Gamma)$ holding $\mathbb{P}_x$-almost surely. Moreover, the hitting times of right processes into Borel sets are always stopping times. We say that a point $x \in S$ is *irregular for itself with respect to* $\mathcal{Y}$ if we have $\mathbb{P}_x(\inf\{t > 0 : Y_t = x\} > 0) = 1$.

Usually we will be somewhat cavalier and refer to $(\Omega, \mathcal{F}, \{Y_t\}_{t\geq0}, \{\mathbb{P}_x\}_{x \in S})$ or $\{Y_t\}_{t\geq0}$ as the Markov process itself, even though the Markov process really refers to the entire ensemble of objects $\mathcal{Y}$. For the following result, however, we stick with the full rigorous detail:

Lemma 2.21. *Let* $\mathcal{Y} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq0}, \{Y_t\}_{t\geq0}, \{\theta_t\}_{t\geq0}, \{\mathbb{P}_x\}_{x \in S})$ *be a right process in a Radon topological space* S, *and let* $\{L_N\}_{N\in\mathbb{N}}$ *be a sequence of distinct elements of* S *which are irregular for themsevles with respect to* $\mathcal{Y}$. *Then for* $N \in \mathbb{N}$ *define the stopping times* $\tau_0^N := \inf\{t \geq 0 : Y_t \in \{L_1, \dots L_N\}\}$ *and* $\tau_{j+1}^N := \inf\{t > \tau_j^N : Y_t \in \{L_1, \dots L_N\}\}$ *for* $j \in \mathbb{N}_0$. *For* $x \in S$, *we have* $\mathbb{P}_x$-*almost surely that* $\{\tau_j^N\}_{j\in\mathbb{N}_0}$ *are non-decreasing, and strictly increasing whenever they are finite. Moreover, there is a well-defined path* $X_N = \{X_N(j)\}_{j\in\mathbb{N}_0}$ *given by*

$$X_N(j) := \begin{cases} a, & \text{if } \tau_j^N < \infty \text{ and } Y_{\tau_j^N} = L_a \text{ for } a \in [\![1, N]\!], \\ 0, & \text{if } \tau_j^N = \infty, \end{cases} \tag{2.1}$$

and $\mathbf{X} := \{X_N\}_{N\in\mathbb{N}}$ *is a VMC on* $(\Omega, \mathcal{F}, \mathbb{P}_x)$ *for all* $x \in S$.

Proof. Fix $x \in S$ and $N \in \mathbb{N}$, and write $\mathcal{L}_N = \{L_1, \dots L_N\}$ for convenience. Note that $\{\tau_j^N\}_{j\in\mathbb{N}_0}$ are obviously non-increasing, so our first step is to show that, $\mathbb{P}_x$-almost surely, we have $\tau_j^N < \tau_{j+1}^N$ whenever $\tau_j^N < \infty$. To do this, define the function $p_N : S \to [0, 1]$ via

$$p_N(y) = \mathbb{P}_y(\inf\{t > 0 : Y_t \in \mathcal{L}_N\} > 0).$$

For arbitrary $a \in [\![1, N]\!]$, note that every point being irregular for itself implies

$$\begin{aligned} p_N(L_a) &= \mathbb{P}_{L_a}(\inf\{t > 0 : Y_t \in \mathcal{L}_N\} > 0) \\ &= \mathbb{P}_{L_a}(\inf\{t > 0 : Y_t = L_a \text{ or } Y_t \in \mathcal{L}_N \setminus \{L_a\}\} > 0) \\ &= \mathbb{P}_{L_a}(\inf\{t > 0 : Y_t = L_a\} > 0 \text{ and } \inf\{t > 0 : Y_t \in \mathcal{L}_N \setminus \{L_a\}\} > 0) \\ &= \mathbb{P}_{L_a}(\inf\{t > 0 : Y_t \in \mathcal{L}_N \setminus \{L_a\}\} > 0). \end{aligned}$$

Also note that, on the event $\{\inf\{t > 0 : Y_t \in \mathcal{L}_N \setminus \{L_a\}\} = 0\}$, we have $Y_{t_n} \in \mathcal{L}_N \setminus \{L_a\}$ for some sequence of times $\{t_n\}_{n\in\mathbb{N}}$ with $t_n \to 0$. But $\mathcal{L}_N \setminus \{L_a\}$ is closed and Y is right continuous, so this implies $Y_0 \in \mathcal{L}_N \setminus \{L_a\}$, which is obviously a contradiction if $Y_0 = L_a$. Combining this with the above shows $p_N(L_a) = 1$ for all $a \in [\![1, N]\!]$. Then let $j \in \mathbb{N}$ be arbitrary. Note that the definition of $\{\tau_j^N\}_{j\in\mathbb{N}_0}$ implies $\mathbb{1}\{\tau_j^N < \tau_{j+1}^N\} = \mathbb{1}\{\inf\{t > 0 : Y_t \in \mathcal{L}_N\} > 0\} \circ \theta_{\tau_j^N}$, so the strong Markov property gives

$$\mathbb{P}_x\left(\tau_j^N < \tau_{j+1}^N \,\middle|\, \mathcal{F}_{\tau_j^N}\right) = p_N(Y_{\tau_j^N}) = 1.$$

Therefore, we have

$$\begin{aligned} \mathbb{P}_x(\tau_j^N < \tau_{j+1}^N, \tau_j^N < \infty) &= \mathbb{E}_x\left[\mathbb{1}\{\tau_j^N < \infty\}\mathbb{P}_x\left(\tau_j^N < \tau_{j+1}^N \,\middle|\, \mathcal{F}_{\tau_j^N}\right)\right] \\ &= \mathbb{P}_x\left(\tau_j^N < \infty\right). \end{aligned}$$

This shows $\mathbb{P}_x(\tau_j^N < \tau_{j+1}^N \mid \tau_j^N < \infty) = 1$, so intersecting this statement over all $j \in \mathbb{N}_0$ gives the desired result.

To see that (2.1) indeed constructs a well-defined element of $\mathcal{C}_N$, note that Y is right-continuous and $\mathcal{L}_N$ is closed, so we of course have $Y(\tau_j^N) \in \mathcal{L}_N$ holding $\mathbb{P}_x$-almost surely on the event $\{\tau_j^N < \infty\}$. Then since $L_1, \ldots, L_N$ are distinct, there exists a unique $a \in [\![1, N]\!]$ such that $Y(\tau_j^N) = L_a$. The measurability of $X_N(j)$ follows from the facts that X_N is progressively measurable and that τ_j^N is a stopping time.

It only remains to show that X_N is a Markov chain in $[\![1, N]\!]$ with cemetery 0 for each $N \in \mathbb{N}$. Indeed, take any $N \in \mathbb{N}, \ell \in \mathbb{N}_0$, and $a_0, \ldots, a_{\ell+1} \in [\![0, N]\!]$, and let us show

$$\mathbb{P}_x(X_N(\ell+1) = a_{\ell+1} \mid X_N(0) = a_0, \ldots X_N(\ell) = a_\ell)$$
$$= \mathbb{P}_x(X_N(\ell+1) = a_{\ell+1} \mid X_N(\ell) = a_\ell).$$

To do this, first suppose that $a_j = 0$ for some $j \in [\![0, \ell]\!]$. This implies $\tau_j^N = \infty$, hence $\tau_\ell^N = \infty$, so both sides above are equal to $\mathbb{1}\{a_{\ell+1} = 0\}$. Otherwise, we can assume $a_0, \ldots, a_\ell \in [\![1, N]\!]$. Now, if $a_{\ell+1} = 0$, we use the strong Markov property of Y to get:

$$\mathbb{P}_x(X_N(\ell+1) = 0 \mid X_N(0) = a_0, \ldots, X_N(\ell) = a_\ell)$$
$$= \mathbb{P}_x(\tau_{\ell+1}^N = \infty \mid Y_{\tau_0^N} = L_{a_0}, \ldots, Y_{\tau_\ell^N} = L_{a_\ell})$$
$$= \mathbb{P}_x(\tau_{\ell+1}^N = \infty \mid Y_{\tau_\ell^N} = L_{a_\ell})$$
$$= \mathbb{P}_x(X_N(\ell+1) = 0 \mid X_N(\ell) = a_\ell),$$

and, if $a_{\ell+1} \neq 0$, we use the strong Markov property again to get

$$\mathbb{P}_x(X_N(\ell+1) = a_{\ell+1} \mid X_N(0) = a_0, \ldots, X_N(\ell) = a_\ell)$$
$$= \mathbb{P}_x(Y_{\tau_{\ell+1}^N} = L_{a_{\ell+1}} \mid Y_{\tau_0^N} = L_{a_0}, \ldots, Y_{\tau_\ell^N} = L_{a_\ell})$$
$$= \mathbb{P}_x(Y_{\tau_{\ell+1}^N} = L_{a_{\ell+1}} \mid Y_{\tau_\ell^N} = L_{a_\ell})$$
$$= \mathbb{P}_x(X_N(\ell+1) = a_{\ell+1} \mid X_N(\ell) = a_\ell).$$

This finishes the proof of the result. $\square$

First let observe that the VMCs arising from virtual permutations via Example 2.13 can also be constructed via Lemma 2.21.

Example 2.22. Let $\boldsymbol{\sigma} = \{\sigma_N\}_{N \in \mathbb{N}} \in \boldsymbol{\mathcal{S}}$ be a virtual permutation, and set $S := \coprod_{c \in \mathbb{N}} \mathbb{T}_c$ where $\mathbb{T}_c$ is a copy of the circle $\mathbb{R}/\mathbb{Z}$ for each $c \in \mathbb{N}$. The goal will be to define $\{L_N\}_{N \in \mathbb{N}}$ in such a way that each "cycle" in $\boldsymbol{\sigma}$ corresponds to a unique circle $\mathbb{T}_c$ for $c \in \mathbb{N}$.

To do this, first set L_1 to be any point on $\mathbb{T}_1$, and then proceed recursively. For $N \in \mathbb{N}$, consider two cases: If $N+1$ lies in its own cycle in σ_{N+1}, then we set L_{N+1} to be any point in $\mathbb{T}_c$, where $c \in \mathbb{N}$ is the smallest index such that $\coprod_{c'=c+1}^{\infty} \mathbb{T}_c$ contains no elements of $\{L_{N'}\}_{N'=1}^N$. If $N+1$ lies in an existing cycle in σ_{N+1}, then there exist $a, b \in [\![1, N]\!]$ satisfying $\sigma_{N+1}(a) = N+1$ and $\sigma_{N+1}(N+1) = b$, and, by construction, this means there is some $c \in \mathbb{N}$ such that L_a and L_b both lie in $\mathbb{T}_c$, with the additional property that no element of $\{L_{N'}\}_{N'=1}^N$ lies in the clockwise arc between them. In this case, let L_{N+1} be an arbitrary point on this clockwise arc. (To remove the choice from the construction above, we could arbitrarily decide
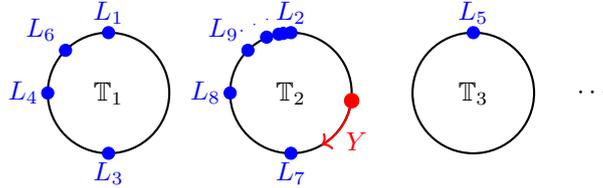
FIGURE 3. As in Example 2.22, we can construct the VMC corresponding to a virtual permutation via Lemma 2.21.

that points in new cycles get sent to 0 mod 1 and that points in old cycles get sent to the midpoints of the arcs they fall into.)

Now consider the non-random process Y which moves clockwise with rate one; let $\{\mathbb{P}_x\}_{x \in S}$ denote the (degenerate) probability measures governing this process. This clearly satisfies the hypotheses of Lemma 2.1, so we construct $\boldsymbol{X}$ by the result therein. Then, for arbitrary $a \in \mathbb{N}$, the VMC $\boldsymbol{X}$ on $(\Omega, \mathcal{F}, \mathbb{P}_{L_a})$ coincides with the virtual chain $\boldsymbol{x}_{\boldsymbol{\sigma},a}$ given by Example 2.13. To visualize this construction, see Figure 3.

Observe that $\boldsymbol{\sigma}$ corresponds to a classical permutation if and only if this construction is such that $\mathbb{T}_c \cap \{L_N\}_{N \in \mathbb{N}}$ is finite for each $c \in \mathbb{N}$. By compactness, this is equivalent to the statement that $\{L_N\}_{N \in \mathbb{N}}$ has no accumulation points.

In many (but not all) of the examples of VMCs that we will see throughout the paper, the cemetery state of 0 is never visited. In the setting of Lemma 2.21, a sufficient condition for this to occur is that the stopping times $\{\tau_j^N\}_{j \in \mathbb{N}_0, N \in \mathbb{N}}$ are all finite; this is in turn implied by the property that the process Y is point-recurrent in the sense that

$$\mathbb{P}_x(H_Y(x') \text{ is unbounded for all } x' \in S) = 1$$

for all $x \in S$, where we have defined the set $H_Y(x') := \{t \geq 0 : Y_t = x'\}$.

We also give another interesting class of examples arising from the theory of regenerative sets.

Example 2.23. Let F the distribution function of a measure on $(0, \infty)$, and let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which is defined a sequence $\{Z_k\}_{k \in \mathbb{N}}$ of i.i.d. random variables with distribution F. Then define the renewal (point) process E via

$$E := \left\{ \sum_{k=1}^K Z_k : K \in \mathbb{N}_0 \right\},$$

and note that E is a regenerative random closed set on $(\Omega, \mathcal{F}, \mathbb{P})$. It also has empty interior $\mathbb{P}$-almost surely, so by [**5**, Theorem 1], it follows that the age process $Y = \{Y_t\}_{t \geq 0}$ of E, given by

$$Y_t := t - \sup\{0 \leq s \leq t : s \in E\}$$

for $t \geq 0$ is a right process in the state space $S = [0, \infty)$. Now let $\{L_N\}_{N \in \mathbb{N}}$ be a strictly decreasing sequence with $L_N \downarrow 0$ as $N \to \infty$. Finally, let $\boldsymbol{X} = \{X_N\}_{N \in \mathbb{N}}$ be the VMC on $(\Omega, \mathcal{F}, \mathbb{P})$ guaranteed by Lemma 2.21 via (2.1).
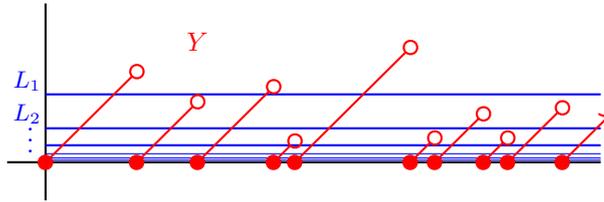
FIGURE 4. A generic sample path of Y and $\boldsymbol{X}$, when Y is the age process of a regenerative set as given in Example 2.23.

Note that, $\mathbb{P}$-almost surely, we have $Z_1 > 0$, hence $X_N(0) = N$ for all $N > \frac{1}{Z_1}$. Thus, $\mathbb{P}$-almost surely, the limit $\lim_{N \to \infty} X_N$ does not exist in $\mathcal{C}$. Therefore, by Lemma 2.9 we have $\mathbb{P}(\boldsymbol{X} \in \iota(\mathcal{C})) = 0$; see Figure 4 for an illustration. Intuitively speaking, then, $\boldsymbol{X}$ is a VMC which tries to "come down from infinity" but "jumps back to infinity" at some random times.

We also note that the construction above works for E any regenerative set with empty interior almost surely. In particular, it seems interesting to consider the case that E is the zero set of a Brownian motion, in which case there are very many short excursions down from infinity.

Example 2.24. Consider the space $S = (-1, 1)$ with the topology inherited form the usual topology on the real line. Now let $(\Omega, \mathcal{F}, \{Y_t\}_{t \geq 0}, \{\mathbb{P}_x\}_{x \in S})$ be the process which moves deterministically at rate one towards the closer of the two endpoints $\{-1, +1\}$, which teleports back to $x = 0$ once it hits one of these endpoints, and which breaks the tie at $x = 0$ by moving into either half of the domain with equal probability.

More concretely, the semigroup $\{P_t\}_{t \geq 0}$ of this process defined via $(P_t f)(x) := \mathbb{E}_x[f(Y_t)]$ for all bounded, measurable $f : S \to \mathbb{R}$ is exactly

$$(P_t f)(x) = \begin{cases} \frac{1}{2}f(t \bmod 1) + \frac{1}{2}f(-(t \bmod 1)), & \text{if } x = 0, \\ f(x + t) & \text{if } x > 0, \text{ and } |x| + t < 1, \\ f(x - t) & \text{if } x < 0, \text{ and } |x| + t < 1, \\ \frac{1}{2}f(x + t \bmod 1) + \frac{1}{2}f(-(x + t \bmod 1)), & \text{if } x > 0 \text{ and } |x| + t \geq 1, \\ \frac{1}{2}f(-x + t \bmod 1) + \frac{1}{2}f(-(-x + t \bmod 1)), & \text{if } x < 0 \text{ and } |x| + t \geq 1. \end{cases}$$

Of course, this process is point-recurrent and every point is irregular for itself. However, it is easy to see that this is not a strong Markov process with respect to the right-continuous augmentation of its natural filtration $\{\mathcal{F}_t\}_{t \geq 0}$. (For example, the strong Markov property fails when applied to the $\{\mathcal{F}_{t+}\}_{t \geq 0}$-stopping time $\tau := \inf\{t \geq \tau_0 : Y_t > 0\}$, where $\tau_0 := \inf\{t \geq 0 : Y_t = 0\}$.)

Now choose any sequence $\{\ell_N\}_{N \in \mathbb{N}}$ in $(0, 1)$ with $\ell_N \downarrow 0$ as $N \to \infty$, and define the points $\{L_N\}_{N \in \mathbb{N}}$ via

$$L_N = \begin{cases} \ell_{(N+1)/2}, & \text{if } N \text{ odd}, \\ -\ell_{N/2}, & \text{if } N \text{ even}. \end{cases}$$

An argument identical to that of Lemma 2.21 still applies to show that $\boldsymbol{X}$ constructed therein is a VMC, since the strong Markov property is only needed at the
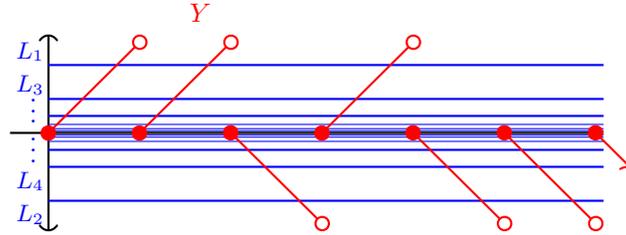
FIGURE 5. A generic sample path of the process Y described by Example 2.24.

hitting times of $\{L_N\}_{N\in\mathbb{N}}$. We regard $\boldsymbol{X} = \{X_N\}_{N\in\mathbb{N}}$ as a VMC on the probability space $(\Omega, \mathcal{F}, \mathbb{P}_0)$.

We of course have $\lim_{N\to\infty} X_N(0) = \infty$ holding $\mathbb{P}_0$-almost surely, hence $\mathbb{P}_0(\boldsymbol{X} \in \iota(\mathcal{C})) = 0$ by Proposition 2.16. See Figure 5 for a generic realization of the sample path of Y and $\boldsymbol{X}$ under $\mathbb{P}_0$. Intuitively speaking, $\boldsymbol{X}$ is a VMC which has "two different ways to come down from infinity"; it chooses one uniformly at random, traverses the (deterministic) path down to the lowest possible state, and then repeats the process over again.

This VMC has a non-trivial probabilistic component as well as non-trivial behavior "at infinity" as desired, but it is distinguished from out previous examples in that it also has non-trivial "probabilistic behavior at infinity". To see this, consider the *tail σ-algbra of $\boldsymbol{X}$*, defined via $\mathcal{T}(\boldsymbol{X}) = \bigcap_{N\in\mathbb{N}} \sigma(X_N, X_{N+1}, \ldots)$. Then, the event

$$A = \{X_N(0) \text{ is even for infinitely many } N \in \mathbb{N}\}$$

is $\mathcal{T}(\boldsymbol{X})$-measurable, and satisfies $\mathbb{P}_0(A) = 1/2$. This is closely related to the fact that the Blumenthal zero-one law fails for the constituent Markov process Y.

