The Pennsylvania State University

The Graduate School

Eberly College of Science

EFFECTIVE THEORY OF LEVY AND FELLER PROCESSES

A Dissertation in

Mathematics

by

Adrian Maler

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December 2015

The dissertation of Adrian Maler was reviewed and approved* by the following:

Stephen G. Simpson
Professor of Mathematics
Dissertation Adviser
Chair of Committee

Jan Reimann
Assistant Professor of Mathematics

Manfred Denker
Visiting Professor of Mathematics

Bharath Sriperumbudur
Assistant Professor of Statistics

Yuxi Zheng
Francis R. Pentz and Helen M. Pentz Professor of Science
Head of the Department of Mathematics

*Signatures are on file in the Graduate School.

ABSTRACT

We develop a computational framework for the study of continuous-time stochastic processes with càdlàg sample paths, then effectivize important results from the classical theory of Lévy and Feller processes.

Probability theory (including stochastic processes) is based on measure. In Chapter 2, we review computable measure theory, and, as an application to probability, effectivize the Skorokhod representation theorem.

Càdlàg (right-continuous, left-limited) functions, representing possible sample paths of a stochastic process, form a metric space called Skorokhod space. In Chapter 3, we show that Skorokhod space is a computable metric space, and establish fundamental computable properties of this space.

In Chapter 4, we develop an effective theory of Lévy processes. Lévy processes are known to have càdlàg modifications, and we show that such a modification is computable from a suitable representation of the process. We also show that the Lévy-Itô decomposition is computable.

In Chapter 5, we extend the effective theory from Lévy processes to the larger class of Feller processes. Feller processes, too, are known to admit càdlàg modifications, and we show that such a modification is computable from a suitable representation of that type of process, which is quite different from how we represent a Lévy process.

In Chapter 6, we outline two areas for future research: an effective theory of càdlàg martingales, and algorithmic randomness for Lévy processes.

Contents

Acknowledgments

This dissertation would not have been possible without the direction of my adviser, Steve Simpson; a great deal of advice from Jason Rute; and a great deal of homework from Manfred Denker. I would also like to thank Jan Reimann, Philip Scott, Donald Richards and Bharath Sriperumbudur, and the math department staff at Penn State, particularly Becky Halpenny.

Chapter 1

Introduction

The purpose of this dissertation is

(i) to develop a computational framework for the study of continuous-time stochastic processes whose sample paths, up to a modification, are càdlàg, which includes Lévy and Feller processes; and

(ii) within this framework, to effectivize important results from classical stochastic process theory, proving

  (a) that a càdlàg modification of a Lévy process not only exists, but is computable from a suitable representation of the process,

  (b) that the components of the Lévy-Itô decomposition are similarly computable from a representation, and

  (c) that a càdlàg modification of a Feller process, too, is computable from a suitable representation of that type of process.

These problems belong to the field of computable analysis, which is the intersection of computability theory and mathematical analysis: the study of how, and to what extent, classical analysis (including measure theory) can be effectivized, meaning carried out by an idealized computer.[1] More precisely, modern probability theory (including stochastic processes) is based on measure theory, so naturally our effective theory of stochastic processes will be based on computable measure theory.

To further explain and motivate our project, this chapter includes a short history of relevant research by other authors in Section 1.1, and an outline of the rest of the dissertation (plus a word on notation) in Section 1.2.

[1] For a brief history of computable analysis, see Avigad and Brattka [5].

1.1 Background

This section is not intended to be a comprehensive review of the literature, but it should at least contain the essential references for our purposes.

Recently, computable aspects of Brownian motion have been studied in the context of algorithmic randomness by Fouché [32, 33], Kjøs-Hanssen and Nerode [49], and Allen, Bienvenu, and Slaman [2], among others,[2] following initial work by Asarin and Pokrovskii [4]. Our own project was conceived as an extension of that theory to the entire class of Lévy processes. (Essentially, Brownian motion is the only continuous Lévy process.) That said, we defer to future research all applications to algorithmic randomness.

