Domains and Stochastic Processes

Michael Mislove

Department of Computer Science
Tulane University, New Orleans, LA 70118

Abstract

Domain theory has a long history of applications in theoretical computer science and mathematics. In this article, we explore the relation of domain theory to probability theory and stochastic processes. The goal is to establish a theory in which Polish spaces are replaced by domains, and measurable maps are replaced by Scott-continuous functions. We illustrate the approach by recasting one of the fundamental results of stochastic process theory – Skorohod’s Representation Theorem – in domain-theoretic terms. We anticipate the domain-theoretic version of results like Skorohod’s Theorem will improve our understanding of probabilistic choice in computational models, and help devise models of probabilistic programming, with its focus on programming languages that support sampling from distributions where the results are applied to Bayesian reasoning.

Keywords: Domain theory, random variables, Skorohod Representation Theorem

1 Introduction

The goal of this paper is establish a strong connection between domain theory and stochastic process theory. This follows the emergence over the past several years of random variables in domain theory as models for probabilistic choice (cf. [25,31,3]) in programming language semantics, and of probabilistic programming semantics, an important tool for modeling programming languages that support sampling from probability distributions to study Bayesian inference (cf. [19,30]). Our aim is to devise domain representations of stochastic processes that are amenable to computational analysis.

Random variables are measurable maps $X : (S, \Sigma_S, \mu) \rightarrow (T, \Sigma_T)$ from a probability space $S$ to a measure space $T$. It is customary to identify a random variable $X$ with its law, $X^* \mu$, the push forward of the measure $\mu$ under $X$. A common setting is that of Polish spaces – completely metrizable separable spaces, because the probability measures on a Polish space also are a Polish space in the weak topology. Our approach expands Polish spaces into domains, which allows us to represent each random variable as a Scott-continuous map. This allows us to use techniques and results from domain theory to prove analogs of results about stochastic processes.
We illustrate our approach by reformulating Skorohod’s Representation Theorem, a fundamental result in stochastic process theory, in the domain setting. In more detail, Skorohod’s Theorem \([32]\) states that any Borel probability measure on a Polish space \(P\) is the law of a random variable \(X: [0, 1] \to P\). That is, if \(\mu\) is a Borel probability measure on a Polish space \(P\) and if \(\lambda\) denotes Lebesgue measure on the unit interval, then there is a measurable map \(X: [0, 1] \to P\) satisfying \(\mu = X_* \lambda\). Furthermore, if \(\mu_n \to_w \mu\) in \(\text{Prob} P\) in the weak topology, then the random variables \(X, X_n: [0, 1] \to P\) can be chosen with laws \(\mu\) and \(\mu_n\), respectively, so that \(X_n \to X\) almost surely wrt \(\lambda\). This allows one to replace arguments about the weak convergence of probability measures on a Polish space with arguments about almost sure convergence of measurable maps from the unit interval to the Polish space in question.

In the domain approach, the unit interval equipped with Lebesgue measure is replaced by a suitable domain equipped with a corresponding probability measure, and the class of Polish spaces is replaced by a suitable category of domains. In outline form, the approach relies on the following:

- It is well known that the Cantor set, \(C \simeq 2^\omega\), regarded as a countable product of two-point groups, is a standard probability space: the canonical quotient map \(\varphi: C \to [0, 1]\) is a Borel isomorphism taking Haar measure \(\nu_C\) on \(C\) to Lebesgue measure (cf., e.g., \([8]\)).

This implies there is a random variable \(X: [0, 1] \to P\) satisfying \(X_* \lambda = \mu\) iff there is a random variable \(X': C \to P\) with \(X'_* \nu_C = \mu\), for any probability measure \(\mu\) on a Polish space \(P\).

- We then expand the Cantor set into the Cantor tree: \(\mathbb{C}_T = \{0, 1\}^\omega \cup \{0, 1\}^\omega = \{0, 1\}^* \cup \{0, 1\}^\omega\), the set of finite and infinite words over \(\{0, 1\}\). This is a computational model for \(C\), since \(\mathbb{C}_T\) is a countably based bounded complete domain when endowed with the prefix order that satisfies \(C \simeq \text{Max} \mathbb{C}_T\).

- In fact, results in domain theory show every Polish space embeds as the space of maximal elements of some countably based bounded complete domain, \(\mathbb{D}_P\), and conversely, the space of maximal elements of any such domain is a Polish space in the relative Scott topology.

Together, the previous results allow us to prove that for every random variable \(X: [0, 1] \to P\), there is a Scott-continuous map \(f: \mathbb{C}_T \to \mathbb{D}_P\) having the same law as the random variable \(X: f_* \nu_C = (f|_C)_* \nu_C = X_* \lambda\).

- Just as every Polish space is the space of maximal elements of a countably based bounded complete domain, \(\mathbb{D}_P\), the family of Borel probability measures on a Polish space \(P\) is the space of maximal elements of the domain of probability measures on the associated countably based bounded complete domain \(\mathbb{D}_P\).

- These associations are topological: the weak topology on the space \(\text{Prob} \text{Max} \mathbb{D}\) of probability measures on a bounded complete domain \(\mathbb{D}\) coincides with the relative Lawson topology on the \(\text{Prob} \mathbb{D}\), viewed as a domain (via the isomorphism with the valuations domain on \(\mathbb{D}\)). In fact, the Lawson topology on \(\text{Prob} \mathbb{D}\), for \(\mathbb{D}\) a
countably based bounded complete domain, coincides with the weak topology.

Combining these results leads to our main theorem:

**Theorem 1. (Skorohod’s Theorem for Bounded Complete Domains)**

Let \( D \) be a countably-based bounded complete domain, and let \( \{\mu_n\}_n \in \text{Prob} D \) be a sequence of Borel probability measures satisfying \( \lim_n \mu_n = \mu \in \text{Prob} D \) in the Lawson topology (which agrees with the weak topology on probability measures). Then there are Scott-continuous maps \( f, f_n : CT \rightarrow D \) satisfying \( f_* \nu_C = \mu, f_n_* \nu_C = \mu_n \) for each \( n \), and \( f_n \rightarrow f \) pointwise wrt to the Scott topologies.

**Remark 1.1** This result is somewhat weaker than Skorohod’s Theorem for Polish spaces, since the convergence of the Scott-continuous functions \( f_n \) is in the Scott topology. The family of functions \( f_n \) described in the Theorem may not be directed, so we cannot conclude that \( \lim f_n(x) = f(x) \) for all \( x \in CT \); we can only conclude that \( \lim \inf_n f_n(x) \geq f(x) \). Still, for those \( x \in CT \) for which \( f(x) \in \text{Max} D \), we do get \( \lim f_n(x) = f(x) \) in the Lawson topology.

Skorohod’s Theorem is a corollary of Theorem 1 as follows. Any Polish space \( P \) has a computational model: a countably-based bounded complete domain \( D_P \) for which \( P \) is homeomorphic to the space \( \text{Max} D_P \) of maximal elements endowed with the relative Scott topology. In fact, \( \text{Max} D_P \) is a \( G_\delta \), hence a Borel subset of \( D_P \). So, if \( \mu_n \) and \( \mu \) are probability measures on \( P \), then they are probability measures on \( D_P \) concentrated on \( \text{Max} D_P \), and so Theorem 1 implies they can be realized as laws for Scott-continuous maps \( f, f_n : CT \rightarrow D_P \), i.e., \( f_* \nu_C = \mu \) and \( f_n_* \nu_C = \mu_n \) for each \( n \geq 0 \). Since \( \mu_n, \mu \) are concentrated on \( \text{Max} D \), we can restrict \( f_n, f \) to those points \( x \in \text{Max} CT \) where \( f_n(x), f(x) \in \text{Max} D \). It follows that \( f_n(x) \) converges to \( f(x) \) wrt the Lawson topology on \( D \) for \( \nu_C \)-almost all \( x \in \text{Max} CT \).

Finally, the canonical surjection \( \phi : C \rightarrow [0, 1] \) preserves all sups and infs, so it has a lower adjoint \( j : [0, 1] \rightarrow C \) that is a Borel isomorphism taking Lebesgue measure to \( \nu_C \). Thus \( (f \circ j)_* \lambda = f_* j_* \lambda = f_* \nu_C = \mu \) and similarly, \( (f_n \circ j)_* \lambda = \mu_n \) are the random variables guaranteed by Skorohod’s classic result.

We also report the results of our research that concerns one of the longstanding problems in domain theory. The Jung-Tix Problem asks whether there is a Cartesian closed category of domains for which the valuations monad is an endofunctor. We cannot answer this question, but we can contribute our knowledge of the structure of \( \forall D \) when \( D \) is a complete chain:

**Theorem 2.** If \( D \) is a countably based complete chain, then \( \text{Prob} D \) and \( \text{SProb} D \) (the family of sub-probability measures on \( D \)) are continuous lattices.

This result significantly expands our knowledge of the domain structure of the family of (sub)probability measures on a domain \( D \). Indeed, up to this point, the only domains \( D \) for which \( \text{SProb} D \) is known to be a domain are: (i) a (rooted) tree, \( T \), for which \( \text{SProb} T \in \text{BCD} \), the category of bounded complete domains, or (ii) a finite reverse tree \( T^{rev} \), in which case \( \text{SProb} T^{rev} \) is in RB [21].

The previous discussion hints at results that also are included in this research.
While our interest in Skorohod’s Theorem results in a focus on the probability measures, we also show that many of our results apply more broadly to the family of subprobability measures on a domain. Regarded as valuations, this is the probabilistic power domain, a much-studied construct in domain theory. We explain the relationship between (sub-)probability measures and valuations in detail, and describe how the more general results concerning these measures follow along the same lines as the arguments we present for probability measures. In each case, the more general result can be obtained by the same proof strategy as the one for probability measures.

1.1 Related Work

Beginning in the mid-1990s, Abbas Edalat developed domain-theoretic approaches to a number of areas, including integration theory [11], stochastic processes [12], dynamical systems and fractals [13], and Brownian motion [4]. The concept of a computational model emerged in Edalat’s work on domain models of spaces arising in real analysis using the domain of compact subsets under reverse inclusion, where the target space arises as the set of maximal elements. The first paper formally presenting such a model was [13], where a domain model for locally compact second countable spaces was given. That paper presents a range of applications of the approach, including dynamical systems, iterated function systems and fractals, a computational model for classical measure theory on locally compact spaces, and a computational generalization of Riemann integration. Related work led to the formal ball model [15] which was tailor-made for modeling metric spaces and Lipschitz functions. Further discussion of these developments occurs in our discussion of Polish spaces in Section 4 below.