3. Canonical Data and the Representation Theorem

Suppose $(\Omega, \mathcal{F}, \mathbb{P})$ is some probability space on which is defined $X = \{X(i)\}_{i\in\mathbb{N}_0}$ a MC in $\mathbb{N}$ with cemetery 0. Then, X has an associated *initial distribution (ID)* $\nu = \mathbb{P} \circ X(0)^{-1}$ and *transition matrix (TM)* $K \in [0,1]^{\mathbb{N}_0 \times \mathbb{N}_0}$ given by $K(a,b) = \mathbb{P}(X(1) = b \,|\, X(0) = a)$ for $a, b \in \mathbb{N}_0$, which also satisfies $K(0,a) = \mathbb{1}\{a = 0\}$ for all $a \in \mathbb{N}_0$. In this case, we say that (ν, K) are the *canonical data* of X. It is classical that for every pair of canonical data there is a probability space on which is defined a MC with these canonical data; moreover, this object is unique in the sense that any two MCs with the same canonical data must have the same law. In other words, the collection of laws of MCs in $\mathbb{N}$ with cemetery 0 is in bijection with the collection of pairs of canonical data.

The goal of this section is show that an analogous statement is true for VMCs (Theorem 3.22), where we suitably define the notions of *virtual transition matrix (VTM*, Subsection 3.1*)* and *virtual initial distribution (VID*, Subsection 3.2*)*, and introduce an important notion of *compatibility* (Subsection 3.3) which is not needed in the MC case. In Subsection 3.4 we explore various examples.

3.1. Virtual transition matrices. First we define the important notion of a virtual transition matrix, which will later be seen to encode most of the "interesting" information about VMCs.

Definition 3.1. Write $\mathcal{K}$ for the space of all transition matrices on $\mathbb{N}_0$ with $K(0,0) = 1$, which we endow with the topology of pointwise convergence.

Moreover, for each $N \in \mathbb{N}$ write $\mathcal{K}_N$ for the space of all transition matrices on $\mathbb{N}_0$ with $K(a,a) = 1$ for all $a \in \mathbb{N}_0 \setminus [\![1,N]\!]$. Observe of course that we have $\mathcal{K}_N \subseteq \mathcal{K}_M \subseteq \mathcal{K}$ for all $N, M \in \mathbb{N}$ with $N \leq M$.

Next we define suitable projection maps for transition matrices from our existing projection maps for paths. Specifically, for each $K \in \mathcal{K}$ and $N \in \mathbb{N}$, we define $P_N(K)$ to be the transition matrix of the Markov chain $P_N(X)$ where X is a Markov chain with transition matrix K. More specifically, for each $K \in \mathcal{K}$ and $a \in \mathbb{N}_0$ write $(\Omega, \mathcal{F}, \mathbb{P}_{K,a})$ for a probability space on which is defined a MC $X = \{X(i)\}_{i \in \mathbb{N}_0}$ with transition matrix K and initial state $X(0) = a$. Then, for each $N \in \mathbb{N}$ and $a, b \in \mathbb{N}_0$, set

$$(P_N(K))(a,b) = \begin{cases} \mathbb{P}_{K,a}((P_N(X))(1) = b), & \text{if } a \in [\![1,N]\!], \\ 1, & \text{else if } a = b, \\ 0, & \text{else if } a \neq b. \end{cases} \tag{3.1}$$

It is easily verified that indeed $P_N(K) \in \mathcal{K}_N$.

The next goal is to develop some properties of these projection maps which parallel (to the extent possible) the analogous results for the projection maps on spaces of paths.

Lemma 3.2. *For $N \in \mathbb{N}$, we have $P_N \circ P_{N+1} = P_N$ on $\mathcal{K}$.*

Proof. For any $K \in \mathcal{K}$ and $a, b \in \mathbb{N}_0$ we have, from the projectivity property $P_N \circ P_{N+1} = P_N$ on $\mathcal{C}$,

$$\begin{aligned} (P_N(K))(a,b) &= \mathbb{P}_{K,a}((P_N(X))(1) = b) \\ &= \mathbb{P}_{K,a}((P_N(P_{N+1}(X)))(1) = b) \\ &= \mathbb{P}_{P_{N+1}(K),a}((P_N(X))(1) = b) \\ &= (P_N(P_{N+1}(K)))(a,b), \end{aligned}$$

as claimed.   □

Lemma 3.3. *For each $K \in \mathcal{K}$, we have $P_N(K) \to K$ as $N \to \infty$.*

Proof. Let $a, b \in \mathbb{N}_0$ be arbitrary. For $N \geq \max\{a,b\}$, we find ourselves in the first case of (3.1), hence we have

$$\lim_{N \to \infty} K_N(a,b) = \lim_{N \to \infty} \mathbb{P}_{K,a}(X_N(1) = b) = \mathbb{P}_{K,a}(X(1) = b) = K(a,b),$$

as needed.   □

Definition 3.4. The set

$$\boldsymbol{\mathcal{K}} := \left\{ \{K_N\}_{N \in \mathbb{N}} \in \prod_{N \in \mathbb{N}} \mathcal{K}_N : P_N(K_{N+1}) = K_N \text{ for all } N \in \mathbb{N} \right\}$$

is called the space of *virtual transition matrices (VTMs)*.

Lemma 3.5. *The map* $\iota : \mathcal{K} \to \boldsymbol{\mathcal{K}}$ *defined in the natural way via* $\iota(K) := \{P_N(K)\}_{N \in \mathbb{N}}$ *is well-defined and injective.*

Proof. To see $\iota(K) \in \boldsymbol{\mathcal{K}}$, simply apply Lemma 3.2. To see injectivity, suppose that $K, K' \in \mathcal{K}$ have $\iota(K) = \iota(K')$, and write $\iota(K) = \{K_N\}_{N \in \mathbb{N}}$ and $\iota(K') = \{K'_N\}_{N \in \mathbb{N}}$. Then, for any $a, b \in \mathbb{N}_0$, we have, by Lemma 3.3,

$$K(a, b) = \lim_{N \to \infty} K_N(a, b) = \lim_{N \to \infty} K'_N(a, b) = K'(a, b),$$

as claimed. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Lemma 3.6. *A VTM* $\boldsymbol{K} = \{K_N\}_{N \in \mathbb{N}} \in \boldsymbol{\mathcal{K}}$ *lies in* $\iota(\mathcal{K})$ *if and only if* $K = \lim_{N \to \infty} K_N$ *exists and* $\boldsymbol{K} = \iota(K)$.

Proof. If $\lim_{N \to \infty} K_N$ exists and $\boldsymbol{K} = \iota(\lim_{N \to \infty} K_N)$, then $\boldsymbol{K} \in \iota(\mathcal{K})$ trivially. Conversely, if we have $\boldsymbol{K} = \iota(K)$ for some $K \in \mathcal{K}$, then the result follows from Lemma 3.3. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Next let us show that the projection $P_N : \mathcal{K}_{N+1} \to \mathcal{K}_N$ for any $N \in \mathbb{N}$ has a particularly simple form which is very amenable to calculations.

Lemma 3.7. *For* $N \in \mathbb{N}$ *and* $K \in \mathcal{K}_{N+1}$, *we have*

$$(P_N(K))(a, b) = \begin{cases} K(a, b) + \frac{K(a, N+1)K(N+1, b)}{1 - K(N+1, N+1)}, & \text{if } K(N+1, N+1) < 1, \\ K(a, 0) + K(a, N+1), & \text{else if } b = 0, \\ K(a, b), & \text{else if } b \neq 0, \end{cases} \qquad (3.2)$$

for $a, b \in [\![0, N]\!]$, *as well as* $(P_N(K))(a, a) = 1$ *for* $a \in \mathbb{N}_0 \setminus [\![1, N]\!]$.

Proof. First observe that the claimed form of $P_N(K)$ is indeed a transition matrix, and moreover that we have $P_N(K) \in \mathcal{K}_N$. Next, note that $a \in \mathbb{N}_0 \setminus [\![1, N]\!]$ implies $(P_N(K))(a, a) = 1$ by definition, so we only need to check (3.2) for $a, b \in [\![1, N]\!]$. Now note, in particular, that $X_{N+1} = X$. Then, if $K(N+1, N+1) < 1$, we have:

$$\begin{aligned} (P_N(K))(a, b) &= \mathbb{P}_{K,a}(X_N(1) = b) \\ &= \mathbb{P}_{K,a}(I_{X,N}(1) < \infty, X(I_{X,N}(1)) = b) \\ &= K(a, b) + K(a, N+1) \sum_{k=1}^{\infty} (K(N+1, N+1))^k K(N+1, b) \\ &= K(a, b) + \frac{K(a, N+1)K(N+1, b)}{1 - K(N+1, N+1)} \\ &= (P_N(K))(a, b), \end{aligned}$$

as needed. Similarly, if $K(N+1, N+1) = 1$, then we have

$$\begin{aligned} (P_N(K))(a, 0) &= \mathbb{P}_{K,a}(X_N(1) = 0) \\ &= \mathbb{P}_{K,a}(I_{X,N} = \infty) + \mathbb{P}_{K,a}(I_{X,N}, \infty, X(I_{X,N}) = 0) \\ &= K(a, 0) + K(a, N+1) \\ &= (P_N(K))(a, 0) \end{aligned}$$

and, for $b \in [\![1, N]\!]$, we have

$$
\begin{aligned}
(P_N(K))(a,b) &= \mathbb{P}_{K,a}(X_N(1) = b) \\
&= \mathbb{P}_{K,a}(I_{X,N} < \infty, X_N(I_{X,N}) = b) \\
&= K(a,b) \\
&= (P_N(K))(a,b).
\end{aligned}
$$

This proves the claim. $\square$

Remark 3.8. The description of the map $P_N : \mathcal{K}_{N+1} \to \mathcal{K}_N$ can be made even more explicit by regarding elements $\mathcal{K}_N$ for $N \in \mathbb{N}$ as suitable block matrices. More specifically, by a slight abuse of notation, we identify an element $K \in \mathcal{K}_{N+1}$ for $N \in \mathbb{N}$ as its top-left $(N+2) \times (N+2)$ submatrix; in this setting, if an element $K_{N+1} \in \mathcal{K}_{N+1}$ is identified with

$$
\begin{pmatrix} 1 & 0 & 0 \\ w & A & u \\ q & v^{\mathrm{T}} & p \end{pmatrix}, \tag{3.3}
$$

where $A \in [0,1]^{N \times N}$ is a matrix, $u, v, w \in [0,1]^N$ are column vectors, and $p, q \in [0,1]$ are real numbers, then the projection $P_N : \mathcal{K}_{N+1} \to \mathcal{K}_N$ is exactly

$$
\begin{pmatrix} 1 & 0 & 0 \\ w & A & u \\ q & v^{\mathrm{T}} & p \end{pmatrix} \mapsto
\begin{cases}
\begin{pmatrix} 1 & 0 \\ w + \frac{q}{1-p}u & A + \frac{1}{1-p}uv^{\mathrm{T}} \end{pmatrix}, & \text{if } p < 1, \\[2em]
\begin{pmatrix} 1 & 0 \\ w + u & A \end{pmatrix}, & \text{if } p = 1.
\end{cases} \tag{3.4}
$$

It does not appear that there is an analogous simple formula for the composite projection map $P_N \circ P_{N+1} : \mathcal{K}_{N+2} \to \mathcal{K}_N$ for $N \in \mathbb{N}$, nor for the general $P_N : \mathcal{K} \to \mathcal{K}_N$ for $N \in \mathbb{N}$.

Remark 3.9. Note that, as $p \to 1$, we have

$$
K_3(p) := \begin{pmatrix} 1 & 0 & 0 \\ 1-p & 0 & p \\ 0 & 1-p & p \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} = K_3(1)
$$

and

$$
P_2(K_3(p)) = \begin{pmatrix} 1 & 0 \\ 1-p & p \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = P_2(K_3(1)).
$$

This example shows that the projections $P_N : \mathcal{K} \to \mathcal{K}_N$ for $N \in \mathbb{N}$ and the inclusion $\iota : \mathcal{K} \to \boldsymbol{\mathcal{K}}$ can fail to be continuous, and also that $\boldsymbol{\mathcal{K}}$ is not closed.

3.2. Virtual initial distributions. While it was relatively straightforward to extend the notion of transition matrices to that of virtual transition matrices, the process of extending the notion of initial distributions (IDs) to an analogous notion for VMCs is somewhat more complicated.

Definition 3.10. The set

$$
\mathcal{I} := \left\{ \{y_N\}_{N \in \mathbb{N}} \in \prod_{N \in \mathbb{N}} [\![0, N]\!] : \begin{array}{c} \text{for all } N \in \mathbb{N}, \text{ either} \\ y_{N+1} = N + 1 \text{ or } y_{N+1} = y_N \end{array} \right\}
$$

is called the space of *virtual initial states (VISs)*.

Observe that $\boldsymbol{\mathcal{I}}$ is closed in $\prod_{N\in\mathbb{N}}[\![0,N]\!]$, hence $\boldsymbol{\mathcal{I}}$ is itself a compact Polish space. For for $N \in \mathbb{N}$, we also define the truncated sets

$$\boldsymbol{\mathcal{I}}_N := \left\{ \{y_M\}_{M=1}^N \in \prod_{M=1}^N [\![0,M]\!] : \begin{array}{l} \text{for all } 1 \leq M < N, \text{ either} \\ y_{M+1} = M+1 \text{ or } y_{M+1} = y_M \end{array} \right\},$$

which are finite sets. Observe that elements of $\boldsymbol{\mathcal{I}}$ correspond to infinite sequences consisting of flat stretches and jumps, where a jump at index $N \in \mathbb{N}$ must jump to level $N \in \mathbb{N}$; elements of $\boldsymbol{\mathcal{I}}_N$ for $N \in \mathbb{N}$ are the analogous finite sequences.

Definition 3.11. A probability measure λ on $\mathcal{M}_1(\boldsymbol{\mathcal{I}})$ is called *Markovian* if for all $N \in \mathbb{N}$ and $\{a_M\}_{M=1}^{N+1} \in \boldsymbol{\mathcal{I}}_{N+1}$ we have $\lambda(y_{N+1} = a_{N+1} \mid y_1 = a_1, \ldots, y_N = a_N) = \lambda(y_{N+1} = a_{N+1} \mid y_N = a_N)$. Then define the set $\boldsymbol{\mathcal{D}}_0 = \{\lambda \in \mathcal{M}_1(\boldsymbol{\mathcal{I}}) : \lambda \text{ is Markovian}\}$, called the space of *virtual distributions (VDs)*.

However, it turns out that the elements of $\boldsymbol{\mathcal{D}}_0$ are in bijection with a simpler space which is much easier to work with. To state this, consider the following:

Definition 3.12. The set

$$\boldsymbol{\mathcal{D}}_1 := \left\{ \{\nu_N\}_{N\in\mathbb{N}} \in \prod_{N\in\mathbb{N}} \mathcal{M}_1([\![0,N]\!]) : \nu_N(a) \geq \nu_{N+1}(a) \text{ for } N \in \mathbb{N}, a \in [\![0,N]\!] \right\}$$

is, by abuse, also called the space of *virtual distributions (VDs)*.

Observe that elements of $\boldsymbol{\mathcal{D}}_1$ are just sequences of marginal distributions, which a priori, do not determine a unique coupling on the entire product space. The main result of this subsection is the following, which establishes the useful fact that $\boldsymbol{\mathcal{D}}_0$ and $\boldsymbol{\mathcal{D}}_1$ are two ways of viewing the same object.

Proposition 3.13. *There is a homeomorphism from $\boldsymbol{\mathcal{D}}_0$ to $\boldsymbol{\mathcal{D}}_1$ given by sending each probability distribution to its sequence of marginal distributions.*

Proof. Note that for $\lambda \in \boldsymbol{\mathcal{D}}_0$ and any $N \in \mathbb{N}$ and $a \in [\![0,N]\!]$, we have

$$\begin{aligned} \lambda(y_N = a) &= \lambda(y_N = a, y_{N+1} = a) + \lambda(y_N = a, y_{N+1} = N+1) \\ &\geq \lambda(y_N = a, y_{N+1} = a) \\ &= \lambda(y_{N+1} = a). \end{aligned}$$

Thus there is a well-defined map $\phi : \boldsymbol{\mathcal{D}}_0 \to \boldsymbol{\mathcal{D}}_1$ sending each element $\lambda \in \boldsymbol{\mathcal{D}}_0$ to $\phi(\lambda)$ its sequence of marginal distributions in $\boldsymbol{\mathcal{D}}_1$.

Next, let us show that ϕ is injective. Take $\lambda, \lambda' \in \boldsymbol{\mathcal{D}}_0$ and write $\phi(\lambda) = \{\nu_N\}_{N\in\mathbb{N}}$ and $\phi(\lambda') = \{\nu_N'\}_{N\in\mathbb{N}}$, and suppose $\phi(\lambda) = \phi(\lambda')$. For $N \in \mathbb{N}$ and $\{a_M\}_{M=1}^N \in \boldsymbol{\mathcal{I}}_N$ write

$$E_N(\{a_M\}_{M=1}^N) := \{\{y_N\}_{N\in\mathbb{N}} \in \boldsymbol{\mathcal{I}} : y_1 = a_1, \ldots, y_N = a_N\},$$

and note that

$$\mathcal{P} := \left\{ E_N(\{a_M\}_{M=1}^N) : N \in \mathbb{N}, \{a_M\}_{M=1}^N \in \boldsymbol{\mathcal{I}}_N \right\}$$

is a π-system in $\boldsymbol{\mathcal{I}}$ which generates the Borel σ-algebra $\mathcal{B}(\boldsymbol{\mathcal{I}})$. So, in order to show $\lambda = \lambda'$ it suffices to show that they agree for all elements of $\mathcal{P}$. To do this, we'll show by induction on $N \in \mathbb{N}$ that they agree for all elements of

$$\mathcal{P}_N := \left\{ E_N(\{a_M\}_{M=1}^N) : \{a_M\}_{M=1}^N \in \boldsymbol{\mathcal{I}}_N \right\}.$$

The base case of $N = 1$ is immediate since $\mathcal{I}_1 = [\![0, 1]\!]$. Before we perform the inductive step, we note that, for any $N \in \mathbb{N}$ and $a \in [\![0, N]\!]$, we have

$$
\begin{aligned}
\lambda(y_{N+1} = a | y_N = a) &= \frac{\lambda(y_{N+1} = a, y_N = a)}{\lambda(y_N = a)} \\
&= \frac{\lambda(y_{N+1} = a)}{\lambda(y_N = a)} \\
&= \frac{\lambda'(y_{N+1} = a)}{\lambda'(y_N = a)} \\
&= \frac{\lambda'(y_{N+1} = a, y_N = a)}{\lambda'(y_N = a)} = \lambda'(y_{N+1} = a | y_N = a),
\end{aligned} \tag{3.5}
$$

and, similarly,

$$
\begin{aligned}
\lambda(y_{N+1} = N+1 | y_N = a) &= \frac{\lambda(y_{N+1} = N+1, y_N = a)}{\lambda(y_N = a)} \\
&= \frac{\lambda(y_N = a) - \lambda(y_{N+1} = a, y_N = a)}{\lambda(y_N = a)} \\
&= \frac{\lambda(y_N = a) - \lambda(y_{N+1} = a)}{\lambda(y_N = a)} \\
&= \frac{\lambda'(y_N = a) - \lambda'(y_{N+1} = N+1)}{\lambda'(y_N = a)} \\
&= \frac{\lambda'(y_N = a) - \lambda'(y_{N+1} = a, y_N = a)}{\lambda'(y_N = a)} \\
&= \frac{\lambda'(y_{N+1} = N+1, y_N = a)}{\lambda'(y_N = a)} \\
&= \lambda'(y_{N+1} = N+1 | y_N = a).
\end{aligned} \tag{3.6}
$$

Now for the inductive step. Take any $N \in \mathbb{N}$ and any $\{a_M\}_{M=1}^{N+1} \in \mathcal{I}_{N+1}$. If $a_N = a_{N+1}$, then we use (3.5) and the inductive hypothesis to compute:

$$
\begin{aligned}
&\lambda(y_1 = a_1, \ldots, y_N = a_N, y_{N+1} = a_N) \\
&= \lambda(y_{N+1} = a_N \,|\, y_N = a_N) \lambda(y_1 = a_1, \ldots y_N = a_N) \\
&= \lambda'(y_{N+1} = a_N \,|\, y_N = a_N) \lambda'(y_1 = a_1, \ldots y_N = a_N) \\
&= \lambda'(y_1 = a_1, \ldots, y_N = a_N, y_{N+1} = a_N).
\end{aligned}
$$

Otherwise we have $a_{N+1} = N + 1$, so we use (3.6) and the inductive hypothesis to compute:

$$
\begin{aligned}
&\lambda(y_1 = a_1, \ldots, y_N = a_N, y_{N+1} = N+1) \\
&= \lambda(y_{N+1} = N+1 \,|\, y_N = a_N) \lambda(y_1 = a_1, \ldots, y_N = a_N) \\
&= \lambda'(y_{N+1} = N+1 \,|\, y_N = a_N) \lambda'(y_1 = a_1, \ldots, y_N = a_N) \\
&= \lambda'(y_1 = a_1, \ldots, y_N = a_N, y_{N+1} = N+1).
\end{aligned}
$$

This completes the induction and shows that λ and λ' agree on $\mathcal{P}_N$, hence on all of $\mathcal{P}$, hence on all of $\mathcal{B}(\mathcal{I})$. Therefore, $\lambda = \lambda'$, so ϕ is injective.