Our extension of the effective theory to Lévy processes, then to the even larger class of Feller processes, is based on computable measure theory, which is a fairly new theory itself. (It was used to study Brownian motion as well.) The recent work by Gács [39], Hoyrup and Rojas [47], and Rute [67] was essential to our project. We should also mention similar and related work by Weihrauch [72], Schröder [68], and Bosserhoff [13].[3]

Quite separately, Chan [18, 19, 20] has developed a constructive theory of continuous-time stochastic processes, including Feller processes, based on a constructive theory of measure which is due to Bishop and Cheng [10] and extends initial work on constructive analysis by Bishop [9]. Our proof that a càdlàg modification of a Feller process is computable is in part an adaption of a constructive argument in Chan [19].

Finally, we should mention other effective theories of stochastic processes of some particular class: Rute [67], discrete-time (sub-, super-, and reverse) martingales; Kjøs-Hanssen, Nguyen, and Rute [51], martingales on Brownian motion; Mukeru [57], stochastic integrals with respect to Brownian motion; Collins [22], discrete-time Markov processes and stochastic integrals with respect to Brownian motion; Edalat [30], discrete-time processes; Bilokon and Edalat [8], continuous-time continuous processes; Freer and Roy [38], discrete-time exchangeable processes; and there may be others.

[2] We might add Fouché [34, 35, 36], Kjøs-Hanssen and Nerode [50], Kjøs-Hanssen and Szabados [52], Davie and Fouché [26], Fouché, Mukeru, and Davie [37], and Allen [1].

[3] Other recent work on or related to computable measure theory, in no particular order: Schröder and Simpson [69], Hoyrup and Rojas [45, 46], Gács, Hoyrup, and Rojas [40, 41], Bienvenu et al. [6], Miyabe [55], Müller [58], Davie [25], Wu [75, 76], Wu and Ding [77, 78], Wu and Weihrauch [79], Hertling and Weihrauch [43, 44], Weihrauch and Tavana-Roshandel [74], and Edalat [29, 31]. Again, this is not intended to be comprehensive.

1.2 Outline

For convenience, we present a section-by-section outline of Chapters 2–6. This information is also included at the start of the chapters themselves.

Chapter 2, "Computable measure theory," is largely review, starting with some concepts from general computability theory in Section 2.1. This is background for computable analysis, which is covered in some detail in Section 2.2. We specialize to computable measure theory in Section 2.3, where we also prove an effective version of the Skorokhod representation theorem as an example of an application to probability theory.

Our interest, again, is mainly in applications to stochastic process theory. Now, in the literature, a stochastic process X is typically presented as a time-indexed family $\{X_t\}$ of random variables $\omega \mapsto X_t(\omega)$, but it can also be treated as a single random variable $\omega \mapsto (t \mapsto X_t(\omega))$ taking values, called sample paths, in some space of functions. In computable analysis, a sample path presentation is more useful, assuming the sample paths are reasonably well behaved, so that they belong to a sensible function space—which is the case for many important classes of stochastic processes.

In Chapter 3, "Effective Skorokhod space," we introduce a computable metric space whose points are càdlàg functions, representing possible sample paths of a stochastic process. Naturally, we begin by reviewing the classical theory in Section 3.1. In Section 3.2, we show that those functions form a computable metric space, and in Section 3.3, we establish the computable properties of càdlàg functions that will be needed in subsequent chapters.

That brings us to our first class of stochastic processes. A Lévy process $\{X_t : t \geq 0\}$ is an $\mathbb{R}^n$-valued stochastic process with independent, stationary increments, so that (i) successive changes in position, say $X_{t_1} - X_{t_0}$, $X_{t_2} - X_{t_1}$, and so on, are statistically independent, and (ii) the probability distribution of the increment $X_t - X_s$ depends only on the elapsed time $t - s$.

According to an important classical theorem, every Lévy process admits a modification whose sample paths are càdlàg, so they live in Skorokhod space. Moreover, the Lévy-Itô decomposition states that every Lévy process can be written as the sum of four processes: a deterministic drift, a Brownian motion, a compound Poisson process, and a pure jump martingale.