Other related work concerns the development of random variable models of probabilistic computational processes. This began with [25], a paper that provided a domain model for finite random variables. Further efforts saw limited success until a few years ago. The model proposed in [17] turned out to be flawed, as was initially shown in [26,27]. But inspired by ideas from [17], Barker [3] devised a monad of random variables that gives an abstract model for randomized algorithms. This line of research was initiated by Scott [31], who showed how the $\mathcal{P}(\mathbb{N})$ model of the lambda calculus could be extended naturally to support probabilistic choice with the aid of random variables $X: [0,1] \rightarrow \mathcal{P}(\mathbb{N})$. Barker’s results generalize Scott’s approach by providing a model of randomized PCF that adds a version of probabilistic choice based on a random variables monad. Notably, this monad leaves important Cartesian closed categories of domains invariant – in particular, the category $\mathbf{BCD}$ of bounded complete domains, as well as the CCC $\mathbf{RB}$ of retracts of bifinite domains invariant, and each enjoys a distributive law with respect to at least one of nondeterminism monads.

The rest of the paper is as follows. In the next section, we review the material we need from a number of areas, including domain theory, topology, and probabil-
ity theory. Section 3 develops results about mappings from the Cantor tree to the space $\text{SProb } D$ of sub-probability measures on a countably-based coherent domain $D$. Section 4 contains the main results of the paper, by first recalling the development of Polish spaces as computational models, and then presenting the main theorems. Section 5 summarizes what's been proved, and discusses future work.

2 Background

In this section we present the background material we need for our main results.

2.1 Domains

Our results rely fundamentally on domain theory. Most of the results that we quote below can be found in [1] or [16]; we give specific references for those that appear elsewhere.

To start, a poset is a partially ordered set. A poset is directed complete if each of its directed subsets has a least upper bound, where a subset $S$ is directed if each finite subset of $S$ has an upper bound in $S$. A directed complete partial order is called a dcpo. The relevant maps between dcpos are the monotone maps that also preserve suprema of directed sets; these maps are usually called Scott continuous.

Restating things topologically, a subset $U \subseteq P$ of a poset is Scott open if (i) $U = \uparrow U \equiv \{x \in P \mid (\exists u \in U) \ u \leq x\}$ is an upper set, and (ii) if $\sup S \in U$ implies $S \cap U \neq \emptyset$ for each directed subset $S \subseteq P$. It is routine to show that the family of Scott-open sets forms a topology on any poset; this topology satisfies $\downarrow x \equiv \{y \in P \mid y \leq x\} = \overline{x}$ is the closure of a point, so the Scott topology is always $T_0$, and it is $T_1$ iff $P$ is a flat poset. A mapping between dcpos is Scott continuous in the order-theoretic sense iff it is a monotone map that is continuous with respect to the Scott topologies on its domain and range. We let DCPO denote the category of dcpos and Scott-continuous maps; DCPO is a Cartesian closed category.

If $P$ is a dcpo, and $x, y \in P$, then $x$ approximates $y$ iff for every directed set $S \subseteq P$, if $y \leq \sup S$, then there is some $s \in S$ with $x \leq s$. In this case, we write $x \ll y$ and we let $\downarrow y \equiv \{x \in P \mid x \ll y\}$. A basis for a poset $P$ is a family $B \subseteq P$ satisfying $\downarrow y \cap B$ is directed and $y = \sup(\downarrow y \cap B)$ for each $y \in P$. A continuous poset is one that has a basis, and a dcpo $P$ is a domain if $P$ is continuous. An element $k \in P$ is compact if $x \ll x$, and $P$ is algebraic if $KP = \{k \in P \mid k \ll k\}$ forms a basis. Domains are sober spaces in the Scott topology.

Domains admit a Hausdorff refinement of the Scott topology which will play a role in our work. The weak lower topology on $P$ has the sets of the form if $O = P \setminus \uparrow F$ as a basis, where $F \subseteq P$ is a finite subset. The Lawson topology on a domain $P$ is the common refinement of the Scott- and weak lower topologies on $P$. This topology has the family

$$\{U \setminus \uparrow F \mid U \text{ Scott open } \& \ F \subseteq P \text{ finite}\}$$

as a basis. The Lawson topology on a domain is always Hausdorff, and a domain is coherent if its Lawson topology is compact. We denote the closure of a subset
Let $D$ be a countably based domain with countable basis $B_D$, and let $x \in D$. Then:

(i) $x$ is the supremum of a countable chain $\{x_n\}_{n \in \mathbb{N}}$ with $x_n \ll x$ for each $n$.

(ii) If $D$ is coherent, then there is a countable chain of Lawson-open sets $U_n = \{\uparrow x_n \setminus \uparrow F_n \mid x_n \ll x \not\in F_n \subseteq B_D \text{ finite}\}$ with $x = \bigcap_n U_n$.

**Proof.** (i): Since $D$ has a countable base, there is a countable directed set $B \subseteq \downarrow x$ with $\sqcup D = x$. If we enumerate $B = \{b_0, b_1, \ldots\}$, then we define the desired sequence $x_n$ as follows: $x_0 = b_0$, and if $x_0 \ll x_1 \ll \ldots \ll x_n$ have been chosen from $B$, then we choose $x_{n+1} \in B$ with $b_i \ll x_{n+1}$ for each $i \leq n$ and $x_n \ll x_{n+1}$. This extends the sequence, and then a standard maximality argument shows we can choose an countable sequence $x_n$ with $x_n \ll x_{n+1} \ll x$ for each $n$. Finally, $x = \sqcup_n b_n \leq \sqcup_n x_n$, since $b_n \ll x_{n+1}$ for each $n$, but $x_n \ll x$ for each $n$ implies $\sqcup_n x_n \subseteq x$.

(ii): By definition of the Lawson topology, we know $\uparrow x_n \setminus \uparrow F$ is an open set containing $x$ if $x \not\in F \subseteq B_D$, and there are countably many of these sets since $B_D$ is countable. It’s then routine to extract a chain $U_n = \uparrow x_n \setminus \uparrow F_n$ whose intersection also is $x$. \qed

We also need some basic results about Galois adjunctions (cf. Section 0-3 of [16]) in the context of complete lattices. If $L$ and $M$ are complete lattices, a **Galois adjunction** is a pair of mappings $g: L \to M$ and $f: M \to L$ satisfying $f \circ g \leq 1_L$ and $g \circ f \geq 1_M$. In this case, $f$ is the **lower adjoint**, and $g$ is the **upper adjoint**. Lower adjoints preserve all suprema, and upper adjoints preserve all infima. In fact, each mapping $f$ between complete lattices that preserves all suprema is a lower adjoint; its upper adjoint $g$ is defined by $g(y) = \sup f^{-1}[\downarrow y]$. Dually, each mapping $g$ preserving all infima is an upper adjoint; its lower adjoint $f$ is defined by $f(x) = \inf g^{-1}[\uparrow x]$. The cumulative distribution function of a probability measure on $[0,1]$ and its upper adjoint given in the introduction are examples we’ll find relevant.

Finally, we need some detailed information about two Cartesian closed categories of domains. We let DOM denote that category of domains and Scott continuous maps; this is a full subcategory of DCPo, but it is not Cartesian closed. Nevertheless, DOM has several Cartesian closed full subcategories. Two of particular interest to us are the full subcategory BCD of countably based bounded complete domains. Precisely, a **bounded complete domain** is a domain in which every non-empty subset has a greatest lower bound. An equivalent statement is that every subset having an upper bound has a least upper bound.

A second CCC of domains in which we are interested is RB, the category of countably based domains that are retracts of bifinite domains, and Scott-continuous maps. The simplest way to make this definition precise is by saying that $D \in RB$ iff there is a countable family $\{f_n\}_n$ of deflations of $D$ satisfying $\mathrm{Id}_D = \sup_n f_n$, where $f_n: D \to D$ is a deflation if $f_n$ is Scott continuous and $f_n(D)$ is finite. Moreover,
since \( f_n(x) \ll x \) for each deflation (cf. [16] Lemma II-2.16), and \( f_n(D) \) is finite for each \( n \), the family \( B_D = \bigcup_n f_n(D) \) is a countable basis for \( D \).

Finally, note that \( \text{BCD} \) is a full subcategory of \( \text{RB} \), and that \( \text{BCD} \) and \( \text{RB} \) consist of coherent domains.

### 2.2 The probabilistic power domain

A **continuous valuation** on a domain \( D \) is a mapping \( \mu: \mathcal{O}(D) \rightarrow [0,1] \) from the family of Scott-open sets to the interval satisfying:

- (Strictness) \( \mu(\emptyset) = 0 \),
- (Modularity) \( \mu(U \cup V) + \mu(U \cap V) = \mu(U) + \mu(V) \), for \( U, V \in \mathcal{O}(D) \),
- (Scott continuity) If \( \{U_i\}_{i \in I} \subseteq \mathcal{O}(D) \) is directed, then \( \mu(\bigcup_i U_i) = \sup_i \mu(U_i) \).

Valuations are ordered pointwise: \( \mu \leq \nu \) if \( \mu(U) \leq \nu(U) \) for all \( U \in \mathcal{O}(D) \). We denote the set of valuations over a domain \( D \) with this order by \( \mathcal{V}_D \). This is often referred to as the probabilistic power domain of \( D \). We also denote by \( \mathcal{V}_1 D \) the valuations \( \mu \) satisfying \( \mu(D) = 1 \). As we will see, valuations correspond to sub-probability measures, while members of \( \mathcal{V}_1 D \) correspond to probability measures. The following is called the Splitting Lemma, it is fundamental for understanding the domain structure of \( \mathcal{V} D \) and of \( \mathcal{V}_1 D \).

**Theorem 2.2** (Splitting Lemma [20]) Let \( D \) be a domain and let \( \mu = \sum_{x \in F} r_x \delta_x \) and \( \nu = \sum_{y \in G} s_y \delta_y \) be simple valuations on \( D \). Then the following are equivalent:

1. \( \mu \leq \nu \in \mathcal{V} D \),
2. There is a family \( \{t_{x,y}\}_{(x,y) \in F \times G} \subseteq [0,1] \) of transport numbers satisfying:
   - \( r_x = \sum_{y \in G} t_{x,y} \) for each \( x \in F \),
   - \( \sum_{x \in F} t_{x,y} \leq s_y \) for each \( y \in G \),
   - \( t_{x,y} > 0 \Rightarrow x \leq y \).

Moreover, \( \mu \ll \nu \in \mathcal{V} D \) iff (i) \( \sum_{x \in F} r_x < s_y \) for each \( y \in G \) and (ii) \( t_{x,y} \) satisfies \( t_{x,y} > 0 \) implies \( x \ll y \in D \) for each \( x \in F, y \in G \).

This result can be used to show that, given a basis \( B_D \) for \( D \), the family \( \{\sum_{x \in F} r_x \delta_x \mid F \subseteq B_D, \sum_{x \in F} r_x < 1\} \) forms a basis for \( \mathcal{V} D \); in particular, each sub-probability measure is the directed supremum of simple measures way-below it, so \( \mathcal{V} D \) is a domain if \( D \) is one. Moreover, Jung and Tix [21] showed that \( \mathcal{V} D \) is a coherent domain if \( D \) is.