The next step is to show that ϕ is surjective. Take any $\boldsymbol{\nu} = \{\nu_N\}_{N \in \mathbb{N}} \in \boldsymbol{\mathcal{D}}_1$, and let us construct a $\lambda \in \boldsymbol{\mathcal{D}}_0$ such that $\phi(\lambda) = \boldsymbol{\nu}$. To do this , we recursively define a sequence $\{\lambda_N\}_{N \in \mathbb{N}}$ with $\lambda_N \in \mathcal{M}_1(\prod_{M=1}^{N}[\![0, M]\!])$ for each $N \in \mathbb{N}$. To start, we set $\lambda_1 = \nu_1$ on $\mathcal{M}_1([\![0, 1]\!])$. Then, for $N \in \mathbb{N}$ define $\lambda_{N+1} \in \mathcal{M}_1(\prod_{M=1}^{N+1}[\![0, M]\!])$ via

$$\lambda_{N+1}(y_1 = a_1, \ldots, y_{N+1} = a_{N+1})$$

$$= \begin{cases} \frac{\nu_{N+1}(a_N)}{\nu_N(a_N)} \lambda_N(y_1 = a_1, \ldots, y_N = a_N), & \text{if } a_{N+1} = a_N, \\ \frac{\nu_N(a_N) - \nu_{N+1}(a_N)}{\nu_N(a_N)} \lambda_N(y_1 = a_1, \ldots, y_N = a_N), & \text{if } a_{N+1} = N+1, \\ 0, & \text{if } a_{N+1} \notin \{a_N, N+1\}. \end{cases}$$

Note that the definition of $\boldsymbol{\mathcal{D}}_1$ guarantees $\nu_N(a_N) - \nu_{N+1}(a_N) \geq 0$, hence these are well-defined probabilities. We claim that $\{\lambda_N\}_{N \in \mathbb{N}}$ is projective, that is, that we have $\lambda_{N+1}(\cdot \times [\![0, N+1]\!]) = \lambda_N(\cdot)$ as measures on $\prod_{M=1}^{N}[\![0, M]\!]$; since $\prod_{M=1}^{N}[\![0, M]\!]$ is a finite set, it suffices to check that the two measures assign the same value to each point. For $N = 1$ this follows since for any $a_1 \in [\![0, 1]\!]$ we have

$$\lambda_2(y_1 = a_1) = \lambda_2(y_1 = a_1, y_2 = a_1) + \lambda_2(y_1 = a_1, y_2 = 2)$$

$$= \frac{\nu_2(a_1)}{\nu_1(a_1)} \nu_1(a_1) + \frac{\nu_1(a_1) - \nu_2(a_1)}{\nu_1(a_1)} \nu_1(a_1) = \nu_1(a_1)$$

by construction. For general $N \in \mathbb{N}$, this follows since for any $\{a_M\}_{M=1}^{N} \in \prod_{M=1}^{N}[\![0, M]\!] \setminus \boldsymbol{\mathcal{I}}_N$ we have

$$\lambda_{N+1}(y_1 = a_1, \ldots, y_N = a_N)$$
$$= \lambda_N(y_1 = a_1, \ldots, y_N = a_N) = 0.$$

Moreover, for any $N \in \mathbb{N}$ and $\{a_M\}_{M=1}^{N} \in \boldsymbol{\mathcal{I}}_N$ we have

$$\lambda_{N+1}(y_1 = a_1, \ldots, y_N = a_N)$$
$$= \lambda_{N+1}(y_1 = a_1, \ldots, y_N = a_N, y_{N+1} = a_{N+1})$$
$$\quad + \lambda_{N+1}(y_1 = a_1, \ldots, y_N = a_N, y_{N+1} = N+1)$$
$$= \frac{\nu_{N+1}(a_N)}{\nu_N(a_N)} \lambda_N(y_1 = a_1, \ldots, y_N = a_N)$$
$$\quad + \frac{\nu_N(a_N) - \nu_{N+1}(a_N)}{\nu_N(a_N)} \lambda_N(y_1 = a_1, \ldots, y_N = a_N)$$
$$= \lambda_N(y_1 = a_1, \ldots, y_N = a_N).$$

Thus, by [**6**, Theorem 6.14], there is a measure λ on $\prod_{N \in \mathbb{N}}[\![0, N]\!]$ whose projection onto $\prod_{M=1}^{N}[\![0, M]\!]$ is λ_N for each $N \in \mathbb{N}$.

Next we claim that $\lambda(\boldsymbol{\mathcal{I}}) = 1$. To do this, we first prove $\lambda_N(\boldsymbol{\mathcal{I}}_N) = 1$ for each $N \in \mathbb{N}$. We proceed by induction, and we note that the base case of $N = 1$ is immediate by construction. For the inductive step, let $N \in \mathbb{N}$ be arbitrary, and take any $\{a_M\}_{M=1}^{N+1} \in \prod_{M=1}^{N+1}[\![0, M]\!] \setminus \boldsymbol{\mathcal{I}}_{N+1}$. If $a_{M+1} \notin \{a_M, M+1\}$ for some $M \in [\![0, N-1]\!]$, then

$$\lambda_{N+1}(y_1 = a_1, \ldots, y_{N+1} = a_{N+1}) \leq \lambda_{N+1}(y_1 = a_1, \ldots, y_{M+1} = a_{M+1})$$
$$= \lambda_{M+1}(y_1 = a_1, \ldots, y_{M+1} = a_{M+1}) = 0,$$

by projectivity and the induction hypothesis, and, if $a_{N+1} \notin \{a_N, N+1\}$, then

$$\lambda_{N+1}(y_1 = a_1, \ldots, y_{N+1} = a_{N+1}) = 0$$

by the definition of λ_{N+1}. This completes the induction. Then note that $\mathcal{I} = \bigcap_{N \in \mathbb{N}}(\mathcal{I}_N \times \prod_{M=N+1}^{\infty} [\![0, M]\!])$, so we have

$$\lambda(\mathcal{I}) = \lambda \left(\bigcap_{N \in \mathbb{N}} \left(\mathcal{I}_N \times \prod_{M=N+1}^{\infty} [\![0, M]\!] \right) \right)$$

$$= \lim_{N \to \infty} \lambda \left(\mathcal{I}_N \times \prod_{M=N+1}^{\infty} [\![0, M]\!] \right) = \lim_{N \to \infty} \boldsymbol{\lambda}_N(\mathcal{I}_N) = 1,$$

as needed.

Next, we claim that λ is Markovian. That is, we need to show

$$\lambda(y_{N+1} = a_{N+1} | y_1 = a_1, \ldots, y_N = a_N) = \lambda(y_{N+1} = a_{N+1} | y_N = a_N)$$

for all $N \in \mathbb{N}$ and $\{a_M\}_{M=1}^{N+1}$. If $a_{N+1} \notin \{a_N, N+1\}$, then this follows from $\lambda(\mathcal{I}) = 1$ since both sides are equal to zero. Thus, if we show the claim for $a_{N+1} = a_N$, then it automatically holds for $a_{N+1} = N + 1$ by taking complements. Therefore it suffices to show the desired identity for $a_{N+1} = a_N$, and, to do this, show that both sides are equal to $\nu_{N+1}(a_N)/\nu_N(a_N)$. On the one hand, we compute, using projectivity and the definition of $\{\lambda_N\}_{N \in \mathbb{N}}$:

$$\lambda(y_{N+1} = a_N \,|\, y_1 = a_1, \ldots, y_N = a_N)$$

$$= \frac{\lambda(y_1 = a_1, \ldots, y_N = a_N, y_{N+1} = a_N)}{\lambda(y_1 = a_1, \ldots, y_N = a_N)}$$

$$= \frac{\lambda_{N+1}(y_1 = a_1, \ldots, y_N = a_N, y_{N+1} = a_N)}{\lambda_N(y_1 = a_1, \ldots, y_N = a_N)} = \frac{\nu_{N+1}(a_N)}{\nu_N(a_N)}.$$

On the other hand, we compute, by summing over all $\{a'_M\}_{M=1}^{N}$ with $a'_N = a_N$ and using projectivity and the definition of $\{\lambda_N\}_{N \in \mathbb{N}}$:

$$\lambda(y_{N+1} = a_N \,|\, y_N = a_N)$$

$$= \frac{\lambda(y_N = a_N, y_{N+1} = a_N)}{\lambda(y_N = a_N)}$$

$$= \frac{\lambda_{N+1}(y_N = a_N, y_{N+1} = a_N)}{\lambda_N(y_N = a_N)}$$

$$= \frac{\sum \lambda_{N+1}(y_1 = a'_1, \ldots, y_N = a'_N, y_{N+1} = a_N)}{\sum \lambda_N(y_1 = a'_1, \ldots, y_N = a'_N)}$$

$$= \frac{\sum \frac{\nu_{N+1}(a_N)}{\nu_N(a_N)} \lambda_{N+1}(y_1 = a'_1, \ldots, y_N = a'_N)}{\sum \lambda_N(y_1 = a'_1, \ldots, y_N = a'_N)}$$

$$= \frac{\nu_{N+1}(a_N)}{\nu_N(a_N)}.$$

This shows that λ is indeed Markovian.

Finally, we claim that $\phi(\lambda) = \boldsymbol{\nu}$, that is, that

$$\lambda \left(\prod_{M=1}^{N-1} [\![0, M]\!] \times \cdot \times \prod_{M=N+1}^{\infty} [\![0, M]\!] \right) = \nu_N(\cdot)$$

holds as measures on $[\![0, N]\!]$ for each $N \in \mathbb{N}$. By projectivity, it suffices to show that

$$\lambda_N \left(\prod_{M=1}^{N-1} [\![0, M]\!] \times \cdot \right) = \nu_N(\cdot)$$

holds as measures on $[\![0, N]\!]$ for each $N \in \mathbb{N}$. Since this is a finite set, it suffices to check that both measures assign the same value to each element. We proceed by induction, and again we note that the base case of $N = 1$ is immediate. For the inductive step, let $N \in \mathbb{N}$ be arbitrary. Note by summing and taking complements that it further suffices to prove the claim only for $a_{N+1} \in [\![0, N]\!]$. To do this, we sum over all $\{a'_M\}_{M=1}^N \in \mathcal{I}_N$ with $a'_N = a_{N+1}$, and apply the definition of λ_{N+1} and the inductive hypothesis:

$$\begin{aligned}
&\lambda_{N+1}(y_{N+1} = a_{N+1}) \\
&= \sum \lambda_{N+1}(y_1 = a'_1, \ldots, y_N = a'_N, y_{N+1} = a_{N+1}) \\
&= \sum \frac{\nu_{N+1}(a_{N+1})}{\nu_N(a_{N+1})} \lambda_N(y_1 = a'_1, \ldots, y_{N-1} = a'_{N-1}, y_N = a_{N+1}) \\
&= \frac{\nu_{N+1}(a_{N+1})}{\nu_N(a_{N+1})} \sum \lambda_N(y_1 = a'_1, \ldots, y_{N-1} = a'_{N-1}, y_N = a_{N+1}) \\
&= \frac{\nu_{N+1}(a_{N+1})}{\nu_N(a_{N+1})} \lambda_N(y_N = a_{N+1}) \\
&= \frac{\nu_{N+1}(a_{N+1})}{\nu_N(a_{N+1})} \nu_N(a) \\
&= \nu_{N+1}(a_{N+1}).
\end{aligned}$$

This completes the induction and proves $\phi(\lambda) = \boldsymbol{\nu}$. Thus, ϕ is a bijection.

Finally, note that ϕ is obviously continuous. Moreover, $\mathcal{D}_0$ is compact and $\mathcal{D}_1$ is Hausdorff, so it is classical that ϕ, being a continuous bijection, is actually a homeomorphism. $\qquad\square$

Example 3.14. For $N \in \mathbb{N}$, write $\mathrm{U}[\![1, N]\!]$ for the uniform measure on $[\![1, N]\!]$. Then consider the sequence $\mathbf{U} = \{\mathrm{U}[\![1, N]\!]\}_{N \in \mathbb{N}}$. For any $N \in \mathbb{N}$ we clearly have $\mathrm{U}[\![1, N]\!](a) = \frac{1}{N} > \frac{1}{N+1} = \mathrm{U}[\![1, N+1]\!](a)$ when $a \in [\![1, N]\!]$, and also $\mathrm{U}[\![1, N]\!](0) = 0 = \mathrm{U}[\![1, N+1]\!](0)$. Thus we have $\mathbf{U} \in \mathcal{D}_1$, and an illustrative example is to exactly identify the measure $\phi^{-1}(\mathbf{U}) \in \mathcal{D}_0$.

To do this, we note that there is a simple correspondence between $\mathcal{I}$-valued random variables $\boldsymbol{a} = \{a_N\}_{N \in \mathbb{N}}$ and sequences of $\{0, 1\}$-valued random variables $\{H_N\}_{N \in \mathbb{N}}$, given by

$$H_{N+1} = \begin{cases} 0, & \text{if } a_{N+1} = a_N, \\ 1, & \text{if } a_{N+1} = N+1, \end{cases} \qquad a_{N+1} = \begin{cases} a_N, & \text{if } H_N = 0, \\ N+1, & \text{if } H_N = 1. \end{cases} \qquad (3.7)$$

We use this by letting $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which is defined a sequence $\{H_N\}_{N\in\mathbb{N}}$ of independent random variables with with $H_N = \mathrm{Ber}(\frac{1}{N})$ for all $N \in \mathbb{N}$; then let $\boldsymbol{a} = \{a_N\}_{N\in\mathbb{N}}$ be defined as above, and set $\lambda = \mathbb{P} \circ \boldsymbol{a}^{-1}$. Of course for any $N \in \mathbb{N}$ and $\{a_M\}_{M=1}^N \in \boldsymbol{\mathcal{I}}_N$, the values $\lambda(y_{N+1} = N+1 \,|\, y_1 = a_1, \dots, y_N = a_N)$ and $\lambda(y_{N+1} = N+1 | y_N = a_N)$ are both equal to $\mathbb{P}(H_{N+1}) = \frac{1}{N+1}$. Taking complements shows that $\lambda(y_{N+1} = a_N \,|\, y_1 = a_1, \dots, y_N = a_N)$ and $\lambda(y_{N+1} = a_N \,|\, y_N = a_N)$ are also equal. Hence we have $\lambda(y_{N+1} = a_{N+1} \,|\, y_1 = a_1, \dots, y_N = a_N)$ and $\lambda(y_{N+1} = a_{N+1} \,|\, y_N = a_N)$ for all $N \in \mathbb{N}$ and $\{a_M\}_{M=1}^{N+1} \in \boldsymbol{\mathcal{I}}_{N+1}$, so λ is Markovian.

We claim that $\phi(\lambda) = \{\nu_N\}_{N\in\mathbb{N}}$ equals $\mathbf{U} = \{\mathrm{U}[\![1, N]\!]\}_{N\in\mathbb{N}}$. To see this, we use induction on $N \in \mathbb{N}$, wherein the base case of $N = 1$ is clear. For the inductive step, let $N \in \mathbb{N}$ be arbitrary. Note that by summing and taking complements it suffices to show $\nu_{N+1}(a) = \frac{1}{N+1}$ for all $a \in [\![1, N+1]\!]$. Indeed, this follows by independence:

$$\begin{aligned}
\nu_{N+1}(j) &= \mathbb{P}(a_{N+1} = j) \\
&= \mathbb{P}(a_N = j, H_{N+1} = 0) \\
&= \mathbb{P}(a_N = j)\mathbb{P}(H_{N+1} = 0) = \frac{1}{N} \cdot \frac{N}{N+1} = \frac{1}{N+1}.
\end{aligned}$$

This completes the induction and proves $\phi(\lambda) = \mathbf{U}$. We call $\mathbf{U} = \{\mathrm{U}[\![1, N]\!]\}_{N\in\mathbb{N}} \in \boldsymbol{\mathcal{D}}$ the *virtual uniform measure*. In the following section we will study $\mathbf{U}$ and a collection of related VMCs more carefully.

3.3. Compatibility. As we have hinted at above, VDs are important since they will generalize the notion of initial distributions for VMCs. However, not all VDs can serve as an "initial distribution" for a VMC. The reason for this phenomenon is that, for a VMC $\boldsymbol{X} = \{X_N\}_{N\in\mathbb{N}}$, each chain X_N for $N \in \mathbb{N}$ progresses at its own rate; one way to inrerpret this is that VMCs "have no natural time-scale". Thus, for $i \in \mathbb{N}$, the slice of values $\{X_N(i)\}_{N\in\mathbb{N}}$ does not, in general, represent a collection of values at the same point in time. The exception is $i = 0$, since all of the chains agree that this is the initial time. Hence, at time $i = 0$ we require extra information about the coupling of the process at all levels. This leads us to study an important notion of compatibility between VDs and VTMs.

Definition 3.15. Take any $\boldsymbol{\nu} = \{\nu_N\}_{N\in\mathbb{N}} \in \boldsymbol{\mathcal{D}}_1$ and $\boldsymbol{K} = \{K_N\}_{N\in\mathbb{N}} \in \boldsymbol{\mathcal{K}}$. We say that $\boldsymbol{\nu}$ is *compatible with* $\boldsymbol{K}$, or that the pair $(\boldsymbol{\nu}, \boldsymbol{K})$ is *compatible*, if we have

$$\nu_N(a) = \begin{cases} \nu_{N+1}(a) + \nu_{N+1}(N+1)\frac{K_{N+1}(N+1,a)}{1-K_{N+1}(N+1,N+1)}, & \text{if } K_{N+1}(N+1, N+1) < 1, \\ \nu_{N+1}(0) + \nu_{N+1}(N+1), & \text{else if } a = 0, \\ \nu_{N+1}(a), & \text{else if } a \neq 0, \end{cases}$$

(3.8)

for all $N \in \mathbb{N}$ and $a \in [\![0, N]\!]$.