We effectivize that much of the classical theory in Chapter 4, "Effective theory of Lévy processes." Section 4.1 is review, of course, including the two theorems just mentioned. In Section 4.2, we show that a càdlàg modification of a Lévy process not only exists, but is computable from a representation of the process. In Section 4.3, building on that result, we show that the components of the Lévy-Itô decomposition are computable as well.

Every Lévy process is also a Feller process. A Feller process is a type of Markov process, which is to say a process with no memory: the distribution of its future state depends, not on its entire past, but only on its present state. Since every Feller process has a càdlàg modification, it seems natural to extend the effective theory in this direction. One difference is that, in many cases, we would rather treat a Feller process, not as a given collection of sample paths, but as a so-called transition function describing how the process, whatever it may be, changes from one time to another. We would then construct sample paths that are consistent with the transition function.

For that reason alone, the material in Chapter 5, "Effective theory of Feller processes," differs considerably from the previous chapter. As usual, Section 5.1 is review. In Section 5.2, we effectivize all the basic objects from the classical theory, including transition functions. Finally, in Section 5.3, we show that a càdlàg modification of a Feller process, too, is computable from a representation of the process—which, being based on a transition function, is quite different from our representation of a Lévy process.

With our computational framework in place, and with some idea of how to effectivize (some parts of) stochastic process theory, there are a number of potential extensions and applications. In Chapter 6, "Future directions," we mention two: an effective theory of càdlàg martingales in Section 6.1, and algorithmic randomness for Lévy processes in Section 6.2.

One last thing before we begin: a word on notation. Obviously, different areas of mathematics use their own, which sometimes clash—especially in this dissertation. In probability theory, X and Y denote stochastic processes; in computable analysis, they denote metric spaces. In both cases, μ and λ most likely denote measures; in the context of càdlàg functions, they denote so-called time changes—and so on. We will try to avoid overloading.

We might add that in the literature there appears to be no accepted standard term for an effectively compact set (Definition 2.2.52), a càdlàg function (Definition 3.1.3), the metric on the space of measurable functions (Definition 2.3.21), either of the two metrics on the space of càdlàg functions (Definitions 3.1.22 and 3.1.35), a probability kernel (Definition 5.1.3), or a Feller process (Definition 5.1.44)—but these are minor points.

Chapter 2

Computable measure theory

Modern probability theory, including stochastic processes, is developed in a framework of measure theory. Accordingly, we develop our effective theory of stochastic processes in a framework of computable measure theory.

Section 2.1 briefly reviews some concepts from general computability theory as background for computable analysis, which is covered in Section 2.2. We specialize to computable measure theory in Section 2.3, where we also prove an effective version of the Skorokhod representation theorem as an application to probability theory; otherwise, this chapter is largely review.

2.1 Computability

For a general introduction to computability, see, for example, Cutland [24] or Rogers Jr. [64]. The fundamental concepts are

(i) a *computable function* $f :\subseteq \mathbb{N}^n \to \mathbb{N}^m$ (for any given $n, m \in \mathbb{N}$), and

(ii) a *computable operator* $\Phi :\subseteq \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}$.

We assume the reader is familiar with both. In this section, we review the concept of an effective representation, as background for computable analysis. Representations were introduced by Kreitz and Weihrauch [53]; for details, see Weihrauch [73]. The material is somewhat technical.

As a preliminary, we establish a notational convention:

Notation 2.1.1. Angle brackets denote finite sequences, also known as strings. Round brackets denote infinite sequences (and open intervals).

2.1.1 Representations

Taking as fixed the definition of a computable function $f :\subseteq \mathbb{N}^n \to \mathbb{N}^m$, we make the standard extensions in Definitions 2.1.2–2.1.5.