Our interest is in countably-based coherent domains, in which case we can refine the Splitting Lemma 2.2 and Lemma 2.1.

**Notation 2.3** We let \( \text{Dyad} \) denote the positive dyadic rationals in the unit interval. Likewise, \( \text{Dyad}_n = \{ \frac{k}{2^n} \mid 0 \leq k \leq 2^n \} \) is the set of dyadics with denominator \( 2^n \), for each \( n \geq 1 \).

**Proposition 2.4** Let \( D \) be a coherent domain with countable basis \( B_D \). Then:

1. \( \mathcal{V} D \) is a countably-based coherent domain with basis...
\[ B_{V_D} = \{ \sum_{x \in F} r_x \delta_x \mid F \subseteq B_D \text{ finite} \& \ r_x \in \text{Dyad} \ \forall x \in F \}. \]

Moreover, if \( \sum_{x \in F} r_x \delta_x \leq \sum_{y \in G} s_y \delta_y \in B_{V_D} \), then the family \( \{ t_{x,y} \}_{(x,y) \in F \times G} \) of transport numbers from the Splitting Lemma 2.2 satisfy \( t_{x,y} \in \text{Dyad} \) for all \( (x,y) \in F \times G \).

(ii) The family \( \forall_1 D = \{ \mu \in \text{VD} \mid \mu(D) = 1 \} \) is a countably-based coherent domain with basis
\[ B_{\forall_1 D} = \{ \sum_{x \in F} r_x \delta_x \mid \bot \in F \subseteq B_D \text{ finite}, \sum_x r_x = 1 \& \ r_x \in \text{Dyad} \ \forall x \in F \}. \]

Moreover, \( \sum_{x \in F} r_x \delta_x \leq (\text{resp.,} \leq) \sum_{y \in G} s_y \delta_y \) in \( \forall_1 D \) iff the transport numbers \( \{ t_{x,y} \} \) from 2.2 satisfy:
- \( r_x = \sum_y t_{x,y} \) for each \( x \in F \), and
- \( \sum_x t_{x,y} < (\text{resp.,} \leq) s_y \) if \( y \not= \bot \) for each \( y \in Y \).
In particular, \( r_\bot > (\text{resp.,} \geq) s_\bot \).

Finally, each \( \mu \in \text{VD} \) (respectively, \( \forall_1 D \)) is the supremum of a countable chain \( \mu_n \in B_{\text{VD}} \) (respectively, \( \forall_1 D \)).

**Proof.** (i): It is shown in [21] that \( \text{VD} \) is coherent if \( D \) is, and the Splitting Lemma 2.2 implies \( B_{\text{VD}} \) is a basis for \( \text{VD} \).

We next outline the proof of the second point – that the transport numbers \( t_{x,y} \) between comparable simple measures all belong to \( \text{Dyad} \) if the coefficients of the measures do. This follows from the proof of the Splitting Lemma 2.2 as presented in [20]: That proof is an application of the Max Flow – Min Cut Theorem to the directed graph \( G = (E, N) \) which has a “source node,” \( \bot \), connected by an outgoing edge of weight \( r_x \) to each “node” \( x \in F \), a “sink node,” \( \top \), with an incoming edge of weight \( s_y \) from each element \( y \in G \), and edges from \( x \in F \) to \( y \in G \) of large weight (say, 1), if \( x \leq y \).

A **flow** is an assignment \( f : E \rightarrow \mathbb{R}_+ \) of non-negative numbers to each edge so that \( f(u, v) \leq c(u, v) \) for nodes \( u, v \), where \( c(u, v) \) is the weight as defined above, and satisfying \( \sum_u f(u, v) = \sum_t f(v, t) \) for each node \( t \not= \bot, \top \). The **value** of a flow \( f \) is \( \text{val} f = \sum_u f(\bot, u) \), the total amount of flow out of \( \bot \) using \( f \). A **cut** is a partition of \( N = S \cup T \) with \( \bot \in S \) and \( \top \in T \). The value of a flow across the cut \( T \) is \( \sum_{(u,v) \in S \times T \cap E} f(u,v) \).

The Max Flow–Min Cut Theorem asserts that the maximum flow on a directed graph is equal to the minimum cut. It is proved by applying the Ford–Fulkerson Algorithm [6]. The algorithm starts by assigning the minimum flow \( f(u, v) = 0 \) for all edges \( (u, v) \in E \), and then iterates a process of selecting a path from \( \bot \) to \( \top \), calculating the residual capacity of each edge in the path, defining a residual graph \( G_f \), augmenting the paths in \( G_f \) to include additional flow, and then iterating. The result of the algorithm is the set of flows along edges across the cut, which are the transport numbers \( t_{x,y} \) in our case. Since the calculations of new edge weights involve only arithmetic operations, and since the dyadic rationals form a subsemigroup of \( \mathbb{R}_+ \), the resulting transport numbers \( t_{x,y} \) are dyadic rationals if the coefficients of the input distributions are dyadic. Moreover, if all the weights are rational, then the algorithm halts producing the maximum flow across the network.
(ii): Let \( \varphi: \mathbb{V}D \rightarrow \mathbb{V}_1D \) by \( \varphi(\mu) = \mu + (1 - \mu(D))\delta_1 \). Then \( \varphi \) is Scott continuous, since \( \sup_n \mu_n = \mu \) implies \( \sup_n \mu_n(D) = \mu(D) \). \( \varphi \) also is a projection: \( \varphi \circ \varphi = \varphi \). So, Theorem I-1.22 of \([16]\) implies \( \mathbb{V}_1D = \varphi(\mathbb{V}D) \) is continuous, and \( \mu \ll \nu \in \mathbb{V}_1D \) iff \( (\exists \mu' \in \mathbb{V}D) \mu \leq \varphi(\mu') \) \& \( \mu' \ll_{\mathbb{V}D} \nu \). This implies \( B_{\mathbb{V}_1D} \) is a basis for \( \mathbb{V}_1D \), since \( B_{\mathbb{V}D} \) is a basis is one for \( \mathbb{V}D \) by part (i). The final claim follows from these results and the characterization of the transport numbers \( \{t_{x,y}\} \) from Theorem 2.2.

### 2.3 The Jung-Tix Problem and the special case of chains

We now turn our attention to a longstanding problem in domain theory. The Jung-Tix Problem asks whether there are any Cartesian closed categories of domains for which the valuations monad \( \mathbb{V} \) is an endofunctor. We do not have an answer, but we can offer insight to the question. We are able to show that the probabilistic power domain of any complete chain is a continuous lattice. We include the result here because we discovered a proof of this result during our research on how to express Skorohod’s Theorem in domain-theoretic terms.

**Notation 2.5** Throughout this section, we assume \( D \) is a chain.

**Definition 2.6** Let \( \mu \) be a sub-probability measure on \( D \). The cumulative distribution function \( F_\mu: D \rightarrow [0,1] \) is defined by \( F_\mu(x) = \mu([x) \).  

**Proposition 2.7** For each \( \mu \in SProbD \), \( F_\mu \) preserves all infima.

**Proof.** Let \( \mu \) be a sub-probability measure. If \( x \leq y \in D \), then \( \downarrow x \subseteq \downarrow y \), so \( F_\mu(x) = \mu([x) \leq \mu([y) = F_\mu(y) \). So \( F_\mu \) is monotone, and since \( D \) is a chain, this means \( F_\mu \) also preserves finite infima. Now, any filtered set \( A \subseteq D \) is totally ordered because \( D \) is. Then \( \downarrow \inf A = \bigcap_{x \in A} \downarrow x \), and so 

\[
F_\mu(\downarrow \inf A) = F_\mu(\bigcap_{x \in A} \downarrow x) = \mu(\bigcap_{x \in A} \downarrow x) = \inf_{x \in A} \mu(\downarrow x) = \inf_{x \in A} F_\mu(x),
\]

where the next-to-last equality follows from the fact that, being a Scott-continuous valuation on \( D \), \( \mu \) preserves directed unions of open sets, so it preserves filtered intersections of closed sets, such as \( \{\downarrow x \mid x \in A\} \). This shows \( F_\mu \) also preserves filtered infima, and so it preserves all infima. \( \square \)

Since \( F_\mu \) preserves all infima and \( D \) is a continuous lattice, it follows that \( F_\mu \) is an upper adjoint, so it has a unique lower adjoint \( G_\mu: [0,1] \rightarrow D \) defined by 

\[
G_\mu(r) = \inf F_\mu^{-1}(\uparrow r).
\]

We denote this relationship by \( F_\mu \dashv G_\mu \).

We recall some facts about such adjoint pairs; for more detail, see Chapter 0 of \([16]\). First, each component of an adjoint pair \( f: L \rightarrow M \), \( g: M \rightarrow L \) with \( f \dashv g \) determines the other. The formula for \( G \) above shows how to define the lower adjoint, given an upper adjoint: \( g(x) = \inf f^{-1}(\uparrow x) \). Conversely, given a lower adjoint \( g \), the upper adjoint \( f \) is given by \( f(y) = \sup g^{-1}(\downarrow y) \). Upper adjoints preserve all infima, and lower adjoints preserve all suprema. Moreover, if \( f \dashv g \) and \( f' \dashv g' \), then \( f \leq f' \) iff \( g' \geq g \). Finally, the components earn their names because of the relationship \( f \circ g \geq 1_M \) and \( g \circ f \leq 1_L \).
Proposition 2.8 If $\mu$ is a sub-probability measure on $D$ with cumulative distribution function $F_\mu$, then the upper adjoint, $G_\mu: [0,1] \rightarrow D$ satisfies $G_\mu \lambda = \mu$, where $\lambda$ denotes Lebesgue measure.

Proof. If $x \in D$, then
\[
G_\mu \lambda(\downarrow x) = \lambda(G_\mu^{-1}(\downarrow x)) = \lambda(\downarrow F_\mu(x)) = F_\mu^{-1}(G_\mu(x)) = G_\lambda(x).
\]
Since $G_\mu \lambda$ and $\mu$ agree on Scott-closed sets, it follows that $G_\mu \lambda = \mu$. \qed

Theorem 2.9 If $D$ is a chain and $KD = \{\bot\}$, then $G \mapsto G \lambda: [[0,1] \rightarrow D] \rightarrow SProb D$ is an order-isomorphism. Therefore, $SProb D$ is a continuous lattice, and the same is true of $Prob D$ is a domain.