Remark 3.16. As in the case of the projection operation for transition matrices, we can make the notion of compatibility slightly more concrete by viewing the elements of $\mathcal{K}$ as suitable block matrices. Again let us for each $N \in \mathbb{N}$ identify

$K_{N+1} \in \mathcal{K}_{N+1}$ with the block matrix

$$
\begin{pmatrix}
1 & 0 & 0 \\
w & A & u \\
q & v^{\mathrm{T}} & p
\end{pmatrix},
\tag{3.3}
$$

for $A \in [0,1]^{N \times N}, u, v, w \in [0,1]^N$, and $p, q \in [0,1]$, and let us also identify ν_{N+1} with the column vector

$$
\begin{pmatrix}
z \\
r
\end{pmatrix}
\tag{3.9}
$$

for $z \in [0,1]^{N+1}$ and $r \in [0,1]$, Then compatibility says that the column vector $\nu_N \in [0,1]^{N+1}$ must satisfy the identity

$$
\nu_N = \begin{cases}
z + \frac{r}{1-p} \begin{pmatrix} q \\ v \end{pmatrix}, & \text{if } p < 1, \\[2ex]
z + \begin{pmatrix} r \\ 0 \end{pmatrix}, & \text{if } p = 1.
\end{cases}
$$

The goal of the next few results is to show that every VMC is uniquely described by a compatible pair in $\mathcal{D}_1 \times \mathcal{K}$. The first direction is simple:

Lemma 3.17. *Suppose $\boldsymbol{X} = \{X_N\}_{N \in \mathbb{N}}$ is a VMC on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then, the sequence $\boldsymbol{\nu} = \{\mathbb{P} \circ X_N(0)^{-1}\}_{N \in \mathbb{N}}$ is a VD, the sequence $\boldsymbol{K} = \{K_N\}_{N \in \mathbb{N}}$ defined via*

$$
K_N(a, b) = \begin{cases}
\mathbb{P}(X_N(1) = b \mid X_N(0) = a), & \text{if } a, b \in [\![0, N]\!], \\
1, & \text{else if } a = b, \\
0, & \text{else if } a \neq b,
\end{cases}
\tag{3.10}
$$

for $N \in \mathbb{N}$ is a VTM, and the pair $(\boldsymbol{\nu}, \boldsymbol{K})$ is compatible.

Proof. First let's show $\boldsymbol{\nu} \in \mathcal{D}_1$. To do this, let $N \in \mathbb{N}$ be arbitrary. Clearly, $\mathbb{P} \circ X_N(0)^{-1}$ is a probability measure on $[\![0, N]\!]$. Moreover, note that $X_{N+1}(0) < N + 1$ implies $I_{X_{N+1}, N}(0) = \inf\{i \in \mathbb{N}_0 : X_{N+1}(0) \in [\![0, N]\!]\} = 0$, hence $X_N(0) = X_{N+1}(I_{X_{N+1}, N}(0)) = X_{N+1}(0)$. Thus, for any $a \in [\![0, N]\!]$ we have

$$
\begin{aligned}
&\mathbb{P}(X_N(0) = a) \\
&= \mathbb{P}(X_N(0) = a, X_{N+1}(0) = a) + \mathbb{P}(X_N(0) = a, X_{N+1}(0) = N + 1) \\
&\geq \mathbb{P}(X_N(0) = a, X_{N+1}(0) = a) \\
&= \mathbb{P}(X_{N+1}(0) = a),
\end{aligned}
$$

as needed.

Next, we need to show $\boldsymbol{K} \in \mathcal{K}$. To do this, let $N \in \mathbb{N}$ be arbitrary, and note that K_N is clearly a transition matrix on $\mathbb{N}_0$ with state a absorbing for all $a \in \mathbb{N}_0 \setminus [\![0, N]\!]$. Thus, it only remains to check (3.2) for all $a, b \in [\![0, N]\!]$. That this holds for $a = 0$ is immediate, so we can assume $a \in [\![1, N]\!]$, and in this case we have $\{X_N(0) = a\} = \{I_{X_{N+1}, N}(0) < \infty, X_{N+1}(I_{X_{N+1}, N}(0)) = a\}$. Then, by the strong

Markov property, we have

$$
\begin{aligned}
K_N(a,b) &= \mathbb{P}(X_N(1) = b \,|\, X_N(0) = i) \\
&= \mathbb{P}((P_N(X_{N+1}))(1) = b \,|\, X_N(0) = i) \\
&= \mathbb{P}((P_N(X_{N+1}))(1) = b \,|\, I_{X_{N+1},N}(0) < \infty, X_N(I_{X_{N+1},N}(0)) = a) \\
&= \mathbb{P}((P_N(X_{N+1}))(1) = b \,|\, X_{N+1}(0) = a).
\end{aligned}
$$

Now consider some cases. If $K_{N+1}(N+1, N+1) < 1$, then we of course have $\mathbb{P}(I_{X_{N+1},N}(1) < \infty | X_{N+1}(0) = a) = 1$, hence

$$
\begin{aligned}
K_N(a,b) &= \mathbb{P}((P_N(X_{N+1}))(1) = j \,|\, X_{N+1}(0) = a) \\
&= \mathbb{P}(I_{X_{N+1},N}(1) < \infty, X_{N+1}(I_{X_{N+1},N}(1)) = b \,|\, X_{N+1}(0) = a) \\
&= \mathbb{P}(X_{N+1}(I_{X_{N+1},N}(1)) = b \,|\, X_{N+1}(0) = a) \\
&= K_{N+1}(a,b) + K_{N+1}(a, N+1) \sum_{k=1}^{\infty} (K_{N+1}(N+1, N+1))^{k-1} K_{N+1}(N+1, b) \\
&= K_{N+1}(a,b) + \frac{K_{N+1}(a, N+1) K_{N+1}(N+1, b)}{1 - K_{N+1}(N+1, N+1)}.
\end{aligned}
$$

If instead $K_{N+1}(N+1, N+1) = 1$, then we consider the value of $b \in [\![0, N]\!]$. If $b = 0$ then

$$
\begin{aligned}
K_N(a,0) &= \mathbb{P}((P_N(X_{N+1}))(1) = 0 \,|\, X_{N+1}(0) = a) \\
&= \mathbb{P}(I_{X_{N+1},N}(1) < \infty, X_{N+1}(I_{X_{N+1},N}(1)) = 0 \,|\, X_{N+1}(0) = a) \\
&\quad + \mathbb{P}(I_{X_{N+1},N}(1) = \infty \,|\, X_{N+1}(0) = a) \\
&= K_{N+1}(a,0) + K_{N+1}(a, N+1),
\end{aligned}
$$

and if $b \neq 0$ then

$$
\begin{aligned}
K_N(a,b) &= \mathbb{P}((P_N(X_{N+1}))(1) = b \,|\, X_{N+1}(0) = b) \\
&= \mathbb{P}(I_{X_{N+1},N}(1) < \infty, X_{N+1}(I_{X_{N+1},N}(1)) = b \,|\, X_{N+1}(0) = a) \\
&= K_{N+1}(a,b).
\end{aligned}
$$

Therefore, we have $\boldsymbol{K} \in \mathcal{K}$.

Finally, we need to check that $\boldsymbol{\nu}$ is compatible with $\boldsymbol{K}$, which amounts to checking (3.8) for all $N \in \mathbb{N}$ and $a \in [\![0, N]\!]$. To do this, write

$$
\mathbb{P}(X_N(0) = a)
$$
$$
= \mathbb{P}(X_{N+1}(0) = a) + \mathbb{P}(X_N(0) = a \,|\, X_{N+1}(0) = N+1)\mathbb{P}(X_{N+1}(0) = N+1),
$$

and let us compute $\mathbb{P}(X_N(0) = a \,|\, X_{N+1}(0) = N+1)$ by considering the necessary cases. If $K_{N+1}(N+1, N+1) = 1$, then we clearly have

$$
\mathbb{P}(X_N(0) = a \,|\, X_{N+1}(0) = N+1) = \begin{cases} 1, & \text{if } a = 0, \\ 0, & \text{if } a \neq 0, \end{cases}
$$

hence

$$
\mathbb{P}(X_N(0) = j) = \begin{cases} \mathbb{P}(X_{N+1}(0) = 0) + \mathbb{P}(X_{N+1}(0) = N+1), & \text{if } a = 0, \\ \mathbb{P}(X_{N+1}(0) = a), & \text{if } a \neq 0. \end{cases}
$$

Otherwise we have $K_{N+1}(N+1, N+1) < 1$, so we can compute

$$\mathbb{P}(X_N(0) = a \,|\, X_{N+1}(0) = N+1)$$
$$= \mathbb{P}(I_{X_{N+1},X}(0) < \infty, X_{N+1}(I_{X_{N+1},X}(0)) = a \,|\, X_{N+1}(0) = N+1)$$
$$= \sum_{k=0}^{\infty} (K_{N+1}(N+1, N+1))^k K_{N+1}(N+1, a)$$
$$= \frac{K_{N+1}(N+1, a)}{1 - K_{N+1}(N+1, N+1)},$$

hence

$$\mathbb{P}(X_N(0) = a)$$
$$= \mathbb{P}(X_{N+1}(0) = a) + \mathbb{P}(X_{N+1}(0) = N+1)\frac{K_{N+1}(N+1, a)}{1 - K_{N+1}(N+1, N+1)},$$

as needed. This finishes the proof of the result. $\qquad\square$

Definition 3.18. In the setting of Lemma 3.17, we say that $\boldsymbol{\nu}$ and $\boldsymbol{K}$ are the *virtual initial distribution (VID)* and VTM of $\boldsymbol{X}$, respectively. Collectively, we say that $(\boldsymbol{\nu}, \boldsymbol{K})$ are the *canonical data* of $\boldsymbol{X}$.

Remark 3.19. It would be slightly more precise to call $\boldsymbol{\nu}$ the *initial virtual distritbution* of $\boldsymbol{X}$, but we with will use the slightly imprecise term because it sounds more natural (cf. Remark 2.15).

The next goal is to establish the converse of Lemma 3.17, namely that for any compatible pair $(\boldsymbol{\nu}, \boldsymbol{K}) \in \boldsymbol{\mathcal{D}}_1 \times \boldsymbol{\mathcal{K}}$ there is a VMC on some probability space with these as its canonical data.

Proposition 3.20. *If $(\boldsymbol{\nu}, \boldsymbol{K}) \in \boldsymbol{\mathcal{D}}_1 \times \boldsymbol{\mathcal{K}}$ is any compatible pair, then there is a VMC with canonical data $(\boldsymbol{\nu}, \boldsymbol{K})$.*

Proof. Write $\boldsymbol{\nu} = \{\nu_N\}_{N \in \mathbb{N}}$ and $\boldsymbol{K} = \{K_N\}_{N \in \mathbb{N}}$. Then, for each $N \in \mathbb{N}$, let $(\Omega_N, \mathcal{F}_N, \mathbb{P}_N)$ be a probability space on which is defined a MC $X_N = \{X_N(i)\}_{i \in \mathbb{N}_0}$ on $[\![0, N]\!]$ with initial distribution ν_N and transition matrix K_N. Let μ_N denote the law of $\{P_M(X_N)\}_{M=1}^{N}$ on the product space $\prod_{M=1}^{N} \mathcal{C}_M$.

Now let us introduce some notation. For $N \in \mathbb{N}$, write

$$\boldsymbol{\mathcal{C}}_N := \left\{ \{x_M\}_{M=1}^{N} \in \prod_{M=1}^{N} \mathcal{C}_M : P_M(x_{M+1}) = x_M \text{ for } 1 \le M < N \right\},$$

and write $\mathcal{B}(\boldsymbol{\mathcal{C}}_N)$ for the Borel σ-algebra on $\boldsymbol{\mathcal{C}}_N$. Next, for $\ell \in \mathbb{N}$ and $a_0, \dots a_\ell \in \mathbb{N}$, write

$$E_N(a_0, \dots a_\ell) := \left\{ \{x_M\}_{M=1}^{N} \in \boldsymbol{\mathcal{C}}_N : x_N(0) = a_0, \dots, x_N(\ell) = a_\ell \right\},$$

and define

$$\mathcal{P}_N := \{E_N(a_0, \dots a_\ell) : \ell \in \mathbb{N}, a_0, \dots, a_\ell \in [\![0, N]\!]\}.$$

Observe that $\mathcal{P}_N$ is a π-system with $\sigma(\mathcal{P}_N) = \mathcal{B}(\boldsymbol{\mathcal{C}}_N)$. Also, $\mathcal{B}(\prod_{M=1}^{N} \mathcal{C}_M \setminus \boldsymbol{\mathcal{C}}_N)$ is a π-system (in fact, a σ-algebra), and $\mathcal{P}_N$ and $\mathcal{B}(\prod_{M=1}^{N} \mathcal{C}_M \setminus \boldsymbol{\mathcal{C}}_N)$ are disjoint apart from that they both contain the empty set. Therefore, $\mathcal{P}_N \cup \mathcal{B}(\prod_{M=1}^{N} \mathcal{C}_M \setminus \boldsymbol{\mathcal{C}}_N)$ is a π-system which generates $\mathcal{B}(\prod_{M=1}^{N} \mathcal{C}_M)$

Now we claim that $\{\mu_N\}_{N\in\mathbb{N}}$ is projective, that is, that $\mu_{N+1}(\cdot \times \mathcal{C}_{N+1}) = \mu_N(\cdot)$ on $\mathcal{B}(\prod_{M=1}^N \mathcal{C}_M)$ for all $N \in \mathbb{N}$. By the previous paragraph, it is enough to check this for sets in $\mathcal{P}_N \cup \mathcal{B}(\prod_{M=1}^N \mathcal{C}_M \setminus \boldsymbol{\mathcal{C}}_N)$; we note that $A \in \mathcal{B}(\prod_{M=1}^N \mathcal{C}_M \setminus \boldsymbol{\mathcal{C}}_N)$ has $\mu_{N+1}(A \times \mathcal{C}_{N+1}) = 0 = \mu_N(A)$, so it only remains to check this for sets in $\mathcal{P}_N$. More precisely, we need to show that for $N \in \mathbb{N}, \ell \in \mathbb{N}_0$ and $a_0, \ldots a_\ell \in [\![0, N]\!]$, we have

$$
\begin{aligned}
&\mu_{N+1}(E_N(a_0, \ldots, a_\ell) \times \mathcal{C}_{N+1})\\
&= \mathbb{P}_{N+1}((P_N(X_{N+1}))(0) = a_0, \ldots, (P_N(X_{N+1}))(\ell) = a_\ell)\\
&= \mathbb{P}_N(X_N(0) = a_0, \ldots, X_N(\ell) = a_\ell)\\
&= \mu_N(E_N(a_0, \ldots, a_\ell)).
\end{aligned}
$$

Since 0 is an absorbing state for both X_N and $P_N(X_{N+1})$, it suffices to assume that $a_0, \ldots, a_{\ell-1} \in [\![1, N]\!]$ and $a_\ell \in [\![0, N]\!]$, since otherwise both sides above are equal to zero.

To do this, we need to check a few different cases, and, in each case, we simply apply the definitions compatibility and of VTM. If $K_{N+1}(N+1, N+1) < 1$, then we have

$$
\begin{aligned}
&\mathbb{P}_{N+1}((P_N(X_{N+1}))(0) = a_0, \ldots, (P_N(X_{N+1}))(\ell) = a_\ell)\\
&= \left(\nu_{N+1}(a_0) + \nu_{N+1}(N+1) \frac{K_{N+1}(N+1, a_0)}{1 - K_{N+1}(N+1, N+1)} \right)\\
&\qquad \times \prod_{k=1}^{\ell} \left(K_{N+1}(a_{k-1}, a_k) + \frac{K_{N+1}(a_{k-1}, N+1) K_{N+1}(N+1, a_k)}{1 - K_{N+1}(N+1, N+1)} \right)\\
&= \nu_N(a_0) \prod_{k=1}^{\ell} K_N(a_{k-1}, a_k)\\
&= \mathbb{P}_N(X_N(0) = a_0, \ldots, X_N(\ell) = a_\ell),
\end{aligned}
$$

as needed. Otherwise $K_{N+1}(N+1, N+1) = 1$, and we need to split into a few more cases. First suppose $\ell = 0$, and let us further consider the value of a_0. If $a_0 = 0$, then

$$
\begin{aligned}
\mathbb{P}_{N+1}((P_N(X_{N+1}))(0) = 0) &= \nu_{N+1}(0) + \nu_{N+1}(N+1)\\
&= \nu_N(0) = \mathbb{P}_N(X_N(0) = 0),
\end{aligned}
$$

and, if $a_0 \neq 0$, then

$$
\begin{aligned}
\mathbb{P}_{N+1}((P_N(X_{N+1}))(0) = a_0) &= \nu_{N+1}(a_0)\\
&= \nu_N(a_0) = \mathbb{P}_N(X_N(0) = a_0),
\end{aligned}
$$

as needed. Now suppose $\ell \geq 1$, and let us further consider the value of a_ℓ. If $a_\ell = 0$, then

$$\mathbb{P}_{N+1}((P_N(X_{N+1}))(0) = a_0, \ldots, (P_N(X_{N+1}))(\ell - 1) = a_{\ell-1}, (P_N(X_{N+1}))(\ell) = 0)$$

$$= \nu_{N+1}(a_0) \prod_{k=1}^{\ell-1} K_{N+1}(a_{k-1}, a_k)(K_{N+1}(a_{\ell-1}, 0) + K_{N+1}(a_{\ell-1}, N + 1))$$

$$= \nu_N(a_0) \prod_{k=1}^{\ell-1} K_N(a_{k-1}, a_k) K_N(a_{\ell-1}, 0)$$

$$= \mathbb{P}_N(X_N(0) = a_0, \ldots, X_N(\ell - 1) = a_{\ell-1}, X_N(\ell) = 0),$$

and, if $a_\ell \neq 0$, then

$$\mathbb{P}_{N+1}((P_N(X_{N+1}))(0) = a_0, \ldots, (P_N(X_{N+1}))(\ell) = a_\ell)$$

$$= \nu_{N+1}(a_0) \prod_{k=1}^{\ell-1} K_{N+1}(a_{k-1}, a_k) K_{N+1}(a_{\ell-1}, a_\ell)$$

$$= \nu_N(a_0) \prod_{k=1}^{\ell-1} K_N(a_{k-1}, a_k) K_N(a_{\ell-1}, 0)$$

$$= \mathbb{P}_N(X_N(0) = a_0, \ldots, X_N(\ell) = a_\ell),$$

as needed. This shows that $\{\mu_N\}_{N \in \mathbb{N}}$ is indeed projective. Thus, by [**6**, Theorem 6.14], there exists a probability measure μ on $\prod_{N \in \mathbb{N}} \mathcal{C}_N$ with projections onto $\prod_{M=1}^{N} \mathcal{C}_M$ given by μ_N for all $N \in \mathbb{N}$.

Now note that, since $\mu_N(\mathcal{C}_N) = 1$ for all $N \in \mathbb{N}$ and $\mathcal{C} = \bigcap_{N \in \mathbb{N}}(\mathcal{C}_N \times \prod_{M=N+1}^{\infty} \mathcal{C}_M)$, we have

$$\mu(\mathcal{C}) = \mu \left(\bigcap_{N \in \mathbb{N}} \left(\mathcal{C}_N \times \prod_{M=N+1}^{\infty} \mathcal{C}_M \right) \right)$$

$$= \lim_{N \to \infty} \mu \left(\mathcal{C}_N \times \prod_{M=N+1}^{\infty} \mathcal{C}_M \right) = \lim_{N \to \infty} \mu_N(\mathcal{C}_N) = 1.$$

In other words, μ is the law of a $\mathcal{C}$-valued random variable, and, by construction, its marginal distribution on $\mathcal{C}_N$ is the law of a MC with initial distribution ν_N and transition matrix K_N, for each $N \in \mathbb{N}$. This proves that μ is a VMC with canonical data $(\boldsymbol{\nu}, \boldsymbol{K})$, as claimed. $\square$

Finally, we show that the VMC guaranteed by Proposition 3.20 is unique in law.

Proposition 3.21. *Suppose that* $(\Omega, \mathcal{F}, \mathbb{P})$ *and* $(\Omega', \mathcal{F}', \mathbb{P}')$ *are probability spaces on which are defined VMCs* $\boldsymbol{X}$ *and* $\boldsymbol{X}'$, *respectively, with canonical data* $(\boldsymbol{\nu}, \boldsymbol{K})$ *and* $(\boldsymbol{\nu}', \boldsymbol{K}')$, *respectively. If* $(\boldsymbol{\nu}, \boldsymbol{K}) = (\boldsymbol{\nu}', \boldsymbol{K}')$, *then we have* $\mathbb{P} \circ \boldsymbol{X}^{-1} = \mathbb{P}' \circ (\boldsymbol{X}')^{-1}$ *as measures on* $\mathcal{C}$.

Proof. For $M \in \mathbb{N}$ and $B_1 \in \mathcal{B}(\mathcal{C}_1), \ldots, B_M \in \mathcal{B}(\mathcal{C}_M)$, define

$$E_M(B_1, \ldots, B_M) := \{\{x_N\}_{N \in \mathbb{N}} \in \mathcal{C} : x_1 \in B_1, \ldots, x_M \in B_M\}$$

and also

$$\mathcal{P}_M := \{E_M(B_1, \ldots, B_M) : B_1 \in \mathcal{B}(\mathcal{C}_1), \ldots, B_M \in \mathcal{B}(\mathcal{C}_M)\}.$$

Then note that $\mathcal{P} := \bigcup_{M \in \mathbb{N}} \mathcal{P}_M$ is a π-system with $\sigma(\mathcal{P}) = \mathcal{B}(\boldsymbol{\mathcal{C}})$. Hence, it suffices to show that $\mathbb{P} \circ \boldsymbol{X}^{-1}$ and $\mathbb{P}' \circ (\boldsymbol{X}')^{-1}$ agree for all sets in $\mathcal{P}$.

To do this, write $\boldsymbol{X} = \{X_N\}_{N \in \mathbb{N}}$, and recall that, on $\mathrm{supp}(\mathbb{P} \circ \boldsymbol{X}^{-1})$, the random variables $\{X_N\}_{N \in [\![1, M-1]\!]}$ are a measurable function of the random variable X_M, for all $M \in \mathbb{N}$; the analogous claim is of course also true for $\boldsymbol{X}' = \{X'_N\}_{N \in \mathbb{N}}$. Also write $\boldsymbol{\nu} = \{\nu_N\}_{N \in \mathbb{N}}$ and $\boldsymbol{\nu}' = \{\nu'_N\}_{N \in \mathbb{N}}$, well as $\boldsymbol{K} = \{K_N\}_{N \in \mathbb{N}}$ and $\boldsymbol{K}' = \{K'_N\}_{N \in \mathbb{N}}$. Then take any $N \in \mathbb{N}$ and $B_1 \in \mathcal{B}(\mathcal{C}_1), \ldots, B_N \in \mathcal{B}(\mathcal{C}_N)$, and note that the claim is immediate if $E_N(B_1, \ldots, B_N)$ is empty. Otherwise, we have

$$
\begin{aligned}
(\mathbb{P} \circ \boldsymbol{X}^{-1})(E_N(B_1, \ldots, B_N)) &= \mathbb{P}(X_1 \in B_1, \ldots, X_N \in B_N) \\
&= \mathbb{P}(X_N \in B_N) \\
&= \mathbb{P}'(X'_N \in B_N) \\
&= \mathbb{P}'(X'_1 \in B_1, \ldots, X'_N \in B_N) \\
&= (\mathbb{P}' \circ (\boldsymbol{X}')^{-1})(E_N(B_1, \ldots, B_N)),
\end{aligned}
$$

where $\mathbb{P}(X_N \in B_N) = \mathbb{P}'(X'_N \in B_N)$ follows from the classical result applied to $\nu_N = \nu'_N$ and $K_N = K'_N$. This finishes the proof. $\qquad \square$

Summarizing the partial results shown so far, we come to the main result of this section, which provides a representation theorem describing that each VMC is uniquely described by a compatible pair of a VD and a VTM.

Theorem 3.22. *There is a homeomorphism from the space of laws of VMCs to the space of compatible pairs of VID and VTM given by sending each law of a VMC to its canonical data.*

Proof. Note that Lemma 3.17 implies that there is a map

$$
\Phi : \boldsymbol{\mathcal{M}} \to \{(\boldsymbol{\nu}, \boldsymbol{K}) \in \boldsymbol{\mathcal{D}}_1 \times \boldsymbol{\mathcal{K}} : (\boldsymbol{\nu}, \boldsymbol{K}) \text{ is compatible}\}
$$

which sends each law of a VMC to its canonical data. Then, Proposition 3.20 implies that Φ is surjective and Proposition 3.21 implies that Φ is injective. Now we apply the usual trick: The domain of Φ is compact by Lemma 2.18, and the range of Φ is Hausdorff (it is a subset of a metrizable space, hence metrizable itself). Moreover, Φ is clearly continuous, and a continuous bijection with these properties is automatically a homeomorphism. $\qquad \square$

3.4. Examples. Finally, let us see some illustrative examples of the theory just developed. To do this we introduce one last notational convention: In most interesting examples of VMCs, the VTM $\boldsymbol{K} = \{K_N\}_{N \in \mathbb{N}}$ has $K_N(a, 0) = 0$ for all $N \in \mathbb{N}$ and $a \in [\![1, N]\!]$, so displaying the zeroth row and column does not add any information. So, in order to compactify the presentation, we will use matrices with parentheses as boundaries to denote TMs including the zeroth row and column and matrices with brackets as boundaries to denote TMs excluding the zeroth row and column; in LaTeX, these correspond to the `pmatrix` and `bmatrix` environments, respectively.

Example 3.23. Let $\boldsymbol{\sigma} = \{\sigma_N\}_{N \in \mathbb{N}} \in \boldsymbol{\mathcal{S}}$ be a virtual permutation and take any $a \in \mathbb{N}$, and let $\boldsymbol{X}$ be the VMC on $(\Omega, \mathcal{F}, \mathbb{P}_{L_a})$ constructed in Example 2.22. Its VTM $\boldsymbol{K} = \{K_N\}_{N \in \mathbb{N}}$ is a *virtual permutation matrix* in the sense that K_N is a permutation matrix for each $N \in \mathbb{N}$, and its VID is exactly $\boldsymbol{\nu} = \{\delta_{a_N}\}_{N \in \mathbb{N}}$ for a

suitable sequence $\{a_N\}_{N\in\mathbb{N}} \in \prod_{N\in\mathbb{N}}[\![0,N]\!]$. We write $\boldsymbol{K}(\boldsymbol{\sigma})$ for the VTM of this VMC, which, in particular, does not depend on $a \in \mathbb{N}_0$.

Example 3.24. Consider the VMC $\boldsymbol{X}$ on $(\Omega,\mathcal{F},\mathbb{P})$ constructed via Example 2.23, where F and $\{L_N\}_{N\in\mathbb{N}}$ are left general. We define $q_N = (1-F(L_N))/(1-F(L_{N+1}))$ for $N \in \mathbb{N}$, and then we note that its VTM $\boldsymbol{K} = \{K_N\}_{N\in\mathbb{N}}$ is exactly given by

$$K_N = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \\ q_1 & 0 & 0 & 0 & \cdots & 0 & 0 & 1-q_1 \\ 0 & q_2 & 0 & 0 & \cdots & 0 & 0 & 1-q_2 \\ 0 & 0 & q_3 & 0 & \cdots & 0 & 0 & 1-q_3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1-q_{N-3} \\ 0 & 0 & 0 & 0 & \cdots & q_{N-2} & 0 & 1-q_{N-2} \\ 0 & 0 & 0 & 0 & \cdots & 0 & q_{N-1} & 1-q_{N-1} \end{bmatrix}$$

for each $N \in \mathbb{N}$, and its VID $\boldsymbol{\nu}$ of $\boldsymbol{X}$ is just $\boldsymbol{\nu} = \{\delta_N\}_{N\in\mathbb{N}} \in \mathcal{D}_1$. Note also that $K = \lim_{N\to\infty} K_N$ does not exist in $\mathcal{K}$, hence by Lemma 3.6 there is no TM corresponding to this VTM.

Example 3.25. Consider the VMC $\boldsymbol{X}$ on $(\Omega,\mathcal{F},\mathbb{P}_0)$ constructed via Example 2.24. Its VTM $\boldsymbol{K} = \{K_N\}_{N\in\mathbb{N}}$ and VID $\boldsymbol{\nu} = \{\nu_N\}_{N\in\mathbb{N}}$ are given by

$$K_N = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1/2 & 1/2 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \nu_N = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 1/2 \\ 1/2 \end{bmatrix}$$

for all $N \in \mathbb{N}$. Note, in particular, that neither the VID not the VTM depend on the choice of $\{\ell_N\}_{N\in\mathbb{N}}$. Also observe $K = \lim_{N\to\infty} K_N$ does not exist in $\mathcal{K}$, hence by Lemma 3.6 there is no TM corresponding to this VTM.

Example 3.26. For each $N \in \mathbb{N}$, set

$$K_N = \begin{bmatrix} 1/2 & 1/2 & 0 & \cdots & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & \cdots & 0 & 0 & 0 \\ 0 & 1/2 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1/2 & 0 \\ 0 & 0 & 0 & \cdots & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & \cdots & 0 & 1/2 & 1/2 \end{bmatrix},$$

which is an element of $[0,1]^{(N+1)\times(N+1)}$. It is easy to verify that $\boldsymbol{K} = \{K_N\}_{N\in\mathbb{N}}$ is indeed a VTM. In fact, $\boldsymbol{K} = \iota(K)$, where

$$K = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 & 0 & \cdots \\ 1/2 & 0 & 1/2 & 0 & 0 & \cdots \\ 0 & 1/2 & 0 & 1/2 & 0 & \cdots \\ 0 & 0 & 1/2 & 0 & 1/2 & \cdots \\ 0 & 0 & 0 & 1/2 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

In other words, $\boldsymbol{K}$ is the VTM representing the simple symmetric random walk on $\mathbb{N}$ where the boundary state 1 holds and reflects with equal probability.

In particular, observe that the VID $\boldsymbol{\nu} = \{\delta_N\}_{N\in\mathbb{N}}$ is compatible with $\boldsymbol{K}$. It follows that there is a VMC $\boldsymbol{X} = \{X_N\}_{N\in\mathbb{N}}$ with canonical data $(\boldsymbol{\nu}, \boldsymbol{K})$, and it satisfies $X_N(0) = N$ for all $N \in \mathbb{N}$. Thus we have $\lim_{N\to\infty} X_N(0) = \infty$ hence $\lim_{N\to\infty} X_N$ does not exist in $\mathcal{C}$. Therefore, $\boldsymbol{X} \notin \iota(\mathcal{C})$ almost surely, by Lemma 2.9. Intuitively speaking, $\boldsymbol{X}$ is the simple random walk on $\mathbb{N}$ where the state 1 holds and reflects with equal probability, but "started from infinity". This example demonstrates that even if a VTM corresponds to a classical TM, a compatible VID can be such that the resulting VMC does not correspond to a classical MC.

4. Some Aspects of Convexity

In the classical setting, even though the law of a Markov chain (MC) is uniquely characterized by its transition matrix (TM) and initial distribution (ID), one usually regards the TM as containing the rich information and the ID as playing a secondary role. We adopt this perspective in the present context, where, although the law of a virtual Markov chain (VMC) is uniquely characterized by its virtual transition matrix (VTM) and virtual initial distribution (VID), we will regard the VTM as containing the important information and the VID as playing a secondary role.

This viewpoint naturally leads us to study a few different problems of convexity in infinite-dimensional spaces related to the study of VMCs. In Subsection 4.1, we consider the compact convex space of all VIDs that are compatible with a given VTM, and in Subsection 4.2 we consider two different compact convex spaces of VIDs that generalize the notions of equilibrium distribution and stationary distribution for classical MCs; in Subsection 4.3 we compute some examples of the preceding results. In Subsection 4.4 we aim to understand the convexity structure of the *virtual Birkhoff polytope*, and we show, in particular, that the classical Birkhoff-von Neumann theorem fails in the virtual setting.

4.1. Compatibility revisited. Motivated by Theorem 3.22, we are led to study the space of all VIDs that are compatible with a given VTM. As we will show in this subsection, this is a compact convex set whose extreme points can, in many cases, be concretely understood.

Definition 4.1. For $\boldsymbol{K} \in \mathcal{K}$, write $\boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$ for the set of elements of $\boldsymbol{\mathcal{D}}_1$ which are compatible with $\boldsymbol{K}$.

We also introduce some notation to simplify the requisite conditions of VTM and of compatibility. That is, for $\boldsymbol{K} = \{K_N\}_{N\in\mathbb{N}} \in \prod_{N\in\mathbb{N}} \mathcal{K}_N$, $N \in \mathbb{N}$, and $a \in [\![0,N]\!]$,

we define the constants

$$C_{N,a}^{\boldsymbol{K}} := \begin{cases} \frac{K_{N+1}(N+1,a)}{1-K_{N+1}(N+1,N+1)}, & \text{if } K_{N+1}(N+1,N+1) < 1, \\ 1, & \text{else if } a = 0, \\ 0, & \text{else if } a \neq 0. \end{cases}$$

Also define $C_{0,0}^{\boldsymbol{K}} = 0$ by convention. Then, observe that $\boldsymbol{K}$ is a VTM if and only if we have $K_N(a,b) = K_{N+1}(a,b) + K_{N+1}(a,N+1)C_{N,b}^{\boldsymbol{K}}$ for all $N \in \mathbb{N}$ and $a \in [\![0,N]\!]$ and that a VD $\boldsymbol{\nu} = \{\nu_N\}_{N \in \mathbb{N}} \in \boldsymbol{\mathcal{D}}_1$ is compatible with a VTM $\boldsymbol{K} \in \mathcal{K}$ iff we have $\nu_N(a) = \nu_{N+1}(a) + \nu_{N+1}(N+1)C_{N,a}^{\boldsymbol{K}}$ for all $N \in \mathbb{N}$ and $a \in [\![0,N]\!]$. Also note that row-stochasticity implies that we have

$$\sum_{a=0}^{N} C_{N,a}^{\boldsymbol{K}} = 1 \tag{4.1}$$

for any $N \in \mathbb{N}$.

Lemma 4.2. *For $\boldsymbol{K} \in \mathcal{K}$, the set $\boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$ is compact, and convex.*

Proof. Observe that we can write

$$\boldsymbol{\mathcal{D}}_1(\boldsymbol{K}) = \bigcap_{\substack{M \in \mathbb{N} \\ a \in [\![0,M]\!]}} \left\{ \{\nu_N\}_{N \in \mathbb{N}} \in \boldsymbol{\mathcal{D}}_1 : \nu_N(a) = \nu_{N+1}(a) + \nu_{N+1}(N+1)C_{N,a}^{\boldsymbol{K}} \right\}.$$

In other words, $\boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$ is defined as the solution set to countably many linear equations whose coefficients depend on $\boldsymbol{K}$. Of course this implies that $\boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$ is convex. Moreover, it implies that $\boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$ is closed in $\boldsymbol{\mathcal{D}}_1$, which, since $\boldsymbol{\mathcal{D}}_1$ is a compact space, implies that $\boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$ is a compact space. □

By Choquet's theorem, we know that every element of $\boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$ can be written as a mixture of elements of the extreme points of this set, denoted $\mathrm{ex}(\boldsymbol{\mathcal{D}}_1(\boldsymbol{K}))$. We are hence motivated to understand the structure of this set of extreme points.

Lemma 4.3. *Let $\boldsymbol{K} \in \mathcal{K}$ be a VTM and let $M \in \mathbb{N}_0$ be arbitrary. Then, there is a unique VID $\boldsymbol{\nu} = \{\nu_N\}_{N \in \mathbb{N}} \in \boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$ such that $\nu_N = \delta_M$ holds for all $N \in \mathbb{N}$ with $N \geq M$.*

Proof. For existence, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which is defined $X = \{X(i)\}_{i \in \mathbb{N}_0}$ a MC in $\mathbb{N}_0$ with TM K_M and ID δ_M. By Proposition 2.16 and Lemma 3.17, the VMC $\iota(X)$ has some canonical data $\boldsymbol{\nu} = \{\nu_N\}_{N \in \mathbb{N}} \in \boldsymbol{\mathcal{D}}_1$ and $\boldsymbol{K} = \{K_N\}_{N \in \mathbb{N}} \in \mathcal{K}$ with $(\boldsymbol{\nu}, \boldsymbol{K})$ compatible. Note that, for $N \geq M$, we have $P_N(X) = X$, hence

$$\nu_N(a) = \mathbb{P}((P_N(X)) = a) = \mathbb{P}(X = a) = \delta_M(a)$$

for all $a \in [\![0,N]\!]$, as desired. Also, by construction, we have $\boldsymbol{K} = \iota(K_M)$. Thus, it only remains to show that $(\boldsymbol{\nu}, \iota(K_M))$ being compatible implies that $(\boldsymbol{\nu}, \boldsymbol{K})$ is compatible. To do this, write $\iota(K_M) = \{K_N'\}_{N \in \mathbb{N}}$, and note that, for $N \in \mathbb{N}$, we have that $N \geq M$ implies $K_N' = K_M$ and that $N \leq M$ implies $K_N' = K_N$. Now let $N \in \mathbb{N}$ and $a \in [\![0,N]\!]$ be arbitrary. If $N \leq M$ then we have $C_{N,a}^{\boldsymbol{K}} = C_{N,a}^{\iota(K_M)}$, hence

$$\nu_N(a) = \nu_{N+1}(a) + \nu_{N+1}(N+1)C_{N,a}^{\iota(K_M)}$$
$$= \nu_{N+1}(a) + \nu_{N+1}(N+1)C_{N,a}^{\boldsymbol{K}}.$$

If instead $N \geq M$ then we have $\nu_{N+1}(N+1) = 0$, hence

$$\nu_N(a) = \delta_M(a) = \nu_{N+1}(a) + \nu_{N+1}(N+1)C_{N,a}^{\boldsymbol{K}}.$$

Therefore, existence is proved.

For uniqueness, take $\boldsymbol{\nu}, \boldsymbol{\nu}' \in \boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$ and write $\boldsymbol{\nu} = \{\nu_N\}_{N \in \mathbb{N}}$ and $\boldsymbol{\nu}' = \{\nu_N'\}_{N \in \mathbb{N}}$, and suppose that $\nu_N = \nu_N' = \delta_N$ holds for all $N \in \mathbb{N}$ with $N \geq M$. Then applying the compatibility relation $\nu_N(a) = \nu_{N+1}(a) + \nu_{N+1}(N+1)C_{N,a}^{\boldsymbol{K}}$ for all $N \in \mathbb{N}$ and $a \in [\![0, N]\!]$ inductively shows that we also have $\nu_N = \nu_N'$ for all $N \in \mathbb{N}$ with $N \leq M$. Thus $\nu_N = \nu_N'$ for all $N \in \mathbb{N}$, hence $\boldsymbol{\nu} = \boldsymbol{\nu}'$. $\qquad\square$

For any $\boldsymbol{K} \in \mathcal{K}$ and $M \in \mathbb{N}_0$, write $\delta_M^{\boldsymbol{K}}$ for the element of $\boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$ guaranteed by the result above. Observe also that, if $\boldsymbol{K}$ arises from the setting of Lemma 2.21, then we simply have $\delta_M^{\boldsymbol{K}} = \{\mathbb{P}_{L_M} \circ X_N(0)^{-1}\}_{N \in \mathbb{N}}$ for $M \in \mathbb{N}$. The next result confirms the intuition that $\delta_M^{\boldsymbol{K}}$ are extreme points of $\boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$ for all $M \in \mathbb{N}_0$.

Lemma 4.4. *For arbitrary $\boldsymbol{K} \in \mathcal{K}$, we have*

$$\mathrm{ex}(\boldsymbol{\mathcal{D}}_1(\boldsymbol{K})) \supseteq \{\delta_N^{\boldsymbol{K}} : N \in \mathbb{N}_0\}.$$

Proof. Take any $M \in \mathbb{N}$ and suppose that $\boldsymbol{\nu}, \boldsymbol{\nu}' \in \boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$ and $\alpha \in (0,1)$ are such that we have $(1-\alpha)\boldsymbol{\nu} + \alpha\boldsymbol{\nu}' = \delta_M^{\boldsymbol{K}}$. Now write $\boldsymbol{\nu} = \{\nu_N\}_{N \in \mathbb{N}}$ and $\boldsymbol{\nu}' = \{\nu_N'\}_{N \in \mathbb{N}}$. Note that for any $N \in \mathbb{N}$ with $N \geq M$ we have $(1-\alpha)\nu_N + \alpha\nu_N' = \delta_N$, hence $\nu_N = \nu_N' = \delta_N$. Thus by Lemma 4.3 we have $\boldsymbol{\nu} = \boldsymbol{\nu}' = \delta_N^{\boldsymbol{K}}$, as claimed. $\qquad\square$

For most interesting examples of $\boldsymbol{K} \in \mathcal{K}$, it is intuitive that $\boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$ should have more extreme points than just those included in $\{\delta_N^{\boldsymbol{K}} : N \in \mathbb{N}_0\}$. (Indeed, consider Example 3.24: The VMC $\boldsymbol{X}$ starts "from infinity", and this does not appear to be a mixture of any other starting states.) In order to get a more complete understanding of such extreme points of $\boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$, we introduce an additional characterization of VIDs.