Proof. Each Scott-continuous map $G: [0,1] \rightarrow D$ preserves all suprema, since the domain $D$ is a chain. And each such map determines a sub-probability measure $G \lambda$. Then the cumulative distribution $F_{G \lambda}: D \rightarrow [0,1]$ satisfies $F_{G \lambda}(x) = G \lambda(\downarrow x) = \lambda(G^{-1}(\downarrow x)) = sup G^{-1}(\downarrow x)$. This means $F_{G \lambda}$ is the upper adjoint of $G$. Since upper and lower adjoints uniquely determine one another, the mapping $G \mapsto G \lambda$ has an inverse sending $\mu$ to the lower adjoint of $F_\mu$.

For the order structure, suppose $G \leq G'$ in $[[0,1] \rightarrow D]$. We show $G \lambda \leq G' \lambda$: Then given $x \in D$ and $r \in [0,1]$, if $G'(r) \leq x$, then $G(r) \leq x$; said another way, $G^{-1}(\downarrow x) \subseteq G'^{-1}(\downarrow x)$, so
\[
G' \lambda(\downarrow x) = \lambda(G'^{-1}(\downarrow x)) = sup G'^{-1}(\downarrow x) \leq sup G^{-1}(\downarrow x) = \lambda(G^{-1}(\downarrow x)) = G \lambda(\downarrow x).
\]
If $x = \bot$, $G \lambda(\uparrow x) = G \lambda(D) \leq G' \lambda(D) = G' \lambda(\uparrow x)$. On the other hand, since $KD = \{\bot\}$, then $x > \bot$ implies $D = \downarrow x \cup \uparrow x$, so we have
\[
G \lambda(\uparrow x) = G \lambda(D) - G \lambda(\downarrow x) \leq G \lambda(D) - G' \lambda(\downarrow x) \leq G' \lambda(D) - G' \lambda(\uparrow x) = G' \lambda(\uparrow x).
\]
Since $D$ is a chain, every Scott-open set has the form $\uparrow x$ for some $x \in D$, so $G \lambda \leq G' \lambda$.

Conversely, if $\mu \leq \nu$, then $\mu(\downarrow x) \geq \nu(\downarrow x)$ by the same argument we used above, so $F_\mu(x) = sup \mu(\downarrow x) \geq sup \nu(\downarrow x) = F_\nu(x)$. It follows that $G_\mu \leq G_\nu$ from our remarks about adjoint pairs.

Thus, the correspondence $G \mapsto G \lambda$ is an order-isomorphism. Since $[0,1]$ and $D$ are continuous lattices, they are both bounded complete domains, so $[[0,1] \rightarrow D]$ is a bounded complete domain. But $x \mapsto \top$ is the largest element of $[[0,1] \rightarrow D]$, so this is a continuous lattice. It follows that $SProb D$ is a continuous lattice as well.

For the final claim, the mapping $\mu \mapsto \mu + (1 - \mu(D))\delta_\bot: SProb D \rightarrow Prob D$ is a closure operator that preserves directed sups, and the image of a continuous lattice under such a closure operator is a continuous lattice (cf. [16], Definition 0-2.10ff.). \qed

2.4 Valuations versus sub-probability measures

It is straightforward to show that each Borel sub-probability measure on a domain $D$ restricts to a Scott-continuous valuation on the Scott-open sets of $D$. The converse,
that each Scott-continuous valuation on a dcpo extends to a unique Borel sub-
probability measure was shown by Alvarez-Manilla, Edalat and Saheb-Djorhomi [2].

The next step is to link the order-structure of \( VD \) to the family \( \text{SProb} \, D \) of
sub-probability measures on \( D \), and this requires the next result. We recall that
a simple sub-probability measure on a space \( X \) is a finite convex sum \( \sum_{x \in F} r_x \delta_x \),
where \( F \subseteq X \) is finite, \( r_x \geq 0 \) for each \( x \in F \), and \( \sum_{x \in F} r_x \leq 1 \). We also recall that
the real numbers, \( \mathbb{R} \), are a continuous poset whose Scott topology has the intervals
\((a, \infty)\) as a basis, and whose Lawson topology is the usual topology.

**Proposition 2.10** Let \( D \) be a coherent domain, and let \( \mu, \nu \) be sub-probability mea-
sures on \( D \). Then the following conditions are equivalent:

(i) \( \mu \leq \nu \in VD \).

(ii) For each Scott-continuous map \( f : D \rightarrow \mathbb{R}_+ \), \( \int f \, d\mu \leq \int f \, d\nu \).

(iii) For each monotone Lawson-continuous \( f : D \rightarrow \mathbb{R}_+ \), \( \int f \, d\mu \leq \int f \, d\nu \).

**Proof.** We show the result for simple measures, which then implies it holds for all
measures since \( VD \) is a domain – so its partial order is (topologically) closed – in
which the simple measures are dense.

So, suppose \( \mu = \sum_{x \in F} r_x \delta_x \) and \( \nu = \sum_{y \in G} s_y \delta_y \) are simple measures on \( D \).

(i) implies (ii): Suppose that \( \mu \leq \nu \in VD \). If \( f : D \rightarrow \mathbb{R}_+ \), then \( \int f \, d\mu = \sum_{x \in F} r_x \cdot f(x) = \sum_{x \in F, y \in G} t_{x,y} \cdot f(x) \)
and \( \int f \, d\nu = \sum_{y \in G} s_y \cdot f(y) \). Since \( \mu \leq \nu \), there are \( t_{x,y} \in [0,1] \)
guaranteed by the Splitting Lemma 2.2, and so

\[
\int f \, d\mu = \sum_{x \in F} r_x \cdot f(x) = \sum_{x \in F, y \in G} t_{x,y} \cdot f(x)
\leq \sum_{x \in F, y \in G} t_{x,y} \cdot f(y) \leq \sum_{y \in G} s_y \cdot f(y) = \int f \, d\nu,
\]

where the first inequality follows from the facts that \( t_{x,y} > 0 \) implies \( x \leq y \) and \( f \)
is monotone. This shows (i) implies (ii).

(ii) implies (iii): Since monotone Lawson continuous maps are Scott continuous,
this is obvious.

(iii) implies (i): Let \( U \) be a Scott-open subset of \( D \), and let \( H = (F \cup G) \setminus U \). Using
the facts that \( D \) is coherent, so its Lawson topology is compact Hausdorff, and that
\( H \) is finite, we define a family \( \{ U_d \mid d \in \text{Dyad} \} \) of Scott-open upper sets indexed by
\( \text{Dyad} \), the dyadic numbers in \([0,1] \) as follows: We let \( U_0 = D \setminus \downarrow H, U_1 = U \), and for
\( d < d' \), we recursively choose \( U_d \supseteq \overline{U_{d'}} \), the Lawson-closure of \( U_{d'} \). Then define a
mapping

\[
f : D \rightarrow [0,1] \text{ by } f(x) = 0 \text{ if } x \in \downarrow H, \text{ and otherwise } f(x) = \inf \{ d \mid x \in U_d \}.
\]

Since the family \( \{ U_d \} \) consists of Scott-open sets satisfying \( U_d \supseteq \overline{U_{d'}} \) for \( d < d' \),
this mapping is monotone, and the standard Urysohn Lemma argument (cf. Theorem
33.1 [28]) shows it is Lawson continuous. So, \( \int f \, d\mu \leq \int f \, d\nu \) by assumption.
Since $\mu$ and $\nu$ are simple, $\int f d\mu = \sum_{x \in F \setminus H} r_x : f(x)$ and $\int f d\nu = \sum_{y \in G \setminus H} s_y : f(y)$. By construction, $(F \cup G) \setminus H \subseteq U = U_1$, so
$$\sum_{x \in F \setminus H} r_x : f(x) = \sum_{x \in F \setminus H} r_x = \mu(U), \quad \text{and} \quad \sum_{y \in G \setminus H} s_y : f(y) = \sum_{y \in G \setminus H} s_y = \nu(U),$$
and so $\mu(U) = \int f d\mu \leq \int f d\nu = \nu(U)$, as required. \qed

Proposition 2.10 tells us we can realize the domain order structure of $\forall D$ on $SProb D$ using the classical approach of integration against functions $f : D \rightarrow \mathbb{R}_+$. Put another way, the mapping $\psi : \forall D \rightarrow SProb D$ that realizes a valuation as a sub-probability measure is not only a bijection, but an order isomorphism if one equips $SProb D$ with the order described in the Proposition. It should be noted the same is true for $\forall D$, with essentially the same proof. Moreover, the isomorphisms are also homeomorphisms:

**Theorem 2.11** \cite{14,7} If $D$ is a coherent domain, then the Lawson topology on $SProb D$ is the weak topology. The same holds for $\forall D$.

**Remark 2.12** Thus, Theorem 2.11 says we can regard $\forall D$ and $SProb D$ as one and the same, from a domain-theoretic perspective, and the same holds for $\forall_1 D$ and $\forall D$. We sometimes will “confuse” these two views of valuations / sub-probability measures without explicitly noting the identification.

### 3 Domain Mappings from the Cantor Tree

The **Cantor tree** is the family $\mathbb{C}T = \{0,1\}^* \cup \{0,1\}^\omega$ of finite and infinite words over $\{0,1\}$ in the prefix order. Equivalently, $\mathbb{C}T$ is the full rooted binary tree which is directed complete, and since it is countably based, this means every directed supremum can be achieved as the supremum of an increasing countable chain. $\mathbb{C}T$ will play the role of the unit interval in our approach to generalizing Skorohod’s Theorem to the domain setting. For that, we need some preliminary definitions.

An **antichain** is a non-empty subset $A \subseteq \mathbb{C}T$ satisfying $a, b \in A$ implies $a$ and $b$ do not compare in the prefix order.

**Notation 3.1** We establish some notation for what follows:

(i) We let $\mathcal{C}_n \simeq 2^n$ be the set of $n$-bit words in $\mathbb{C}T$, which forms an antichain. Recall that there is a well-defined retraction mapping $\pi_n : \mathcal{C} \rightarrow \mathcal{C}_n$ from the Cantor set onto $\mathcal{C}_n$ sending each infinite binary word to its $n$-bit prefix. In addition, if $m \leq n$, then there is a map $\pi_{m,n} : \mathcal{C}_m \rightarrow \mathcal{C}_n$ that sends each $n$-bit word to its $m$-bit prefix.

(ii) Since $\mathcal{C}_n$ is finite, the set $\downarrow \mathcal{C}_n \subseteq \mathbb{C}T$ is Scott closed. Then both $\pi_n$ and $\pi_{m,n}$ extend to mappings $\hat{\pi}_n : \mathbb{C}T \rightarrow \downarrow \mathcal{C}_n$ and $\hat{\pi}_{m,n} : \downarrow \mathcal{C}_n \rightarrow \downarrow \mathcal{C}_m$ that send each element of $\downarrow \mathbb{C}T$ to its largest prefix in $\downarrow \mathcal{C}_n$, respectively its largest prefix in $\downarrow \mathcal{C}_m$. Note that $\hat{\pi}_m = \hat{\pi}_{m,n} \circ \hat{\pi}_n$ if $m \leq n$.