Definition 4.5. For any $\boldsymbol{K} \in \mathcal{K}$, we set

$$\boldsymbol{\mathcal{D}}_2(\boldsymbol{K}) := \left\{\{p_a\}_{a \in \mathbb{N}_0} \in [0,1]^{\mathbb{N}_0} : p_a \geq \sum_{M=a}^{\infty} C_{M,a}^{\boldsymbol{K}} p_{M+1} \text{ for } a \in \mathbb{N}_0, \text{ and } p_0 = 1\right\}.$$

Lemma 4.6. *For any $\boldsymbol{K} \in \mathcal{K}$, the set $\boldsymbol{\mathcal{D}}_2(\boldsymbol{K})$ is compact and convex.*

Proof. Observe that $\boldsymbol{\mathcal{D}}_2(\boldsymbol{K})$ is just the solution set to countably many non-strict inequalities, hence it is closed in $[0,1]^{\mathbb{N}_0}$. But $[0,1]^{\mathbb{N}_0}$ is compact by Tychonoff, hence $\boldsymbol{\mathcal{D}}_2(\boldsymbol{K})$ is compact. Moreover, the inequalities are linear, so $\boldsymbol{\mathcal{D}}_2(\boldsymbol{K})$ is convex. $\quad\square$

Proposition 4.7. *For any $\boldsymbol{K} \in \mathcal{K}$, the map $\psi_{\boldsymbol{K}} : \boldsymbol{\mathcal{D}}_2(\boldsymbol{K}) \to \boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$ defined via $\psi_{\boldsymbol{K}}(\{p_a\}_{a \in \mathbb{N}_0}) = \{\nu_N\}_{N \in \mathbb{N}}$ with*

$$\nu_N(a) = p_a - \sum_{M=a}^{N-1} C_{M,a}^{\boldsymbol{K}} p_{M+1}$$

for $N \in \mathbb{N}$ and $a \in [\![0, N]\!]$ is a well-defined linear homeomorphism. Its inverse $\psi_{\boldsymbol{K}}^{-1} : \boldsymbol{\mathcal{D}}_1(\boldsymbol{K}) \to \boldsymbol{\mathcal{D}}_2(\boldsymbol{K})$ is given by $\psi_{\boldsymbol{K}}^{-1}(\{\nu_N\}_{N \in \mathbb{N}}) = (\{p_a\}_{a \in \mathbb{N}_0})$ with $p_0 = 1$ and $p_a = \nu_a(a)$ for $a \in \mathbb{N}$.

Proof. Let us show that $\psi_{\boldsymbol{K}}$ is well-defined. To do this, first take arbitrary $N \in \mathbb{N}$ and $a \in [\![0, N]\!]$, and note that, on the one hand, we have

$$\nu_N(a) = p_a - \sum_{M=a}^{N-1} C_{M,a}^{\boldsymbol{K}} p_{M+1} \geq p_a - \sum_{M=a}^{\infty} C_{M,a}^{\boldsymbol{K}} p_{M+1} \geq 0,$$

while, on the other hand, we have

$$\nu_N(a) = p_a - \sum_{M=a}^{N-1} C_{M,a}^{\boldsymbol{K}} p_{M+1} \leq p_a \leq 1.$$

This shows $0 \leq \nu_N(a) \leq 1$. Next note that, we have

$$\nu_N(a) = p_a - \sum_{M=a}^{N-1} C_{M,a}^{\boldsymbol{K}} p_{M+1} \geq p_a - \sum_{M=a}^{N} C_{M,a}^{\boldsymbol{K}} p_{M+1} = \nu_{N+1}(a),$$

hence $\nu_N(a) \geq \nu_{N+1}(a)$ as needed. Finally, note that for any $N \in \mathbb{N}$ we have, by switching the order of summation and applying (4.1):

$$\begin{aligned}
\sum_{a=0}^{N} \nu_N(a) &= \sum_{a=0}^{N} \left(p_a - \sum_{M=a}^{N-1} C_{M,a}^{\boldsymbol{K}} p_{M+1} \right) \\
&= \sum_{a=0}^{N} p_a - \sum_{a=0}^{N} \sum_{M=a}^{N-1} C_{M,a}^{\boldsymbol{K}} p_{M+1} \\
&= \sum_{a=0}^{N} p_a - \sum_{M=0}^{N-1} p_{M+1} \sum_{a=0}^{M} C_{M,a}^{\boldsymbol{K}} \\
&= \sum_{a=0}^{N} p_a - \sum_{M=0}^{N-1} p_{M+1} = p_0 = 1.
\end{aligned}$$

Thus, $\psi_{\boldsymbol{K}} : \boldsymbol{\mathcal{D}}_2(\boldsymbol{K}) \to \boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$ is indeed a well-defined map. It is also immediate from the definition that this map is linear.

Now let us show that $\psi_{\boldsymbol{K}}$ is injective. Suppose $\{p_a\}_{a \in \mathbb{N}_0}, \{p_a'\}_{a \in \mathbb{N}_0} \in \boldsymbol{\mathcal{D}}_2(\boldsymbol{K})$ have $\psi_{\boldsymbol{K}}(\{p_a\}_{a \in \mathbb{N}_0}) = \psi_{\boldsymbol{K}}(\{p_a'\}_{a \in \mathbb{N}_0})$. Write $\psi_{\boldsymbol{K}}(\{p_a\}_{a \in \mathbb{N}_0}) = \{\nu_N\}_{N \in \mathbb{N}}$ and $\psi_{\boldsymbol{K}}(\{p_a'\}_{a \in \mathbb{N}_0}) = \{\nu_N'\}_{N \in \mathbb{N}}$, and note that we have $\nu_N(N) = p_N$ and $\nu_N'(N) = p_N'$ for all $N \in \mathbb{N}$. Hence, we have $p_a = p_a'$ for all $a \in \mathbb{N}$, and we of course also have $p_0 = p_0' = 1$. Thus, $\{p_a\}_{a \in \mathbb{N}_0} = \{p_a'\}_{a \in \mathbb{N}_0}$ so $\psi_{\boldsymbol{K}}$ is injective.

To see that $\psi_{\boldsymbol{K}}$ is surjective, take any $\{\nu_N\}_{N \in \mathbb{N}} \in \boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$, and define $\{p_a\}_{a \in \mathbb{N}_0}$ via $p_0 = 1$ and $p_a = \nu_a(a)$ for $a \in \mathbb{N}$. Then take any $N \in \mathbb{N}$ and $a \in [\![1, N]\!]$, and compute, using compatibility:

$$\begin{aligned}
p_a - \sum_{M=a}^{N-1} C_{M,a}^{\boldsymbol{K}} p_{M+1} &= \nu_a(a) - \sum_{M=0}^{N-1} C_{M,a}^{\boldsymbol{K}} \nu_{M+1}(M+1) \\
&= \nu_a(a) - \sum_{M=a}^{N-1} (\nu_M(a) - \nu_{M+1}(a)) = \nu_N(a).
\end{aligned}$$

By summing over these and taking complements, we also get

$$p_0 - \sum_{M=0}^{N-1} C^{\boldsymbol{K}}_{M,0} p_{M+1} = \nu_N(0),$$

hence $\psi_{\boldsymbol{K}}(\{p_a\}_{a \in \mathbb{N}_0}) = \{\nu_N\}_{N \in \mathbb{N}}$. Therefore, $\psi_{\boldsymbol{K}}$ is surjective. This proves that $\psi_{\boldsymbol{K}}$ is in fact a bijection.

Observe now that the inverse map $\psi_{\boldsymbol{K}}^{-1} : \boldsymbol{\mathcal{D}}_1(\boldsymbol{K}) \to \boldsymbol{\mathcal{D}}_2(\boldsymbol{K})$ is given exactly by $\psi_{\boldsymbol{K}}^{-1}(\{\nu_N\}_{N \in \mathbb{N}}) = \{p_a\}_{a \in \mathbb{N}_0}$ where $p_0 = 1$ and $p_a = \nu_a(a)$ for $a \in \mathbb{N}$, and that this map is clearly continuous. Since $\boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$ is compact by Lemma 4.2 and $\boldsymbol{\mathcal{D}}_2(\boldsymbol{K})$ is Hausdorff, it is classical that $\psi_{\boldsymbol{K}}^{-1}$ is in fact a homeomorphism, hence that $\psi_{\boldsymbol{K}}$ is a homeomorphism. This finishes the proof. $\qquad \square$

Let us build the following diagram to succintly summarize the three different perspectives on VIDs that we have developed thus far:

$$\mathcal{M}_1(\boldsymbol{\mathcal{I}}) \qquad\qquad \prod_{N \in \mathbb{N}} \mathcal{M}_1(\llbracket 0, N \rrbracket) \qquad\qquad [0,1]^{\mathbb{N}_0}$$

$$\text{\rotatebox{90}{$\subset$}} \qquad\qquad\qquad\qquad \text{\rotatebox{90}{$\subset$}} \;\; \text{conv.}$$

$$\boldsymbol{\mathcal{D}}_0 \xleftarrow{\quad\text{homeo.}\quad} \boldsymbol{\mathcal{D}}_1 \qquad\qquad\qquad \text{\rotatebox{90}{$\subset$}} \;\; \text{conv.}$$

$$\text{\rotatebox{90}{$\subset$}} \;\; \text{conv.}$$

$$\boldsymbol{\mathcal{D}}_1(\boldsymbol{K}) \xleftarrow{\text{lin. homeo.}} \boldsymbol{\mathcal{D}}_2(\boldsymbol{K})$$

In more detail, the space $\boldsymbol{\mathcal{D}}_0$ describes the law of an entire Markovian probability distribution on $\boldsymbol{\mathcal{I}}$, and the space $\boldsymbol{\mathcal{D}}_1$ describes only the marginal sequences of such probability distributions; however, we saw in Proposition 3.13 that these spaces were, in fact, homeomorphic. Now consider fixing some $\boldsymbol{K} \in \boldsymbol{\mathcal{K}}$. The space $\boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$ describes the (convex) subspace of $\boldsymbol{\mathcal{D}}_1$ consisting of the sequences of probability distrubutions which are compatible with $\boldsymbol{K}$, and the space $\boldsymbol{\mathcal{D}}_0(\boldsymbol{K})$ describes each marginal distribution with a single real number; however, we saw in Proposition 4.7 that these spaces were, in fact, linearly homeomorphic. (The latter correspondence is predicated on the observation that knowing that $\boldsymbol{\nu}$ is compatible with $\boldsymbol{K}$ renders much of its information as redundant.)

Now we return to the task at hand of understanding the extreme points of $\boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$. The importance of Proposition 4.7 is that it shows that it suffices to find the extreme points of $\boldsymbol{\mathcal{D}}_2(\boldsymbol{K})$ and then simply compute their image under $\psi_{\boldsymbol{K}}^{-1}$. Since $\boldsymbol{\mathcal{D}}_2(\boldsymbol{K})$ is just a subset of $[0,1]^{\mathbb{N}_0}$ which is defined by countably many linear inequalities, its extreme point arise by setting "as many of them as possible" to equalities. While this is by no means easy in general, the plan can be executed in many simple examples. The explicit tools for doing so are the following results:

Lemma 4.8. *The maps*

$$\psi_{\boldsymbol{K}} : \mathrm{ex}(\boldsymbol{\mathcal{D}}_2(\boldsymbol{K})) \to \mathrm{ex}(\boldsymbol{\mathcal{D}}_1(\boldsymbol{K}))$$

and

$$\psi_{\boldsymbol{K}} : \overline{\mathrm{ex}(\boldsymbol{\mathcal{D}}_2(\boldsymbol{K}))} \to \overline{\mathrm{ex}(\boldsymbol{\mathcal{D}}_1(\boldsymbol{K}))}$$

are well-defined homeomorphisms.

Proof. First let us show that $\psi_{\boldsymbol{K}} : \mathrm{ex}(\boldsymbol{\mathcal{D}}_2(\boldsymbol{K})) \to \mathrm{ex}(\boldsymbol{\mathcal{D}}_1(\boldsymbol{K}))$ is well-defined. Indeed, suppose that $p = \{p_a\}_{a \in \mathbb{N}_0} \in \mathrm{ex}(\boldsymbol{\mathcal{D}}_2(\boldsymbol{K}))$ and that $\boldsymbol{\nu}, \boldsymbol{\nu}' \in \boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$ and

$\alpha \in (0,1)$ have $(1-\alpha)\boldsymbol{\nu} + \alpha\boldsymbol{\nu}' = \psi_{\boldsymbol{K}}(p)$. Then take $\psi_{\boldsymbol{K}}^{-1}$ of both sides and use the linearity guaranteed by Proposition 4.7 to get $(1-\alpha)\psi_{\boldsymbol{K}}^{-1}(\boldsymbol{\nu}) + \alpha\psi_{\boldsymbol{K}}^{-1}(\boldsymbol{\nu}') = p$. Since $p \in \operatorname{ex}(\boldsymbol{\mathcal{D}}_2(\boldsymbol{K}))$, this implies $\psi_{\boldsymbol{K}}^{-1}(\boldsymbol{\nu}) = \psi_{\boldsymbol{K}}^{-1}(\boldsymbol{\nu}')$, hence $\boldsymbol{\nu} = \boldsymbol{\nu}'$. Therefore, $\psi_{\boldsymbol{K}}(p) \in \operatorname{ex}(\boldsymbol{\mathcal{D}}_1(\boldsymbol{K}))$, so $\psi_{\boldsymbol{K}} : \operatorname{ex}(\boldsymbol{\mathcal{D}}_2(\boldsymbol{K})) \to \operatorname{ex}(\boldsymbol{\mathcal{D}}_1(\boldsymbol{K}))$ is indeed a well-defined map. Since $\psi_{\boldsymbol{K}} : \boldsymbol{\mathcal{D}}_2(\boldsymbol{K}) \to \boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$ is a homeomorphism by Proposition 4.7, it follows that $\psi_{\boldsymbol{K}} : \operatorname{ex}(\boldsymbol{\mathcal{D}}_2(\boldsymbol{K})) \to \operatorname{ex}(\boldsymbol{\mathcal{D}}_1(\boldsymbol{K}))$ is also a homeomorphism. Moreover, $\psi_{\boldsymbol{K}} : \boldsymbol{\mathcal{D}}_2(\boldsymbol{K}) \to \boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$ being a homeomorphism implies that the same result extends to the closures. $\square$

Corollary 4.9. *For any $\boldsymbol{K} \in \boldsymbol{\mathcal{K}}$ and $N \in \mathbb{N}$, the element $\psi_{\boldsymbol{K}}^{-1}(\delta_N^{\boldsymbol{K}})$ is the unique $\{p_a\}_{a \in \mathbb{N}_0} \in \boldsymbol{\mathcal{D}}_2(\boldsymbol{K})$ satisfying $p_N = 1$ and $p_a = 0$ for all $a > N$.*

Proof. Immediate from Lemma 4.3 and from the characterization of the map $\psi_{\boldsymbol{K}}^{-1} : \boldsymbol{\mathcal{D}}_1(\boldsymbol{K}) \to \boldsymbol{\mathcal{D}}_2(\boldsymbol{K})$ given in the proof of Proposition 4.7. $\square$

At the end of this section we will see some examples of using these tools to compute the extreme points of $\boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$ for some specific choices of $\boldsymbol{K}$.

4.2. Equilibrium and stationary distributions. Fix $N \in \mathbb{N}$ and let $K \in \mathcal{K}_N$ be any TM. We say that $\nu \in \mathcal{M}_1(\llbracket 0, N \rrbracket)$ is a *stationary distribution* if we have $\nu K = \nu$. Of course, the name comes from the fact that, if X is the MC with canonical data (ν, K), then it is a stationary stochastic process. It is classical that, if $K \in \mathcal{K}_N$ is irreducible and aperiodic, then there exists a unique stationary distribution ν, and it moreover has the property that if X is a MC with canonical data (ν', K) for any $\nu' \in \mathcal{M}_1(\llbracket 0, N \rrbracket)$, then the law of $X(i)$ converges to ν in distribution as $i \to \infty$. For this latter reason, stationary distributions are sometimes called *equilibrium distributions*. In this subsection we explore analogous notions of equilibrium and stationary distributions for VTMs. Interestingly, in the virtual case it becomes important to distinguish between these notions.

For any $\boldsymbol{K} \in \boldsymbol{\mathcal{K}}$, let us say that $\boldsymbol{\nu} = \{\nu_N\}_{N \in \mathbb{N}} \in \boldsymbol{\mathcal{D}}_1$ is an *equilibrium distribution for* $\boldsymbol{K} = \{K_N\}_{N \in \mathbb{N}} \in \boldsymbol{\mathcal{K}}$ if we have $\nu_N K_N = \nu_N$ for all $N \in \mathbb{N}$, and let us say that it is a *stationary distribution for* $\boldsymbol{K}$ if it is an equilibrium distribution for $\boldsymbol{K}$ which is also compatible with $\boldsymbol{K}$. We write $\boldsymbol{\mathcal{D}}_1^{\mathrm{eq}}(\boldsymbol{K})$ for the space of VIDs that are equilibrium distributions for $\boldsymbol{K}$, and write $\boldsymbol{\mathcal{D}}_1^{\mathrm{st}}(\boldsymbol{K})$ for the space of VIDs that are stationary distributions for $\boldsymbol{K}$.

Lemma 4.10. *For any $\boldsymbol{K} \in \boldsymbol{\mathcal{K}}$, the spaces $\boldsymbol{\mathcal{D}}_1^{eq}(\boldsymbol{K})$ and $\boldsymbol{\mathcal{D}}_1^{st}(\boldsymbol{K})$ are compact and convex.*

Proof. By Lemma 4.2 and the fact $\boldsymbol{\mathcal{D}}_1^{\mathrm{st}}(\boldsymbol{K}) = \boldsymbol{\mathcal{D}}_1^{\mathrm{eq}}(\boldsymbol{K}) \cap \boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$, it suffices to show that $\boldsymbol{\mathcal{D}}_1^{\mathrm{eq}}(\boldsymbol{K})$ is compact and convex. To do this, we simply write

$$\boldsymbol{\mathcal{D}}_1^{\mathrm{eq}}(\boldsymbol{K}) = \bigcap_{N \in \mathbb{N}} \left\{ \{\nu_N\}_{N \in \mathbb{N}} \in \boldsymbol{\mathcal{D}}_1 : \nu_N K_N = \nu_N \right\},$$

and the result follows. $\square$

Let us now give two results which justify the nomenclature of the terms just introduced. We say that a VTM $\boldsymbol{K} = \{K_N\}_{N \in \mathbb{N}} \in \boldsymbol{\mathcal{K}}$ is *irreducible* if, for each $N \in \mathbb{N}$ we have $K_N(i,0) = 0$ for all $i \in \llbracket 1, N \rrbracket$ and and that $\llbracket 1, N \rrbracket$ is a communicating class of K_N. We say $\boldsymbol{K}$ is *aperiodic* if K_N is aperiodic for each $N \in \mathbb{N}$. The proofs

of both results are immediate from the definitions and from the analagous results for classical MCs.

Lemma 4.11. *If $\boldsymbol{K} = \{K_N\}_{N \in \mathbb{N}} \in \boldsymbol{\mathcal{K}}$ is an irreducible and aperiodic VTM, then there exists a unique equilibrium distribution $\boldsymbol{\nu} = \{\nu_N\}_{N \in \mathbb{N}} \in \boldsymbol{\mathcal{D}}_1$ for $\boldsymbol{K}$. Moreover, if $\boldsymbol{X} = \{X_N\}_{N \in \mathbb{N}}$ is any VMC with VTM $\boldsymbol{K}$, then for each $N \in \mathbb{N}$, the law of $X_N(i)$ converges to ν_N in distribution as $i \to \infty$.*

Lemma 4.12. *Suppose that $\boldsymbol{K} \in \boldsymbol{\mathcal{K}}$ is a VTM with $\boldsymbol{\nu} \in \boldsymbol{\mathcal{D}}_1$ a stationary distribution. Then, the VMC $\boldsymbol{X} = \{X_N\}_{N \in \mathbb{N}}$ with canonical data $(\boldsymbol{\nu}, \boldsymbol{K})$ has the property that X_N is a stationary stochastic process for each $N \in \mathbb{N}$.*

4.3. Examples. Let us see some examples of the ideas developed in this section.

Example 4.13. Consider the VMC $\boldsymbol{X}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ constructed via Example 2.23, whose VTM $\boldsymbol{K}$ is given in Example 3.24, where F and $\{L_N\}_{N \in \mathbb{N}}$ are left general.

As a first step, let us compute the elements of the set $\boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$. By Choquet's theorem, it suffices to find the extreme points of this set. To do this, note that we have $C_{N,N}^{\boldsymbol{K}} = 1$ for all $N \in \mathbb{N}_0$, hence $\boldsymbol{\mathcal{D}}_2(\boldsymbol{K})$ is exactly the space of all nondecreasing sequences in $[0,1]$ with $p_0 = 1$. By Lemma A.2, we know that the extreme points of this space are exactly the sequences $\{1^N 0^\infty : N \in \mathbb{N}\} \cup \{1^\infty\}$. Moreover, we have $\psi_{\boldsymbol{K}}(1^{N+1} 0^\infty) = \delta_N^{\boldsymbol{K}}$ for all $N \in \mathbb{N}_0$ by Lemma 4.9, and also $\lim_{N \to \infty} \delta_N^{\boldsymbol{K}} = \psi_{\boldsymbol{K}}(1^\infty)$ which is extreme by Lemma 4.8. Of course, $\lim_{N \to \infty} \delta_N^{\boldsymbol{K}}$ is just the VID introduced in Example 3.24. Therefore, we have

$$\mathrm{ex}(\boldsymbol{\mathcal{D}}_1(\boldsymbol{K})) = \{\delta_N^{\boldsymbol{K}} : N \in \mathbb{N}_0\} \cup \left\{ \lim_{N \to \infty} \delta_N^{\boldsymbol{K}} \right\}$$

In other words, this VTM has an extremal VID corresponding to starting in each finite level $N \in \mathbb{N}$ as well as a unique extremal VID corresponding to starting at infinity.

Next let us find the points of $\boldsymbol{\mathcal{D}}_1^{\mathrm{eq}}(\boldsymbol{K})$. As long as $q_N > 0$ for all $N \in \mathbb{N}$, it follows that $\boldsymbol{K}$ is irreducible and aperiodic, hence by Lemma 4.11 that $\#(\boldsymbol{\mathcal{D}}_1^{\mathrm{eq}}(\boldsymbol{K})) = 1$. An easy calculation shows the unique element $\boldsymbol{\nu} = \{\nu_N\}_{N \in \mathbb{N}} \in \boldsymbol{\mathcal{D}}_1^{\mathrm{eq}}(\boldsymbol{K})$ is given by

$$\nu_N(a) = z_N^{-1} \prod_{M=1}^{a-1} q_M^{-1} \qquad \text{where} \qquad z_N = \sum_{a=1}^{N} \prod_{M=1}^{a-1} q_M^{-1},$$

for $N \in \mathbb{N}$ and $a \in [\![0, N]\!]$.

Finally, we find $\boldsymbol{\mathcal{D}}_1^{\mathrm{st}}(\boldsymbol{K})$. Note that for $N \in \mathbb{N}$ and $a = 1$ we have

$$\nu_N(1) = z_N^{-1} > z_{N+1}^{-1} = \nu_{N+1}(1) + \nu_{N+1}(N+1) C_{N,1}^{\boldsymbol{K}}$$

which contradicts (3.8), hence $\boldsymbol{\nu}$ is not compatible with $\boldsymbol{K}$. Therefore, $\boldsymbol{\mathcal{D}}_1^{\mathrm{st}}(\boldsymbol{K})$ is empty. In particular, we see that $\boldsymbol{K}$ admits equilibrium distributions but no stationary distributions.

Example 4.14. Consider the VMC $\boldsymbol{X}$ on $(\Omega, \mathcal{F}, \mathbb{P}_0)$ constructed via Example 2.24, whose VTM $\boldsymbol{K}$ is given in Example 3.25. Recall that $\{\ell_N\}_{N \in \mathbb{N}}$ is arbitrary and does not affect the VTM.

To begin, let us compute $\boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$. Note that we have $C_{1,1}^{\boldsymbol{K}} = 1$ as well as $C_{N,N-1}^{\boldsymbol{K}} = 1$ for all $N \in \mathbb{N}$ with $N \geq 2$, hence

$$\boldsymbol{\mathcal{D}}_2(\boldsymbol{K}) = \{\{p_a\}_{a \in \mathbb{N}_0} \in [0,1]^{\mathbb{N}_0} : p_1 \geq p_2 + p_3 \text{ and } p_a \geq p_{a+2} \text{ for } a \geq 2\}.$$

Now we find its extreme points. If $p_1 = 0$ then this forces $p_2 = p_3 = 0$, hence also $p_a = 0$ for all $a \in \mathbb{N}$, and this point is clearly extreme. If $p_1 = 1$, then the resulting point is extreme only when at most one of p_2 or p_3 is equal to 1; hence the terms of some parity are identically equal to zero, while the terms of the other parity form a non-decreasing sequence whose extremality we can characterize via the results of Appendix A. If the even indices are identically zero, then the extreme points of the odd indices are characterized by Lemma A.1, and if the odd indices are identically zero, then the extreme points of the even indices are characterized by Lemma A.2. Moreover, it is easy to show that there are no other extreme points. Transforming these back to $\mathcal{D}_1(\boldsymbol{K})$ via $\psi_{\boldsymbol{K}}$, we find

$$\mathrm{ex}(\mathcal{D}_1(\boldsymbol{K})) = \{\delta_N^{\boldsymbol{K}} : N \in \mathbb{N}_0\} \cup \left\{ \lim_{N \to \infty} \delta_{2N}^{\boldsymbol{K}}, \lim_{N \to \infty} \delta_{2N+1}^{\boldsymbol{K}} \right\}.$$

Of course, the VIDs $\lim_{N \to \infty} \delta_{2N}^{\boldsymbol{K}}$ and $\lim_{N \to \infty} \delta_{2N+1}^{\boldsymbol{K}}$ both correspond to "starting at infinity" but in two different senses.

This point deserves further elaboration. Recall the representation of this VID given in Example 2.24, which states that the sample paths of this VMCs represent repeatedly choosing, uniformly at random, one of two different paths coming down from infinity. These two VIDs then correspond respectively to starting at the two different "trailheads"; geometrically speaking, $\lim_{N \to \infty} \delta_{2N}^{\boldsymbol{K}}$ and $\lim_{N \to \infty} \delta_{2N+1}^{\boldsymbol{K}}$ correspond to starting infinitesimally left of the origin and infinitesimally right of the origin, respectively. At the origin, the corresponding VID is in fact

$$\{\mathbb{P}_0 \circ X_N(0)^{-1}\}_{N \in \mathbb{N}} = \frac{1}{2} \left(\lim_{N \to \infty} \delta_{2N}^{\boldsymbol{K}} + \lim_{N \to \infty} \delta_{2N+1}^{\boldsymbol{K}} \right).$$

This dispels the plausible but false notion that, for any VTM $\boldsymbol{K} \in \mathcal{K}$ arising in the setting of Lemma 2.21, all the elements of $\mathrm{ex}(\mathcal{D}_1(\boldsymbol{K})) \setminus \{\delta_N^{\boldsymbol{K}} : N \in \mathbb{N}_0\}$ are of the form $\{\mathbb{P}_x \circ X_N(0)^{-1}\}_{N \in \mathbb{N}}$ for some $x \in S$ which is an accumulation point of $\{L_N\}_{N \in \mathbb{N}}$.

Next let us find the points of $\mathcal{D}_1^{\mathrm{eq}}(\boldsymbol{K})$. It is clear that $\boldsymbol{K}$ is irreducible and aperiodic, hence by Lemma 4.11 that $\#(\mathcal{D}_1^{\mathrm{eq}}(\boldsymbol{K})) = 1$. Then note that K_N is doubly-stochastic for each $N \in \mathbb{N}$, hence $\mathcal{D}_1^{\mathrm{eq}}(\boldsymbol{K}) = \{\mathbf{U}\}$.

To see that $\mathcal{D}_1^{\mathrm{st}}(\boldsymbol{K})$ is empty, it suffices to show that $\mathbf{U}$ is not compatible with $\boldsymbol{K}$. Indeed, for any $N \in \mathbb{N}$ with $N \geq 2$ and $a = N - 1$ we have

$$\nu_N(N-1) = \frac{1}{N} \neq \frac{2}{N+1} = \nu_{N+1}(N-1) + \nu_{N+1}(N+1)C_{N,N-1}^{\boldsymbol{K}},$$

whence the result.

Example 4.15. Define the VTM $\boldsymbol{K} = \{K_N\}_{N \in \mathbb{N}}$ via

$$K_N = \begin{bmatrix} 1/N & 1/N & \cdots & 1/N & 1/N \\ 1/N & 1/N & \cdots & 1/N & 1/N \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1/N & 1/N & \cdots & 1/N & 1/N \\ 1/N & 1/N & \cdots & 1/N & 1/N \end{bmatrix}$$

for all $N \in \mathbb{N}$. Since $K_N(a,b) \to 0$ as $N \to \infty$ for all $a \in \mathbb{N}$ and $b \in \mathbb{N}_0$, we have by Lemma 3.6 that $\boldsymbol{K} \notin \iota(K)$. Intuitively speaking, $\boldsymbol{K}$ is the VTM representing the "random walk on the infinite clique".

This example is somewhat more complicated than the previous examples in that we can only get a partial understanding of $\boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$. To do this, note that we have $C^{\boldsymbol{K}}_{N,a} = 1/N$ for all $N \in \mathbb{N}$ and $a \in [\![1, N]\!]$ also $C^{\boldsymbol{K}}_{0,0} = 1$ by convention. Hence:

$$\boldsymbol{\mathcal{D}}_2(\boldsymbol{K}) := \left\{ \{p_a\}_{a \in \mathbb{N}_0} \in [0,1]^{\mathbb{N}_0} : \begin{array}{l} p_0 = 1 \geq \sum_{M=1}^{\infty} M^{-1} p_{M+1}, \text{ and} \\ p_a \geq \sum_{M=a}^{\infty} M^{-1} p_{M+1} \text{ for } a \in \mathbb{N} \end{array} \right\}$$

In principle one can use the map $\psi_{\boldsymbol{K}}$ to transform this to an understanding of $\boldsymbol{\mathcal{D}}_1(\boldsymbol{K})$, but, unfortunately, we are not able to characterize the extreme points of this set.

A much easier question is to understand the equilibrium and stationary distributions for $\boldsymbol{K}$. Note that $\boldsymbol{K}$ is clearly irreducible and aperiodic, and moreover that, for each $N \in \mathbb{N}$, the matrix K_N is doubly-stochastic. Hence we have $\boldsymbol{\mathcal{D}}_1^{\mathrm{eq}}(\boldsymbol{K}) = \{\mathbf{U}\}$. One easily checks that $\mathbf{U}$ is in fact compatible with $\boldsymbol{K}$, hence also that $\boldsymbol{\mathcal{D}}_1^{\mathrm{st}}(\boldsymbol{K}) = \{\mathbf{U}\}$.

4.4. The virtual Birkhoff polytope. As we saw in Example 2.22, VMCs can be seen as a generalization of virtual permutations. We also know the classical Birkhoff-von Neumann theorem that, in a fixed dimension, the (closed) convex hull of the TMs of all permutation matrices is exactly the space of doubly-stochastic matrices. Hence, it is natural to ask whether, in some suitable sense, the closed convex hull of the VTMs of all virtual permutations coincides with the space of "doubly-stochastic" VTMs. As we show in this section, the natural generalization is false, but we are able to gain some understanding of the convexity structure of this space.

For each $N \in \mathbb{N}$, write $\mathcal{B}_N \subseteq \mathcal{K}_N$ for the subspace of all doubly-stochastic transition matrices (*DSTMs*) on $[\![1, N]\!]$; these are often called the *Birkhoff polytopes*. Also write $\mathcal{B} \subseteq \mathcal{K}$ for the subspace of all doubly-stochastic transition matrices (*DSTMs*). We say that a VTM $\boldsymbol{K} = \{K_N\}_{N \in \mathbb{N}} \in \boldsymbol{\mathcal{K}}$ is *doubly-stochastic* if K_N is doubly-stochastic for each $N \in \mathbb{N}$, or also that $\boldsymbol{K}$ is a *DSVTM*, and we write $\boldsymbol{\mathcal{B}} \subseteq \boldsymbol{\mathcal{K}}$ for the space of all DSVTMs; this is called the *virtual Birkhoff polytope* although, as we will see, it is not a polytope at all. Note that $\boldsymbol{\mathcal{B}}$ is equivalently characterized as the space of all VTMs which admit the virtual uniform measure $\mathbf{U}$ as an equilibrium distribution.

Note, in particular, that the first row and column of K_N can be ignored for each $N \in \mathbb{N}$, hence can always use matrices with brackets as boundaries to represent these VTMs. Moreover, the projection operation $P_N : \mathcal{B}_{N+1} \to \mathcal{B}_N$ can be written as:

$$\begin{bmatrix} A & u \\ v^{\mathrm{T}} & p \end{bmatrix} \mapsto \begin{cases} \left[A + \frac{1}{1-p} u v^{\mathrm{T}} \right] & \text{if } p < 1 \\[2ex] \left[A \right] & \text{if } p = 1 \end{cases}, \tag{4.2}$$

which follows from (3.4), since (3.3) being doubly-stochastic necessarily implies $w = 0$ and $q = 0$, and since $p = 0$ implies $u = v = 0$. Now we develop the requisite properties of these projection maps.

Lemma 4.16. *For any $N \in \mathbb{N}$, the map $P_N : \mathcal{B}_{N+1} \to \mathcal{B}_N$ is continuous.*

Proof. Suppose that $\{K_n\}_{n \in \mathbb{N}}$ and K in $\mathcal{B}_{N+1}$ have $K_n \to K$. Using the block form (3.3), write $A_n, u_n, v_n,$ and p_n for the blocks of K_n for each $n \in \mathbb{N}$, and write

A, u, v, and p for the blocks of K. Then note that we have $A_n \to A, u_n \to u, v_n \to v$, and $p_n \to p$ in their respective topologies as $n \to \infty$. If $p < 1$, then we have $p_n < 1$ for sufficiently large $n \in \mathbb{N}$, hence we clearly have $P_N(K_n) \to P_n(K)$ as $n \to \infty$. In instead $p = 1$, then we recall the fact that all matrix norms on $\mathcal{K}_N$ are equivalent; hence it suffices to show $\|P_N(K_n) - P_n(K)\|_F \to 0$ as $n \to \infty$ where $\|\cdot\|_F$ denotes the Frobenius norm. To do this, we simply bound

$$\|P_N(K_n) - P_n(K)\|_F \leq \|A_n - A\|_F + \frac{1}{1-p_n}\|u_n v_n^T\|_F$$

$$= \|A_n - A\|_F + \frac{1}{1-p_n}\|u_n\|_2\|v_n\|_2$$

$$\leq \|A_n - A\|_F + \frac{1}{1-p_n}\|u_n\|_1\|v_n\|_1$$

$$\leq \|A_n - A\|_F + 1 - p_n.$$

Since the right side goes to zero as $n \to \infty$, the result is proved. $\qquad\square$

Lemma 4.17. *The map $\iota : \mathcal{B} \to \boldsymbol{\mathcal{B}}$ is a continuous injection.*

Proof. That $\iota : \mathcal{B} \to \boldsymbol{\mathcal{B}}$ is an injection follows from Lemma 3.5 in that $\iota : \mathcal{K} \to \boldsymbol{\mathcal{K}}$ is an injection, and that $\iota : \mathcal{B} \to \boldsymbol{\mathcal{B}}$ is continuous follows from Lemma 4.16. $\qquad\square$

Lemma 4.18. *The space $\boldsymbol{\mathcal{B}}$ is compact.*

Proof. Immediate from Lemma 4.16. $\qquad\square$

Example 4.19. Consider

$$K' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \qquad K'' = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

in $\mathcal{K}_3$, and note these correspond to some elements $\boldsymbol{K} := \{K_N\}_{N \in \mathbb{N}} := \iota(K)$ and $\boldsymbol{K}' := \{K'_N\}_{N \in \mathbb{N}} := \iota(K')$ in $\boldsymbol{\mathcal{K}}$. Since we have

$$P_2(K') = P_2(K'') = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

and also that $K_M = K_N$ and $K'_M = K'_N$ for all $M \geq N$, it follows that $\boldsymbol{K}, \boldsymbol{K}' \in \boldsymbol{\mathcal{B}}$. However, the matrix $K = \frac{1}{2}(K' + K'')$ satisfies

$$K = \begin{bmatrix} 1/2 & 0 & 1/2 \\ 0 & 1/2 & 1/2 \\ 1/2 & 1/2 & 0 \end{bmatrix} \qquad P_2(K) = \begin{bmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{bmatrix}.$$

Since $P_2(\frac{1}{2}(K' + K'')) \neq \frac{1}{2}(P_2(K') + P_2(K''))$, it follows that $\boldsymbol{\mathcal{B}}$ is not convex.

Despite the fact that $\boldsymbol{\mathcal{B}}$ is not convex as in the classical case, we can still understand some of its convexity structure. The following result characterizes exactly when the line segment between two points of $\boldsymbol{\mathcal{B}}$ is fully contained in $\boldsymbol{\mathcal{B}}$, and it implies that the relative interior of any such segment is either fully contained in or fully disjoint from $\boldsymbol{\mathcal{B}}$.

Proposition 4.20. *Take any $\boldsymbol{K}, \boldsymbol{K}' \in \boldsymbol{\mathcal{B}}$ and write $\boldsymbol{K} = \{K_N\}_{N \in \mathbb{N}}$ and $\boldsymbol{K}' = \{K'_N\}_{N \in \mathbb{N}}$. Then, the set of $\alpha \in [0,1]$ for which we have $(1-\alpha)\boldsymbol{K} + \alpha\boldsymbol{K}' \in \boldsymbol{\mathcal{B}}$ is $[0,1]$ if for all $N \in \mathbb{N}$ we have at least one of the properties*

(i) $K_{N+1}(N+1, N+1) = 1$,

(ii) $K'_{N+1}(N+1, N+1) = 1$,

(iii) $K_{N+1}(N+1, N+1) < 1$ and $K_{N+1}(N+1, N+1) < 1$, and

$$\frac{K_{N+1}(a, N+1)}{1 - K_{N+1}(N+1, N+1)} = \frac{K'_{N+1}(a, N+1)}{1 - K'_{N+1}(N+1, N+1)}$$

for all $a \in [\![1, N]\!]$, or

(iv) $K_{N+1}(N+1, N+1) < 1$ and $K_{N+1}(N+1, N+1) < 1$, and

$$\frac{K_{N+1}(N+1, b)}{1 - K_{N+1}(N+1, N+1)} = \frac{K'_{N+1}(N+1, b)}{1 - K'_{N+1}(N+1, N+1)}$$

for all $b \in [\![1, N]\!]$,

and it is $\{0, 1\}$ if there is some $N \in \mathbb{N}$ for which we have none of these properties.