(iii) The projection $\pi_n$ has a corresponding embedding $\iota_n : \mathcal{C}_n \rightarrow \mathcal{C}$ sending an $n$-bit word to the infinite word all of whose coordinates $m > n$ are 0. Then
\[ \pi_n \circ \tau_n = 1_{C_n} \text{ and } \tau_n \circ \pi_n \leq 1_C \text{ form an embedding-projection pair, where we order } C \simeq 2^\omega \text{ in the lexicographic order.} \]

(iv) The set \( C_n \) of \( n \)-bit words also can be given the lexicographic order. Then each dyadic rational \( r \in [0, 1] \) that can be expressed as \( r = \frac{k}{2^n} \), the interval in \( C_n \) from 0 to \( r \), or, equivalently, the first \( k \) \( n \)-bit words.

Moreover, each sequence of such dyadics, \( r_1, \ldots, r_k \) whose sum is at most 1 can be expressed as successive intervals, \( r_1 = [0, \ldots, k_1], r_2 = [k_1 + 1, \ldots, k_1 + k_2], \ldots \), etc.

The proof of the next result relies on a version of Hall’s Marriage Problem [18]. The original version concerns a bipartite graph \( G = (X + Y, E) \), where \( X \) and \( Y \) are the disjoint sets of nodes and all edges in \( E \) connect a node of \( X \) to one of \( Y \). A matching is a subset \( M \subseteq E \) satisfying each node of \( X \) has at most one edge in \( M \), and likewise for \( Y \), and no two edges in \( M \) share any common nodes. A perfect matching is one where every node of \( X \) and every node of \( Y \) has an edge in \( M \). Hall’s Marriage Problem states that there is a perfect matching iff, for each subset \( S \subseteq X \) there are at least \( |S| \) edges in \( E \) from some node of \( S \) to some node of \( Y \).

The generalization we need is for the case of matching each node of \( X \) to \( k \) nodes in \( Y \) so that no two edges in \( M \) share any nodes of \( Y \) (and so each node of \( X \) has edges to \( k \) distinct nodes in \( Y \) none of which is shared with any other node of \( X \)). The generalization of Hall’s Marriage Problem states that such a \( k \) matching exists – i.e., there is a subset \( M \subseteq E \) satisfying each node of \( X \) has at least \( k \) edges in \( M \), and each node of \( Y \) has at most one edge in \( M \) – iff, for each subset \( S \subseteq X \), there are at least \( k \cdot |S| \) edges in \( E \) connecting some node of \( S \) to a node of \( Y \). The generalization follows from the original version by first making \( k \) copies of each node of \( X \), duplicating the edges in \( E \) for each of these new nodes, applying the original version, and then collapsing the resulting perfect matching back to the original graph \( G \).

The following is key to our results:

**Proposition 3.2** Let \( D \) be a domain and let \( \mu = \sum_{x \in F} r_x \delta_x \leq \sum_{y \in G} s_y \delta_y = \mu' \) be simple probability measures on \( D \) with \( r_x, s_y \in Dyad \) for every \( x \) and \( y \). Suppose further that \( f_m: C_m \rightarrow D \) satisfies \( f_m \circ \nu_m = \mu \), where \( \nu_m \) is normalized counting measure on \( C_m \). Then there are \( n > m \) and \( f_n: C_n \rightarrow D \) satisfying \( f_n \circ \nu_n = \mu' \) and \( f_m \circ \pi_{m,n} \leq f_n \).

**Proof.** Since \( \mu \leq \mu' \) are simple probability measures, the Splitting Lemma (Theorem 2.2) implies there are transport numbers \( \{t_{x,y}\} \) satisfying \( r_x = \sum_y t_{x,y} \) and \( \sum_x t_{x,y} = s_y \) for all \( x, y \). Moreover, since \( r_x, s_y \in Dyad \) for all \( x, y \), the transport numbers \( t_{x,y} \in Dyad \) as well, as a proof similar to that for Proposition 2.4 shows. We then choose \( n > m \) such that \( r_x, s_y, t_{x,y} \in Dyad_n \), for all \( x, y \), where we recall \( Dyad_n = \{ \frac{k}{2^n} | 0 \leq k \leq 2^n \} \) (cf. 2.3). Then

\[
\mu = \sum_{x \in F} r_x \delta_x \text{ with } r_x = \frac{k_x}{2^n}, \text{ so } \mu = \sum_{x \in F} \frac{k_x}{2^n} \delta_x
\]

and

\[
\mu' = \sum_{y \in G} s_y \delta_y \text{ with } s_y = \frac{k_y}{2^n}, \text{ so } \mu' = \sum_{y \in G} \frac{k_y}{2^n} \delta_y.
\]

Since \( t_{x,y} \in Dyad_n \) for each \( x, y \), it follows that, for each \( x \in F \), the family \( t_{x,y} \)
distributes the mass \( r_x = \frac{k_x}{2^n} \) associated to \( \delta_x \) in \( \mu \) to \( \frac{k_x}{2^n} \) of the mass in \( \mu' \).

Since \( f_m: C_m \to D \) satisfies \( f_m \star \nu_m = \mu \), we also have \( \mu = \sum_{x \in F} \frac{k_x' \delta_x}{2^n} \), and so \( \frac{k'_x}{2^n} = \frac{k_x}{2^n} \) for each \( x \in F \). That is, \( 2^{n-m}k'_x = k_x \) for each \( x \in F \). So, each \( i \in C_m \) is sent via \( i \mapsto f_m : \frac{1}{2^n} \delta_{f(i)} \mapsto y \) to \( \frac{2^{n-m}}{2^n} \) of the mass of \( \mu' \).

We define a bipartite graph with node sets \( X, Y \) where \( X = C_m \) and \( Y = \bigcup_{y \in G} \{(i, \delta_y) \mid 0 < i \leq k_y\} \), and whose edges are

\[
E = \{(i, (j, \delta_y)) \mid t_{f(i), y} \text{ sends mass at } \delta_{f(i)} \text{ to } (j, \delta_y)\}.
\]

Then each node \( i \in X = C_m \) has \( 2^{n-m} \) edges incident to it in \( E \), and each node \( (j, \delta_y) \in Y \) has exactly one edge incident to it. By construction, for each \( S \subseteq X \), there are \( 2^{n-m} \cdot |S| \) edges in \( E \) from some node in \( S \) to some node in \( Y \). The generalization of Hall’s Marriage Problem described above then implies there is a \( 2^{n-m} \)-matching, and since \( \mu, \mu' \) are probability measures, \( s_y = \sum_{x} t_{x,y} \) for each \( y \in G \), so there is a total function \( \rho: Y \to X \) satisfying \( |\rho^{-1}(i)| = 2^{n-m} \) for each \( i \in X \).

On the other hand, the projection \( \pi_{m,n}: C_n \to C_m \) sends \( 2^{n-m} \) \( n \)-bit words to each element of \( C_m \), so for each \( i \in C_m \) there is a bijection \( b_i: \pi_{m,n}^{-1}(i) \to \rho^{-1}(i) \). Taken together, these \( b_i \)s define a map \( g_n: C_n \to Y \), and if we let \( p_Y: Y \to D \) by \( p_Y(j, \delta_y) = y \) and define \( f_n = p_Y \circ g_n \), then \( f_m \circ \pi_{n,m} \leq f_n \) by construction. That \( f_n \star \nu_n = \mu' \) follows from the fact that \( |f_n^{-1}(y)| = \frac{k_y}{2^n} = s_y \) for each \( y \in G \), again by construction. \( \Box \)

For our next result, we need some information about the weak topology on \( SProb D \). The result we need follows from the Portmanteau Theorem 4.3 (cf., e.g., [5]), and a proof can be found as Corollaries 15 and 16 in [7]:

**Theorem 3.3** Let \( D \) be a countably based coherent domain endowed with the Borel \( \sigma \)-algebra. Then the weak topology on \( SProb D \) is the same as the Lawson topology on \( SProb D \) when viewed as a family of valuations.

Moreover, for a family \( \mu_n, \mu \in SProb D \), the following are equivalent:

(i) \( \mu_n \to w \mu \)

(ii) Both of the following hold:

* \( \limsup_n \mu_n(E) \leq \mu(E) \) for all finitely generated upper sets \( E \subseteq D \).
* \( \liminf_n \mu_n(O) \geq \mu(O) \) for all Scott-open sets \( O \subseteq D \).

(iii) \( \liminf_n \mu_n(O) \geq \mu(O) \) for all Lawson-open sets \( O \subseteq D \).

**Corollary 3.4** If \( D \) is a countably based coherent domain, then the weak topology on \( Prob D \) coincides with the Lawson topology.

**Proof.** The weak topology on \( Prob D \) is compact Hausdorff since the Lawson topology on \( D \) is compact Hausdorff. Then the same argument used in the proof of Corollaries 15 and 16 of [7] shows that the weak topology on \( Prob D \) is finer than the Lawson topology. Since the latter is Hausdorff, the topologies coincide. \( \Box \)

For the next result, we identify \( C \) with the Cantor set of maximal elements in \( CT \), the Cantor tree, and recall that \( \nu_C \) denotes Haar measure on \( C \) viewed as a
countable product of two-point groups.

**Theorem 3.5** Let $D$ be an RB domain with countable basis, $B_D = \bigcup_k d_k(D)$, where $\cdots d_k \leq d_{k+1} \cdots$ is a countable sequence of deflations on $D$ with $\sup_k d_k = \text{Id}_D$ (see the discussion of RB preceding Subsection 2.2).

(i) If $\mu \in \text{Prob}(D)$, then there is a Scott-continuous map $X : \mathbb{C}T \rightarrow D$ satisfying $X_* \nu_C = \mu$.

(ii) Furthermore, if $\mu_n \rightarrow w \mu$ are probability measures on $D$ converging to $\mu$ in the weak topology, then there are Scott-continuous maps $X_n : \mathbb{C}T \rightarrow D$ satisfying $X_n_* \nu_C = \mu_n$ and $X_n \rightarrow X$ pointwise wrt to the Scott topologies.

(iii) Finally, if $X(x) \in \text{Max} D$, then $X_n(x) \rightarrow X(x)$ in the Lawson topology on $D$.

**Proof.** Since $D$ is in RB, the sequence of deflations $d_k : D \rightarrow D$ satisfy $d_k \preccurlyeq d_{k+1}$ for all $k$, and $\sup_k d_k = \text{Id}_D$. Without loss of generality, we assume $d_0 \equiv \bot_D$. Then $\text{Prob} d_k = d_k \circ \text{Proj} D$ projects $\text{Prob} D$ onto $\text{Prob} d_k(D)$. Since $\text{Prob}$ is locally continuous, it follows that $\sup_k d_k = \text{Id}_{\text{Prob} D}$.

(i) To prove (i), we apply Proposition 3.2 recursively. Let $\mu \in \text{Prob} D$, and consider the sequence $\{d_k \mu \mid k \geq 0\}$. Note that $d_k \mu = \sum_{x \in d_k D} d_x \delta_x$ is simple, since $d_k$ is a deflation. We define $f_0 : C_0 \rightarrow D$ by $f_0 \equiv \bot_D$. Then $f_0 \mu = 0 = \delta_\bot = d_0 \mu$.

For the inductive step, assume there are $m_k \geq k$ and $f_k : C_{m_k} \rightarrow D$ satisfying $f_k_* \nu_{m_k} = \sum_{x \in F_k} r_x \delta_x \leq \sum_{y \in G_k} s_y \delta_y = d_k \mu$ with $r_x \in D_yad$ for each $x \in F_k \subseteq B_D$, and

- $\exists \phi_k : F_k \rightarrow G_k$ a bijection with $x \preccurlyeq \phi_k(x)$ and $s_{\phi_k(x)} - r_x < 2^{-m_k}$ for all $x \in F_k$.

Since $f_k_* \mu \leq d_k \mu \leq d_{k+1} \mu$, Proposition 2.4(ii) implies there is a simple measure $\sum_{y \in G_{k+1}} s_y \delta_y = d_{k+1} \mu$ with $r_x \in D_yad$, $F_{k+1} \subseteq B_D$ and $f_k_* \nu_{m_k} \preccurlyeq \sum_{x \in F_{k+1}} r_x \delta_x$, and by making judicious choices, we may assume also satisfies:

- $\exists \phi_{k+1} : F_{k+1} \rightarrow G_{k+1}$ a bijection with $x \preccurlyeq \phi_{k+1}(x)$ and $s_{\phi_{k+1}(x)} - r_x < 2^{-m_{k+1}}$ for all $x \in F_{k+1}$.

Proposition 3.2 implies there are $m_{k+1} > k, m_k, |G_{k+1}|$ and $f_{k+1} : C_{m_{k+1}} \rightarrow D$ satisfying $f_{k+1} \nu_{m_{k+1}} = \sum_{y \in G_{k+1}} s_y \delta_y$ and $f_k \circ \pi_{m_k, m_{k+1}} \leq f_{k+1}$.

The resulting sequence $f_k : C_{m_k} \rightarrow D$ satisfies $\{f_k \nu_{m_k}\}$ is increasing with $\sup_k f_k \nu_{m_k} = \sup_k d_k \mu = \mu$ because of the bulleted items in the definition above.

We now extend the sequence $f_k$ to an increasing sequence of Scott-continuous maps $\hat{f}_k : \mathbb{C}T \rightarrow D$ satisfying $\hat{f}_k \nu_C = f_k \nu_{m_k}$ for each $k$, which implies the supremum $\hat{f} : \mathbb{C}T \rightarrow D$ satisfies $\hat{f}_* \nu_C = \sup_k \hat{f}_k \nu_C = \mu$.

We let $f_0 = f_0 \circ \pi_0$, where $\pi_0 : \mathbb{C}T \rightarrow C_0$ is the projection.

Assume $\hat{f}_k : \mathbb{C}T \rightarrow D$ is defined so that $\hat{f}_k |C_{m_l} = f_l \circ \pi_{m_l, m_k}$ for each $l \leq k$.

---

1 Because of the many functions involved in the proof, we revert to the probability theory approach of denoting random variables by capital letters.
Define \( \tilde{f}_{k+1}(x) = \begin{cases} f_{k+1} \circ \pi_{k+1}(x) & \text{if } x \in \uparrow \mathcal{C}_{m_{k+1}}, \\ \tilde{f}_k \mid \mathcal{C}_{m_k} \circ \pi_{m_k} & \text{otherwise.} \end{cases} \) Then \( \tilde{f}_k \leq \tilde{f}_{k+1} \) for each \( k \) and

\[
\tilde{f}_k \ast \nu_C = (f_k \circ \pi_k) \ast \nu_C = f_{k \ast} (\pi_k \ast \mu_C) = f_{k \ast} \nu_{C_{m_k}} \text{ for each } k, \text{ so } X = \sup_k \tilde{f}_k : \mathbb{C}T \to D \text{ is Scott continuous and satisfies } X \ast \mu_C = \sup_k \tilde{f}_k \ast \nu_C = \mu.
\]

(ii) For part (ii), we appeal to Lemma 2.1 to choose a countable descending chain of Scott-open sets \( U_k \) satisfying \( \bigcap_k U_k = \uparrow \mu \). We also assume without loss of generality that \( U_0 = \uparrow \delta_\bot = \text{Prob} D \).

Fix \( n \), and note that \( f_{k \ast} \nu_{C_{m_k}} \ll \mu_n \) implies \( d_{k \ast} \mu_n \) is in \( U_k \) eventually in \( n \). Assume the argument from part (i) has been used to construct a sequence \( \tilde{g}^n : \mathbb{C}T \to D \) with \( \sup_l \tilde{g}^n \nu_C = \mu_n \), and let \( X_n = \sup_l \tilde{g}^n \). Then \( X_n : \mathbb{C}T \to D \) is Scott continuous and \( X_n \ast \mu_C = \mu_n \).

It remains to show \( X_n \to X \) pointwise wrt the Scott topologies. Let \( x \in \mathbb{C}T \) with \( X(x) \in U \subseteq D \) Scott open. Since \( \sup_k \tilde{f}_n = X \), there is some \( K \) with \( \tilde{f}_k(x) \in U \) for \( k \geq K \). By construction, given \( k \geq K \), there is some \( N' \) with \( \tilde{f}_k \leq \tilde{g}^n \leq X_n \) for each \( n \geq N' \), from which it follows that \( X_n(x) \in U = \uparrow U \) for \( n \geq N' \).

(iii) For the last claim, we note that \( D \) is in \( \text{RB} \) implies the relative Lawson topology and the relative Scott topology agree on \( \text{Max} D \). If \( X(x) \in U \setminus \uparrow F \) is a Scott-open set in \( D \), then there is a Scott-open set \( U' \subseteq U \setminus \uparrow F \) with \( x \in U' \).

Then \( X_n(x) \in U' \) eventually by part (ii), so every limit point of \( \{ X_n(x) \}_n \) is in \( U \). Since \( D \) is Lawson compact, the sequence \( \{ X_n(x) \}_n \) has limit points. Since \( U \setminus \uparrow F \) is an arbitrary Lawson-open set containing \( X(x) \), it follows that every limit point of \( \{ X_n(x) \}_n \) is in \( \bigcap_{X(x) \in U \setminus \uparrow F} U = \uparrow X(x) = \{ X(x) \} \), the last equality following from the fact that \( X(x) \) is maximal. Thus, \( X_n(x) \to X(x) \) in the Lawson topology. \( \square \)

The next result follows from Theorem 3.5 by simply considering the law \( \mu = X \ast \nu_C \).

**Corollary 3.6** If \( D \) is a countably based \( \text{RB} \) domain and \( X : \mathcal{C} \to D \) is a random variable, then there is a Scott-continuous map \( f : \mathbb{C}T \to D \) satisfying \( f \ast \nu_C = X \ast \nu_C \).

**Remark 3.7** Theorem 3.5 still holds if we weaken the hypothesis to assuming only that \( D \) is a countably based coherent domain. The only change in the proof is that the simple measures \( d_{k \ast} \mu \) approximating each measure \( \mu \) must be chosen measure-by-measure, since coherent domains don’t have a sequence of deflations such as the \( d_{k \ast} \) that are available for \( \text{RB} \)-domains. But having a countable basis ensures that each measure \( \mu \) has a countable sequence of simple measures \( \sigma_n \) with \( \mu = \sup_n \sigma_n \) and that also satisfy \( \sigma_n \ll \sigma_{n+1} \ll \mu \). One can use these in place of the “uniformly chosen” simple measures \( d_{k \ast} \mu \). The remainder of the proof proceeds along the same lines, with the simple measures with dyadic coefficients chosen by recursion.

The obvious question is whether Theorem 3.5 extends to subprobability measures. In fact, it does, as we now outline. Given an \( \text{RB} \) domain \( D \), then \( D_\bot \), the lift of \( D \), also is in \( \text{RB} \), and it has a countable basis if \( D \) does. Then, the embedding \( D \hookrightarrow D_\bot \) allows us to define an embedding \( e : \text{SProb} D \to \text{Prob} D_\bot \) by
\(e(\mu) = \mu + (1 - \mu(P))\delta\). Theorem 3.5 then implies there is a Scott-continuous map \(f: \mathbb{CT} \to D\) with \(f_*\nu_C = e(\mu)\). Note that \(f^{-1}(\{\perp\}) = C_f\) is Scott closed, and the restriction \(f' \equiv f|_{D\setminus\{\perp\}}: \mathbb{CT}\setminus C_f \to D\) also is Scott continuous, and it’s routine to show that \(f'_*\nu_C = \mu\). Likewise, any sequence of subprobability measures \(\mu_n \to_w \mu\) converging weakly in \(\text{SProb}\ D\) satisfies the property that there are Scott-continuous partial maps \(f'_n: \mathbb{CT}\setminus C_n \to D\) with \(f'_n\nu_C = \mu_n\). Theorem 3.5(ii) shows \(f_n \to f\) in the Scott topologies, and this in turn implies that \(f'_n \to f'\) in the Scott topologies. We summarize this as follows:

**Corollary 3.8** Let \(D\) be a countably based RB domain. Then:

- If \(\mu \in \text{SProb}\ D\), there is a Scott-continuous map \(f: \mathbb{CT}\setminus C_f \to D\) satisfying \(f_*\nu_C = \mu\), where \(C_f\) is Scott closed.
- If \(\mu_n \to_w \mu\) is a sequence of subprobability measures on \(D\) converging weakly to the subprobability measure \(\mu\), then there is a sequence of Scott-continuous maps \(f_n: \mathbb{CT}\setminus C_n \to D\) satisfying \(f_n\nu_C = \mu_n\), where \(C_n\) is Scott closed, for each \(n\), and \(f_n \to f\) in the Scott topologies.

### 4 Domains, Polish spaces and Skorohod’s Theorem

In this section we present the principal application of our main results. We begin with the necessary background about Polish spaces and random variables, and then outline a representation theorem for Polish spaces topologically embedded as \(G_\delta\)subsets of a domain in the relative Scott topology. With this in place, we focus on probability measures and derive a domain-theoretic proof of Skorohod’s Theorem.

**Definition 4.1** A Polish space is a completely metrizable separable topological space. I.e., \(P\) is Polish iff \(P\) is homeomorphic to a complete metric space that has a countable dense subset.