Proof. To begin, let us define, for $\boldsymbol{K} = \{K_N\}_{N \in \mathbb{N}} \in \prod_{N \in \mathbb{N}} \mathcal{B}_N, N \in \mathbb{N}$, and $a \in [\![0, N]\!]$, the values

$$F^{\boldsymbol{K}}_{N,a,b} := \begin{cases} \frac{K_{N+1}(a, N+1) K_{N+1}(N+1, b)}{1 - K_{N+1}(N+1, N+1)}, & \text{if } K_{N+1}(N+1, N+1) < 1, \\ 0, & \text{if } K_{N+1}(N+1, N+1) = 1. \end{cases}$$

Then note that $\boldsymbol{K}$ is in $\mathcal{B}$ if and only if for all $N \in \mathbb{N}$ and $a, b \in [\![1, N]\!]$ we have $K_N(a, b) = K_{N+1}(a, b) + F^{\boldsymbol{K}}_{N,a,b}$. Thus for $\boldsymbol{K} = \{K_N\}_{N \in \mathbb{N}}$ and $\boldsymbol{K}' = \{K'_N\}_{N \in \mathbb{N}}$ in $\mathcal{B}$ and $\alpha \in [0, 1]$, we have

$$K_N(a, b) = K_{N+1}(a, b) + F^{\boldsymbol{K}}_{N,a,b}, \text{ and}$$

$$K'_N(a, b) = K'_{N+1}(a, b) + F^{\boldsymbol{K}'}_{N,a,b},$$

for all $N \in \mathbb{N}$ and $a, b \in [\![1, N]\!]$, and we want to understand when we have

$$(1 - \alpha) K_N(a, b) + \alpha K'_N(a, b)$$

$$= (1 - \alpha) K_{N+1}(a, b) + \alpha K_{N+1}(a, b) + F^{(1-\alpha)\boldsymbol{K} + \alpha \boldsymbol{K}'}_{N,a,b}$$

for all $N \in \mathbb{N}$ and $a, b \in [\![1, N]\!]$. Plugging the first two equations into the third, we see that $(1 - \alpha)\boldsymbol{K} + \alpha \boldsymbol{K}' \in \mathcal{B}$ if and only if

$$(1 - \alpha) F^{\boldsymbol{K}}_{N,a,b} + \alpha F^{\boldsymbol{K}'}_{N,a,b} = F^{(1-\alpha)\boldsymbol{K} + \alpha \boldsymbol{K}'}_{N,a,b} \tag{4.3}$$

for all $N \in \mathbb{N}$ and $a, b \in [\![1, N]\!]$. We use this characterization to complete the two directions of the proof.

Now let us show that, for each $N \in \mathbb{N}$, any of the properties enumerated above implies (4.3) for all $a, b \in [\![1, N]\!]$. First suppose both (i) and (ii), that $K_{N+1}(N+1, N+1) = K'_{N+1}(N+1, N+1) = 1$, and note that this implies

$$F^{(1-\alpha)\boldsymbol{K} + \alpha \boldsymbol{K}'}_{N,a,b} = F^{\boldsymbol{K}}_{N,a,b} = F^{\boldsymbol{K}'}_{N,a,b} = 0,$$

as needed. Now suppose that (i) holds but (ii) fails, that is, that $K_{N+1}(N+1, N+1) = 1$ and $K'_{N+1}(N+1, N+1) < 1$. Then we have $(1 - \alpha) K_{N+1}(N+1, N+1) + \alpha K'_{N+1}(N+1, N+1) < 1$, and also $K_{N+1}(a, N+1) = K_{N+1}(N+1, b) = 0$ by

double-stochasticity. Hence we can compute:

$$F_{N,a,b}^{(1-\alpha)\boldsymbol{K}+\alpha\boldsymbol{K}'}$$

$$= \frac{((1-\alpha)\cdot 0 + \alpha K_{N+1}'(a, N+1))((1-\alpha)\cdot 0 + \alpha K_{N+1}'(N+1, b))}{1 - (1-\alpha)\cdot 1 - \alpha K_{N+1}'(N+1, N+1)}$$

$$= \frac{\alpha^2 K_{N+1}'(a, N+1) K_{N+1}'(N+1, b)}{\alpha(1 - K_{N+1}'(N+1, N+1))}$$

$$= \frac{\alpha K_{N+1}'(a, N+1) K_{N+1}'(N+1, b)}{1 - K_{N+1}'(N+1, N+1)} = \alpha F_{N,a,b}^{\boldsymbol{K}'},$$

as needed. The same argument works for the case that (i) fails and (ii) holds. Now suppose that both (i) and (ii) fail. Then, (4.3) is equivalent to

$$(1-\alpha)\frac{xy}{1-z} + \alpha\frac{x'y'}{1-z'} = \frac{((1-\alpha)x + \alpha x')((1-\alpha)y + \alpha y')}{1 - (1-\alpha)z - \alpha z'} \qquad (4.4)$$

for the assignment

$$x = K_{N+1}(a, N+1), \qquad x' = K_{N+1}'(a, N+1)$$
$$y = K_{N+1}(N+1, b), \qquad y' = K_{N+1}'(N+1, b)$$
$$z = K_{N+1}(N+1, N+1), \qquad z' = K_{N+1}'(N+1, N+1).$$

Now a small arithmetic miracle occurs, as it turns out that (4.4) is equivalent to

$$\left(\frac{x}{1-z} - \frac{x'}{1-z'}\right)\left(\frac{y}{1-z} - \frac{y'}{1-z'}\right) = 0, \qquad (4.5)$$

which, remarkably, does not depend on α. Now note that either (iii) or (iv) imply that at least one factor in (4.5) is zero, hence (4.4) is satisfied, as needed.

Conversely, it is easy to show for each $N \in \mathbb{N}$ that (4.3) holding for all $a, b \in [\![1, N]\!]$ implies one of the enumerated properties: If both (i) and (ii) fail, then, as we remarked above, (4.3) is equivalent to (4.5), hence at least one of the factors must be equal to zero. This shows that for all $a, b \in [\![1, N]\!]$ we must have either

$$\frac{K_{N+1}(a, N+1)}{1 - K_{N+1}(N+1, N+1)} = \frac{K_{N+1}'(a, N+1)}{1 - K_{N+1}'(N+1, N+1)}$$

or

$$\frac{K_{N+1}(N+1, b)}{1 - K_{N+1}(N+1, N+1)} = \frac{K_{N+1}'(N+1, b)}{1 - K_{N+1}'(N+1, N+1)},$$

which in turn this implies that at least one of (iii) or (iv) holds. $\square$

The following is a direct translation of the conditions (i) to (iv) in terms of the analogous conditions for VTMs of virtual permutations.

Corollary 4.21. *Take any virtual permutations $\boldsymbol{\sigma}, \boldsymbol{\sigma}' \in \mathcal{S}$, and write $\boldsymbol{\sigma} = \{\sigma_N\}_{N\in\mathbb{N}}$ and $\boldsymbol{\sigma}' = \{\sigma_N'\}_{N\in\mathbb{N}}$. Then the set of $\alpha \in [0, 1]$ for which we have $(1-\alpha)\boldsymbol{K}(\boldsymbol{\sigma}) + \alpha\boldsymbol{K}(\boldsymbol{\sigma}') \in \mathcal{B}$ is $[0, 1]$ if for all $N \in \mathbb{N}$ we have at least one of*
 (i) $\sigma_N(N) = N$,
 (ii) $\sigma_N'(N) = N$,
(iii) $(\sigma_N)^{-1}(N) = (\sigma_N')^{-1}(N)$, or
(iv) $\sigma_N(N) = \sigma_N'(N)$,

and it is $\{0,1\}$ *if there is some* $N \in \mathbb{N}$ *for which we have none of these properties.*

Example 4.22. To see that Corollary 4.21 can give non-trivial information, let $\boldsymbol{\sigma} = \{\sigma_N\}_{N \in \mathbb{N}}$ and $\boldsymbol{\sigma}' = \{\sigma'_N\}_{N \in \mathbb{N}}$ in $\boldsymbol{\mathcal{S}}$ be the virtual permutations whose cycle structures are given by

$$\sigma_N = \begin{cases} (1)(23)(45)\cdots(N-1\,N), & \text{if } N \text{ odd,} \\ (1)(23)(45)\cdots(N-2\,N-1)(N), & \text{if } N \text{ even,} \end{cases}$$

and

$$\sigma'_N = \begin{cases} (12)(34)(56)\cdots(N-2\,N-1)(N), & \text{if } N \text{ odd,} \\ (12)(34)(56)\cdots(N-1\,N), & \text{if } N \text{ even,} \end{cases}$$

for all $N \in \mathbb{N}$ with $N \geq 2$. Note that (i) is true for odd N and that (ii) is true for even N, and therefore that we have $(1-\alpha)\boldsymbol{K}(\boldsymbol{\sigma}) + \alpha\boldsymbol{K}(\boldsymbol{\sigma}') \in \boldsymbol{\mathcal{B}}$ for all $\alpha \in [0,1]$. In particular, observe that $\frac{1}{2}\boldsymbol{K}(\boldsymbol{\sigma}) + \frac{1}{2}\boldsymbol{K}(\boldsymbol{\sigma}')$ is exactly the VTM given in Example 3.26, that is, the VTM of the random walk on $\mathbb{N}$ for which the boundary state 1 reflects and holds with equal probability.

Finally, recall that the *kernel* of a set S in a real vector space is the set of all $x \in S$ such that for all $x' \in S$ and $\alpha \in [0,1]$ we have $(1-\alpha)x + \alpha x' \in S$; this is denoted $\ker(S)$. Then, S is convex if and only if $\ker(S) = S$ and S is called *star-shaped* if $\ker(S)$ is non-empty. Hence, the kernel encodes some information about the convexity structure of possibly non-convex sets. Our last result shows that the kernel of the virtual Birkhoff polytope is a sinlgeton.

Theorem 4.23. *We have* $\ker(\boldsymbol{\mathcal{B}}) = \{\boldsymbol{I}\}$*, for* $\boldsymbol{I} = \{I_N\}_{N \in \mathbb{N}}$ *the VTM consisting of the identity matrix* I_N *for each level* $N \in \mathbb{N}$.

Proof. That $\boldsymbol{I}$ is in $\ker(\boldsymbol{\mathcal{B}})$ follows from Proposition 4.20 since (i) is always satisfied. For the converse, suppose $\boldsymbol{K} \in \ker(\boldsymbol{\mathcal{B}})$, and write $\boldsymbol{K} = \{K_N\}_{N \in \mathbb{N}}$. Note that it suffices to show $K_{N+1}(N+1, N+1) = 1$ for all $N \in \mathbb{N}$. To do this, take any $N \in \mathbb{N}$ with $N \geq 2$.

We now define three elements of $\boldsymbol{\mathcal{B}}$: First set

$$K^1_{N+1} := \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \end{bmatrix}$$

in $\mathcal{K}_{N+1}$, and define $\boldsymbol{K}^1 = \iota(K^1_{N+1})$. (Observe that $\boldsymbol{K}^1$ is just the virtual permutation matrix corresponding to the classical permutation with cycle structure $(1\,2\,N+1)$.) Then we have

$$\frac{K^1_{N+1}(1, N+1)}{1 - K^1_{N+1}(N+1, N+1)} = \frac{K^1_{N+1}(N+1, 2)}{1 - K^1_{N+1}(N+1, N+1)} = 0.$$

Second, define $K_{N+1}^2 \in \mathcal{K}_{N+1}$ via

$$
K_{N+1}^2 := \begin{bmatrix}
0 & 1/2 & 0 & 0 & \cdots & 0 & 0 & 0 & 1/2 \\
0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
0 & 1/2 & 0 & 0 & \cdots & 0 & 0 & 0 & 1/2
\end{bmatrix}
$$

and set $\boldsymbol{K}^2 = \iota(K_{N+1}^2)$. Then note that we have

$$
\frac{K_{N+1}^2(1, N+1)}{1 - K_{N+1}^2(N+1, N+1)} = \frac{K_{N+1}^2(N+1, 2)}{1 - K_{N+1}^2(N+1, N+1)} = 1.
$$

Third, consider $\boldsymbol{K}^3$ to be the VTM from Example 4.15, and note that we have

$$
\frac{K_{N+1}^3(1, N+1)}{1 - K_{N+1}^3(N+1, N+1)} = \frac{K_{N+1}^3(N+1, 2)}{1 - K_{N+1}^3(N+1, N+1)} = \frac{1}{N}.
$$

Now we apply Proposition 4.20 to the pair $(\boldsymbol{K}, \boldsymbol{K}^i)$ for $i \in \{1, 2, 3\}$ and we consider the value of $N \in \mathbb{N}$ that we have fixed. Since (ii) always fails by construction, we must have at least one of (i), (iii), or (iv), holding for each $i \in \{1, 2, 3\}$. If (i) were not true, then, by the pigeonhole principle, at least one of (iii) or (iv) would have to hold for two different $i \in \{1, 2, 3\}$, but this is a contradiction since then one of the values

$$
\frac{K_{N+1}(1, N+1)}{1 - K_{N+1}(N+1, N+1)} \qquad \text{or} \qquad \frac{K_{N+1}(N+1, 2)}{1 - K_{N+1}(N+1, N+1)}
$$

would be equal to two distinct values in $\{0, 1, \frac{1}{N}\}$. Therefore, (i) holds, that is, $K_{N+1}(N+1, N+1) = 1$ for all $N \in \mathbb{N}$ with $N \geq 2$.

It still remains to show that $K_2(2, 2) = 1$. To do this, define

$$
K_2(\theta) = \begin{bmatrix} \theta & 1-\theta \\ 1-\theta & \theta \end{bmatrix}
$$

for all $\theta \in (0, 1)$, and set $\boldsymbol{K}(\theta) = \iota(K_2(\theta))$. Again by Proposition 4.20, we must have that one of the enumerated properties holds for each $\theta \in (0, 1)$, and again we see that (ii) always fails. Thus, if (i) were not true, then we could choose three distinct values of $\theta \in (0, 1)$, and the pigeonhole principle would imply that at least one (iii) or (iv) would have to hold for two distinct values of $\theta \in (0, 1)$, which is again a contradiction. Thus, (i) holds, and we have proven $\boldsymbol{K} = \mathbf{I}$. $\qquad \square$

References

[1] A. Borodin and J. Kuan. *Random surface growth with a wall and Plancherel measures for $O(\infty)$*. Comm. Pure Appl. Math., 63(7):831-894, 2010.

[2] P. Bourgade, J. Najnudel, and A. Nikeghbali. *A unitary extension of virtual permutations.* Int. Math. Res. Not. IMRN, (18):4101-4134, 2013.

[3] A. Gnedin and V. Gorin. *Spherically symmetric random permutations.* Random Structures Algorithms, 55(2):342-355, 2019.

[4] V. E. Gorin. *Disjointness of representations arising in harmonic analysis on the infinitedimensional unitary group.* Functional Analysis and Its Applications, 44(2):92-105, 2010.

[5] J. Hoffmann-Jørgensen. *Markov sets.* Math. Scand., 24:145-166 (1970), 1969.

[6] O. Kallenberg. *Foundations of modern probability.* Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[7] A. F. Karr and A. O. Pittenger. *The inverse balayage problem for Markov chains.* Stochastic Process. Appl., 7(2):165-178, 1978.

[8] A. F. Karr and A. O. Pittenger. *The inverse balayage problem for Markov chains. II.* Stochastic Process. Appl., 9(1):35-53, 1979.

[9] S. V. Kerov and N. V. Tsilevich. *Stick breaking process generated by virtual permutations with Ewens distribution.* Journal of Mathematical Sciences, 87(6):4082-4093, 1997.

[10] S. Kerov, G. Olshanski, and A. Vershik. *Harmonic analysis on the infinite symmetric group. A deformation of the regular representation.* C. R. Acad. Sci. Paris Sér. I. Math., 316(8):773-778, 1993.

[11] A. Okounkov and A. Vershik. *A new approach to representation theory of symmetric groups.* Selecta Math. (N.S.), 2(4):581-605, 1996.

[12] M. Sharpe. *General theory of Markov processes,* volume 133 of Pure and Applied Mathematics. Academic Press, Inc., Boston, MA, 1988.

[13] E. Strahov. *Z-measures on partitions related to the infinite Gelfand pair $(S(2, , \infty), H(\infty))$.* Journal of Algebra, 323(2):349–370, 2010.

[14] N. V. Tsilevich. *Distribution of cycle lengths of infinite permutations.* Journal of Mathematical Sciences, 87(6):4072-4081, 1997.

[15] A. M. Vershik and N. V. Tsilevich. *On the Fourier transform on the infinite symmetric group.* Journal of Mathematical Sciences, 138(3):5663-5673, 2006.

Appendix A. The Space of Monotone Sequences

In this appendix we prove some results about extremality in convex spaces of monotone sequences living in a given bounded set of reals. We believe such results are likely well-known, but we could not find a reference. To begin, define

$$D = \left\{ \{x_k\}_{k \in \mathbb{N}} \in [0,1]^{\mathbb{N}} : x_k \geq x_{k+1} \text{ for all } k \in \mathbb{N} \right\},$$

and observe that D is clearly convex. Also define

$$D' = \{\{x_k\}_{k \in \mathbb{N}} \in D : x_1 = 1\},$$

which is also convex. When $[0,1]^{\mathbb{N}}$ is endowed with the product topology which is compact by Tychonoff, then D and D' are compact convex sets. Hence, one naturally inquires about the extreme points of D and D', which are easy to describe.

For each $m \in \mathbb{N}$, define the sequence $1^m 0^\infty := \{x_k\}_{k \in \mathbb{N}}$ via $x_k = \mathbb{1}\{k \leq m\}$ for all $k \in \mathbb{N}$. Also define $1^\infty := \{x_k\}_{k \in \mathbb{N}}$ via $x_k = 1$ for all $k \in \mathbb{N}$ and $0^\infty := \{x_k\}_{k \in \mathbb{N}}$ via $x_k = 0$ for all $k \in \mathbb{N}$.

Lemma A.1. $\mathrm{ex}(D) = \{0^\infty\} \cup \{1^m 0^\infty : m \in \mathbb{N}\} \cup \{1^\infty\}$.

Proof. That D is compact and convex is clear, as is the statement

$$\mathrm{ex}(D) \supseteq \{0^\infty\} \cup \{1^m 0^\infty : m \in \mathbb{N}\} \cup \{1^\infty\},$$

so only the reverse inclusion remains to be shown. Indeed, suppose that $\{x_k\}_{k \in \mathbb{N}}$ is not in the right side. Then the value $k_* = \inf\{k \in \mathbb{N} : x_k < 1\}$ is finite and satisfies $x_{k_*} \in (0, 1)$. Now define

$$k_{**} := \inf\{k \in \mathbb{N} : k \geq k_* \text{ and } x_k > x_{k+1}\}.$$

If $k_{**} = \infty$, then we set $\varepsilon := \min\{x_{k_*}, 1 - x_{k_*}\} > 0$ and define $x^\pm = \{x_k^\pm\}_{k \in \mathbb{N}_0}$ via

$$x_k^\pm = \begin{cases} 1, & \text{if } k < k_*, \\ x_{k_*} \pm \frac{1}{2}\varepsilon, & \text{if } k \geq k_*. \end{cases}$$

Notice that we have $x \notin \{x^+, x^-\}$ and $x = \frac{1}{2}(x^+ + x^-)$, thus x is not extreme. Otherwisewe have $k_{**} < \infty$ and $x_{k_{**}} \in (x_{k_{**}+1}, 1)$, so we can set

$$\varepsilon := \min\{x_{k_*}, 1 - x_{k_*}, x_{k_{**}} - x_{k_{**}+1}, 1 - x_{k_{**}}\} > 0,$$

and we can define $x^\pm = \{x_k^\pm\}_{k \in \mathbb{N}_0}$ via

$$x_k^\pm = \begin{cases} 1, & \text{if } k < k_*, \\ x_{k_*} \pm \frac{1}{2}\varepsilon, & \text{if } k_* \leq k \leq k_{**}, \\ x_k, & \text{if } k > k_{**}. \end{cases}$$

Again we have $x \notin \{x^+, x^-\}$ and $x = \frac{1}{2}(x^+ + x^-)$, so x is not extreme. Therefore, the result is proved. $\qquad\square$

The same argument shows that we also have:

Lemma A.2. $\mathrm{ex}(D') = \{1^m 0^\infty : m \in \mathbb{N}\} \cup \{1^\infty\}$.

Since $1^m 0^\infty \to 1^\infty$ in the product topology as $m \to \infty$, the compact convex sets D and D' both have closed sets of extreme points.

Steven N. Evans
UC Berkeley,
367 Evans Hall,
Berkeley, CA,
United States
evans@stat.berkeley.edu

Adam Q. Jaffe
UC Berkeley,
367 Evans Hall,
Berkeley, CA,
United States
aqjaffe@berkeley.edu