Polish spaces figure prominently in probability theory \([5]\), as well as in descriptive set theory \([22]\). As we commented in the last section, one approach to probability theory \([5]\) begins with metric spaces, and Polish spaces are where the deepest results hold. Since our results all involve separable topological spaces, the measurable sets in all cases are the Borel sets. So, we use the term Borel set instead of measurable set.

The appropriate mappings in probability theory are random variables – measurable maps \(X: (P, \Sigma_P, \mu) \to (S, \Sigma_S)\), where \((P, \Sigma_P, \mu)\) is a probability space, \((S, \Sigma_S)\) is a measure space, and \(X^{-1}(A) \in \Sigma_P\) for each \(A \in \Sigma_S\). If \(X: P \to S\) is a random variable, then the push forward of \(\mu\) by \(X\) is the measure \(X_*\mu\) defined by \(X_*\mu(A) = \mu(X^{-1}(A))\) for all measurable sets \(A \subseteq S\). Equivalently, a function \(f: S \to \mathbb{R}\) is \(X\)-\(\mu\)-integrable iff \(f \circ X\) is \(\mu\)-integrable, and in this case, \(\int f dX_*\mu = \int f \circ X d\mu\). We now show how our results from Section 3 can be applied to Polish spaces that can be topologically embedded as the maximal elements of a domain in the relative Scott topology. We begin by extending some results about probability measures on Polish spaces to sub-probability measures.
Definition 4.2 Let $\mu$ be a probability measure on a space $X$. A Borel set $A \subseteq X$ is a $\mu$-continuity set if $\mu(\overline{A} \setminus A) = 0$. Since $\overline{A}$ is closed and $A$ is Borel, $\overline{A} \setminus A$ is Borel.

Theorem 4.3 (Portmanteau Theorem) Let $X$ be a Polish space, and let $\text{SProb} X$ denote the family of sub-probability measures on $X$ in the weak topology, and let $\mu_n, \mu \in \text{SProb} X$. Then the following are equivalent:

(i) $\mu_n \rightarrow_w \mu$ in the weak topology.
(ii) $\int fd\mu_n \rightarrow \int fd\mu$ for all bounded uniformly continuous $f : X \rightarrow \mathbb{R}$.
(iii) $\liminf_n \mu_n(O) \geq \mu(O)$ for all open sets $O \subseteq X$.
(iv) $\limsup_n \mu_n(F) \leq \mu(F)$ for all closed sets $F \subseteq X$.
(v) $\mu_n(A) \rightarrow \mu(A)$ for all $\mu$-continuity sets $A$.

Proof. Theorem 2.1 of [5] shows these conditions are equivalent if $\mu_n, \mu$ are probability measures on a metric space. We assume the metric $d$ on $X$ satisfies $\text{diam } X < 1$ by normalization, if necessary, and then create a new space $X' = X \cup \{\ast\}$, where $\ast$ is an element not in $X$. If $d$ is the metric on $X$, we extend $d$ to $X'$ by setting $d(\ast, x) = d(x, \ast) = 1$ for all $x \in X$. This makes $X'$ into a Polish space, and there is an embedding $e : \text{SProb} X \hookrightarrow \text{Prob} X'$ by $e(\mu) = \mu + (1 - \mu(X))\delta_*$. Then the conditions (i) – (v) are equivalent for $e(\text{SProb} X)$. But $\mu = e(\mu)|_X$ for all $\mu \in \text{SProb} X$, and $X$ is clopen in $X'$, so they also are equivalent for $\text{SProb} X$. $\square$

For the next result, recall that a measure $\mu$ on a space $X$ is concentrated on the Borel set $A \subseteq X$ if $\mu(X \setminus A) = 0$.

Proposition 4.4 Let $X$ be a Polish space with a topological embedding $\iota : X \rightarrow D$ as a $G_\delta$-subset of a countably-based domain, $D$. Define $e : \text{SProb} X \rightarrow \text{SProb} D$ by $e(\mu) = \iota_* \mu$. Then:

(i) $e$ is one-to-one.
(ii) $e(\text{SProb} X) = \{\mu \in \text{SProb} D \mid \mu \text{ concentrated on } \iota(X)\}$, and $j : e(\text{SProb} X) \rightarrow \text{SProb} X$ by $j(\nu) = \nu \circ \iota$ is inverse to $e$.
(iii) $e(\text{Prob} X) \subseteq \text{Prob} D = \text{Max } \text{SProb} D$ is a Borel subset of $\text{SProb} D$.

Proof. The same result for $\text{Prob} X$ is Proposition 4.2 of [14]. That proof relies on Proposition 4.1 of the same paper, which characterizes properties of the embedding of $X$ into $D$, and hence applies equally to sub-probability measures. With this result in hand, the proofs in [14] apply almost verbatim for sub-probability measures, except that part (ii) uses $\mu(X) = 1$, but this can be replaced by $\mu(X) = ||\mu||$. The final point that $e(\text{SProb} X)$ is a Borel set follows from Lemma 2.3 of [33]. $\square$

Theorem 4.5 Let $X$ be a Polish space and $\iota : X \rightarrow D$ a topological embedding of $X$ into a countably based domain $D$. Then the mapping $e : \text{SProb} X \rightarrow \text{SProb} D$ by $e(\mu) = \iota_* \mu$ is a topological embedding relative to the weak topologies on $\text{SProb} X$ and on $\text{SProb} D$. 

Proof. We first show $e: \text{SProb } X \to \text{SProb } D$ is continuous: Let $\mu_n, \mu \in \text{SProb } X$ with $\mu_n \to_w \mu$. To show $e(\mu_n) \to_w e(\mu)$, let $O \subseteq D$ be Lawson open. Then:

$$\liminf_n e(\mu_n)(O) = \liminf_n \iota \mu_n(O)$$

$$= \liminf_n \mu_n(\iota^{-1}(O))$$

$$\geq \mu(\iota^{-1}(O)) \quad \text{by Theorem 4.3(ii)}$$

$$= e(\mu)(O).$$

Then Theorem 3.3(ii) implies $e(\mu_n) \to_w e(\mu)$.

For the converse, let $\nu_n \to_w \nu$ in $\text{SProb } D$ with $\nu_n, \nu$ all concentrated on $\text{ran } e$. Since $\iota: X \to D$ is a topological embedding, given an open set $O \subseteq X$, there is some Lawson open $O' \subseteq D$ satisfying $\iota(O) = O' \cap \iota(X)$. If $j: \text{SProb } D \to \text{SProb } X$ is the restriction map, then

$$\liminf_n j(\nu_n)(O) = \liminf_n \nu_n(\iota(O))$$

$$= \liminf_n \nu_n(O' \cap \iota(X))$$

$$= \liminf_n \nu_n(O') \quad \nu_n \text{ is concentrated on } \iota(X)$$

$$\geq \nu(O') \quad \text{by Theorem 3.3(ii)}$$

$$= j(\nu)(O).$$

It follows by Theorem 4.3(ii) that $j(\nu_n) \to_w j(\nu)$ in $\text{SProb } X$.}

Remark 4.6 Theorem 4.5 is Corollary 4.1 of [14] extended to the case of sub-probability measures, from the case of probability measures. The proof is identical to the one in [14], except the reasoning has been changed to rely on the results we established for $\text{SProb}$.

4.1 Bounded complete domains and Skorohod’s Theorem

The connection between domains and Polish spaces involves computational models. A computational model for a topological space $X$ is a domain $D$ for which there is a topological embedding $X \simeq \text{Max } D$ of $X$ as the space of maximal elements of $D$ endowed with the relative Scott topology. As described in the Introduction, this notion emerged from the work of Edalat, who developed the first domain models of spaces arising in real analysis using the domain of compact subsets of the space under reverse inclusion. Later, Lawson [23] showed that the space $\text{Max } D$ of maximal elements of a bounded complete countably based domain in the relative Scott topology is a Polish space, and Ciesielski, Flagg and Kopperman [9, 10] showed that every Polish space has such a model. Finally, Martin [24] noted that the space of maximal elements of any countably based, bounded complete domain is a $G_\delta$ in the relative Scott topology. Since these results play a crucial role in our work, we state them formally:

Theorem 4.7 (Lawson [23], Ciesielski, et al. [9, 10], Martin [24]) A space $X$ is representable as $\text{Max } D$ in the relative Scott topology, for $D$ a countably based, bounded
complete domain, iff $X$ is a Polish space. In such a representation, $X$ is a $G_δ$-subspace of $D$ in the Scott topology.

Our goal is to prove Skorohod’s Theorem using our results from Section 3. But before we do that, we need one more preparatory result. W. M. Schmidt was the first to observe that the canonical surjection $π: C \rightarrow [0, 1]$ from the Cantor set, $C \simeq 2^\mathbb{N}$ sends Haar measure $ν_C$ to Lebesgue measure. We need the lower adjoint of that projection for our proof.

**Proposition 4.8** Let $C \simeq 2^\mathbb{N}$ denote the Cantor set regarded as the countable product of two-point groups. Then the canonical projection $π: C \rightarrow [0, 1]$ has a lower adjoint $j: [0, 1] \rightarrow C$ satisfying $j([0, 1]) = C \setminus KC$, and $j \lambda = ν_C$, where $λ$ denotes Lebesgue measure.

**Proof.** For a proof that $π ν_C = λ$, see [8], which also has an extensive discussion of related results, including the following. First, $π$ preserves all sups and all infs, so, in particular, it has a lower adjoint $j: [0, 1] \rightarrow C$ preserving all suprema. Then $j([0, 1]) = C \setminus KC$, where $KC$ is the set of compact elements, which is countable, so $ν_C(KC) = 0$. The pair of maps, $π|_{C \setminus KC}$ and $j$, form a Borel isomorphism.

We claim $j λ = ν_C$: Indeed, if $A \subseteq C$ is a Borel set, then

$$j λ(A) = λ(j^{-1}(A)) = λ(j^{-1}(A \setminus KC)) = ν_C(p^{-1} ◦ j^{-1}(A \setminus KC)) = ν_C(A \setminus KC) = ν_C(A),$$

since $ν_C(KC) = 0$. \(\square\)

A stochastic process on a measure space $(S, Σ_S)$ is a family $\{X_t \mid t ∈ T \subseteq \mathbb{R}^+\}$ of random variables $X_t: (P, Σ_P, ν) \rightarrow (S, Σ_S)$, where $(P, Σ_P, ν)$ is a probability space. It’s often assumed that $(S, Σ_S)$ is a Polish space, in which case $Prob S$ also is Polish in the weak topology. The push forward measure $X_t * ν \in Prob S$ is called the law of $X_t$, and a natural question is the convergence properties of the family $\{X_{t*} ν \mid t \in T\} \subseteq Prob S$. Since $Prob S$ is Polish, convergence can be defined using sequences, and it’s obvious that if $X_{t_n} \rightarrow X_t$ a.e. on $X$, then $X_{t_n*} ν \rightarrow_w X_{t*} ν$ in $Prob S$. Skorohod’s Theorem not only provides a converse to this observation, it also shows the probability space $(P, Σ_P, ν)$ can be assumed to be the unit interval with Lebesgue measure, $λ$.

**Theorem 4.9** (Skorohod’s Theorem [32]) If $P$ is a Polish space and $μ ∈ Prob P$, then there is a random variable $X: [0, 1] \rightarrow P$ satisfying $X_* λ = μ$.

Moreover, if $μ_n, μ ∈ Prob P$ satisfy $μ_n \rightarrow_w μ$ in the weak topology, then the random variables $X, X_n: [0, 1] \rightarrow P$ can be chosen so that $X_* λ = μ$ and $X_{n*} λ = μ_n$ also satisfy $X_n \rightarrow X$ a.s. wrt Lebesgue measure.

**Proof.** Since $P$ is Polish, Theorem 4.7 implies there is a bounded complete domain $D_P$ with a countable base and an embedding $e: P \rightarrow Max D$ and $e(P)$ is a $G_δ$ subset of $D_P$. Then $Prob e = e_*: Prob P \rightarrow Prob D_P$ is an embedding.
Theorem 3.5 then implies there are Scott-continuous maps \( X', X'_n : \mathbb{C} \to D_P \) satisfying \( X'_n \nu_C = \mu, X'_{n*} \nu_C = \mu_n \), and \( X'_n \to X' \) is the Scott topologies. But since \( \mu, \mu_n \in \text{Max } D_P \), it follows that \( X_n \to X \) in the Lawson topologies.

If we restrict the mappings \( X', X'_n \) to those \( x \in \text{Max } \mathbb{C} \) with \( X'_n(x) \in \text{Max } D \), and then precompose with \( j \), we then have random variables \( X_n = X'_n \circ j, X = j \circ X' : [0, 1] \to P \) as desired: \( X_n \lambda = X'_{n*} \nu_C = \mu_n, X \lambda = X'_n \nu_C = \mu \), and \( X_n \to X \) a.s. on \([0, 1]\) wrt \( \lambda \).

\[ \square \]

Remark 4.10 A standard Borel space is a measurable space for which there is a Borel isomorphism onto a Borel subset of a Polish space \([22]\). An obvious example is the unit interval. A standard probability space is then a probability space \( P, \Sigma_P, \nu \) that is isomorphic mod 0 to the unit interval equipped with Lebesgue measure, where “isomorphic mod 0” means there are Borel sets \( A \subseteq P \) and \( B \subseteq [0, 1] \) both of measure 0 so that \( P \setminus A \) is Borel isomorphic to \([0, 1] \setminus B \). This leads to two comments:

(i) The thrust of Skorohod’s Theorem is that the domain for any stochastic process whose range is a Polish space can be assumed to be a standard probability space. Some statements of the theorem simply state it that way, without specifying which standard probability space is being used. But most often, the standard space is assumed to be the unit interval with Lebesgue measure.

(ii) Any two standard Borel spaces that are uncountable are Borel isomorphic (cf. \([22]\)). Clearly \( \mathcal{C} \) is such a space, as is the canonical example, \([0, 1]\). So we could have used \( \mathcal{C} \) in Theorem 4.9 instead of \([0, 1]\).

Finally, just as with Theorem 3.5, on which Theorem 4.9 relies, this result extends verbatim to subprobability measures. T

Corollary 4.11 If \( P \) is a Polish space and \( \mu \) is a subprobability measure on \( P \), then there is a random variable \( X : (0, 1] \to P \) satisfying \( X \lambda = \mu \).

Moreover, if \( \mu_n, \mu \in \text{SProb } P \) satisfy \( \mu_n \to \mu \) in the weak topology, then the random variables \( X, X_n : (0, 1] \to P \) can be chosen so that \( X \lambda = \mu \) and \( X_n \lambda = \mu_n \) also satisfy \( X_n \to X \) a.s. wrt Lebesgue measure.

Proof. Apply Corollary 3.8, and precede the mappings defined there with the lower semicontinuous embedding \( j : (0, 1] \to \mathcal{C} \). \[ \square \]

5 Summary and Future Work

In this paper we have developed a domain-theoretic approach to random variables. Our main results establish some standard results in probability theory using the Cantor tree as a domain, in which measurable maps from the Cantor set are approximated by Scott-continuous maps defined on the tree. We have recast Skorohod’s Theorem in domain-theoretic terms, and we’ve also extended the theorem to subprobability measures using our approach, both for domains and for the classic case of Polish spaces. We also presented a direct proof that the sub-probability measures and the probability measures on a complete chain form a continuous lattice. This
result offers the first new insight in over two decades to the domain structure of \( \text{SProb} D \) and \( \text{Prob} D \), the last such results having appeared in [21].

There are a number of interesting questions to be explored. Finding further results from random variables that can be obtained using the techniques presented here is an obvious issue. Another question we have been exploring is the potential use of the disintegration theory for product measures, in order to understand the domain structure of \( \text{SProb} (D \times E) \), in the case \( D \) and \( E \) are chains. In particular, we’d like to know if we can use the fact that \( \text{SProb} D \) and \( \text{SProb} E \) are continuous lattices to derive some insight into the domain structure of \( \text{SProb} (D \times E) \). Our first idea – that this family of measures also would be a lattice – is debunked by a simple example in [20], so more subtle issues are at play here. Last, we’re interested in investigating the potential application of our ideas to probabilistic programming semantics and Bayesian inference.

Acknowledgement

This work began while the author was a participant in the Logical Methods in Computation Program at the Simons Institute for the Theory of Computing in Fall, 2016. The author thanks the institute and its staff for providing an excellent environment for research. The author also wishes to acknowledge the support of the US AFOSR during the preparation of this work.

References

[1] Abramsky, S. and A. Jung, Domain Theory, in: Handbook of Logic in Computer Science, Clarendon Press (1994), pp. 1–168.

[2] Alvarez-Manilla, M., A. Edalat and N. Saheb-Djarhomi, An extension result for continuous valuations, ENTCS 13 (1998), pp. 2–11.

[3] Barker, T., A monad for randomized algorithms, Tulane PhD dissertation, May, 2016.

[4] Bilokon, P. and A. Edalat, A domain-theoretic approach to Brownian motion and general continuous stochastic processes, Proceedings of CSL–LICS 2014, IEEE Press, pp. 15:1 – 15:10.

[5] Billingsley, P. Convergence of Probability Measures, John Wily & Sons, 1968.

[6] Bollabas, B., Graph Theory, An Introductory Course, Springer-Verlag, 1979.

[7] van Breugal, F., M. Mislove, J. Ouaknine, J. Worrell, Domain theory, testing and simulations for labelled Markov processes, Theoretical Computer Science 333 (2005), pp 171–197.

[8] Brian, W. and M. Mislove, From Haar to Lebesgue via Domain Theory, Lecture Notes in Computer Science 8464 (2014), pp. 214-228.

[9] Ciesielski, K., R. C. Flagg and R. Kopperman, Characterizing topologies with bounded complete computational models, in: Proceedings of MFPS XV, ENTCS 20 (1999).

[10] Ciesielski, K., R. C. Flagg and R. Kopperman, Polish spaces, computable approximation and bitopological spaces, Topology and Its Applications 119 (2002), pp. 241–256.

[11] Edalat, A., Domain theory and integration, Proceedings of LICS 1994, Paris, France, IEEE Press, pp. 115-124.

[12] Edalat, A., Domain theory in stochastic processes, Proceedings of LICS 1995, Paris, France, IEEE Press, pp. 244–254.
[13] Edalat, A., Dynamical systems, measures and fractals in domain theory, Information and Computation 120 (1995), pp. 32–48.

[14] Edalat, A., When Scott is weak on the top, Mathematical Structures for Computer Science 7 (1997), pp. 401–417.

[15] Edalat, A. and R. Heckmann, A computational model for metric spaces, Theoretical Computer Science 193 (1998), pp. 53–73.

[16] Gierz, G., K. H. Hofmann, J. D. Lawson, M. Mislove and D. S. Scott, Continuous Lattices and Domains, Cambridge University Press, 2003.

[17] Goubault-Larrecq, J. and D. Varacca, Continuous Random Variables. LICS 2011, IEEE Press, pp. 97–106.

[18] Hall’s Marriage Theorem, online at https://en.wikipedia.org/wiki/Hall’s_marriage_theorem

[19] Heunen, C, O. Kammar, S. Staton and H. Yang, A convenient category for higher-order probability theory, in: 32nd Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2017, pp. 1–12.

[20] Jones, C., Probabilistic nondeterminism, PhD Thesis, University of Edinburgh, (1988).

[21] Jung, A. and R. Tix, The troublesome probabilistic powerdomain, ENTCS 13 (1998) pp. 70–91.

[22] Kechris, A., Classical Descriptive Set Theory. Springer-Verlag, 1994.

[23] Lawson, J. D., Spaces of maximal points, Mathematical Structures in Computer Science 7 (1997), pp. 543–555.

[24] Martin, K., The space of maximal elements in a compact domain, ENTCS 40 (2001).

[25] Mislove, M., Discrete random variables over domains, Theoretical Computer Science 380, July 2007, pp. 181-198.

[26] Mislove, M., Anatomy of a Domain of Continuous Random Variables I, Theoretical Computer Science 546 (2014), pp. 176–187.

[27] Mislove, M., Anatomy of a Domain of Continuous Random Variables II, in: Bob Coecke, Luke Ong, Prakash Panangaden, editors, Computation, Logic, Games, and Quantum Foundations. The Many Facets of Samson Abramsky. Lecture Notes in Computer Science 7860 (2013), pp. 225–245.

[28] Munkres, Topology, 2nd Edition, Prentice Hall Publishers, 2000, 535 pages.

[29] Plotkin, G. D. and J. Power: Notions of Computation Determine Monads. FoSSaCS 2002, pp 342-356.

[30] Scibior, A., O. Kammar, M. Vákár, S. Staton, H. Yang and Y. Cai, K. Ostermann, S. K. Moss, C. Heunen and Z. Ghahramani, Denotational validation of higher-order Bayesian inference, Proceedings of the ACM on Programming Languages 2 Issue POPL (2018), pp. 1–60.

[31] Scott, D., The stochastic lambda calculus, An extended abstract. J. Applied Logic 12(3): 369-376 (2014).

[32] Skorohod, A. V. , Limit theorems for stochastic processes. Theory of Probability & Its Applications, 1 (1956), pp. 261290.

[33] Varadarajan, V. R., Groups of automorphisms of Borel spaces, Transactions of the American Mathematical Society 109 (1963), pp. 191–220.