Definition 2.1.2. A set $A \subseteq \mathbb{N}^n$ is said to be *computable* if its indicator function $\mathbf{1}_A : \mathbb{N}^n \to \{0,1\} \subseteq \mathbb{N}$ is (total) computable, where as usual

$$\mathbf{1}_A(a) = \begin{cases} 1 & \text{if } a \in A, \\ 0 & \text{otherwise.} \end{cases} \tag{2.1.1}$$

Definition 2.1.3. A sequence $(a_i)_{i\in\mathbb{N}} \subseteq \mathbb{N}^n$ is said to be *computable* if the function $i \mapsto a_i$ is (total) computable of type $\mathbb{N} \to \mathbb{N}^n$.

Definition 2.1.4. A set $A \subseteq \mathbb{N}^n$ is said to be *computably enumerable* if $A = \{a_i : i \in \mathbb{N}\}$ is the image of some computable sequence $(a_i)_{i\in\mathbb{N}}$. In that case, $(a_i)_{i\in\mathbb{N}}$ is said to *computably enumerate* A.

Definition 2.1.5. The elements of a sequence $(f_i)_{i\in\mathbb{N}}$ of functions $\subseteq \mathbb{N}^n \to \mathbb{N}^m$ are said to be *uniformly computable* if there exists a computable function $F :\subseteq \mathbb{N}^{n+1} \to \mathbb{N}^m$ such that $F(i,\cdot) = f_i$ for all $i \in \mathbb{N}$.

In computability theory, our various idealized computers (e.g., Turing machines, register machines) can only handle, as inputs and outputs, *finitary* objects, i.e., natural numbers and anything that can be encoded as such.

Remark 2.1.6. There is a well-known canonical way to encode, as a unique natural number, any string $\langle e_0, \ldots, e_n \rangle$ of (a) natural numbers, (b) integers, or (c) rational numbers. We may identify such a sequence with its code. Thus, for the purposes of computability theory, a function of type, say, $\mathbb{N}^2 \to \mathbb{N}$ or $\mathbb{N} \times \mathbb{Z} \to \mathbb{Q}^2$, etc., may be treated as a function of type $\mathbb{N} \to \mathbb{N}$.

With our basic definitions in hand, we can represent non-finitary objects, such as real numbers, as sequences of natural numbers.

Definition 2.1.7. A *represented set* is a pair $\langle X, \delta \rangle$ where X is a set and $\delta :\subseteq \mathbb{N}^{\mathbb{N}} \to X$ is a surjective function. In that case, δ is said to be a *representation* of X.

Definition 2.1.8. Let $\langle X, \delta \rangle$ be a represented set. A point $x \in X$ is said to be δ-*computable* if there exists a computable sequence $\alpha \in \mathbb{N}^{\mathbb{N}}$ such that $\delta(\alpha) = x$. This can be made uniform: the elements of a sequence $(x_i)_{i\in\mathbb{N}} \subseteq X$ are said to be *uniformly* δ-*computable* if there exists a sequence $(\alpha_i)_{i\in\mathbb{N}}$ of uniformly computable sequences $\alpha_i \in \mathbb{N}^{\mathbb{N}}$ such that $\delta(\alpha_i) = x_i$ for all $i \in \mathbb{N}$.

Remark 2.1.9. If $\langle X, \delta_X \rangle$ and $\langle Y, \delta_Y \rangle$ are represented sets, the Cartesian product $X \times Y$ has an obvious canonical representation, namely

$$\delta_X \times \delta_Y : (\langle n_i, m_i \rangle)_{i \in \mathbb{N}} \mapsto \langle \delta_X((n_i)_{i \in \mathbb{N}}), \delta_Y((m_i)_{i \in \mathbb{N}}) \rangle. \tag{2.1.2}$$

Definition 2.1.10. Let $\langle X, \delta_X \rangle$ and $\langle Y, \delta_Y \rangle$ be represented sets. A function $\phi :\subseteq X \to Y$ is said to be $\langle \delta_X, \delta_Y \rangle$-*computable* on a set $D \subseteq X$ if there exists a computable operator $\Phi :\subseteq \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}$ such that for all $\alpha \in \delta_X^{-1}(D)$,

$$\delta_Y(\Phi(\alpha)) = \phi(\delta_X(\alpha)). \tag{2.1.3}$$

That is, Φ maps each representation of x to some representation of $\phi(x)$.

Notation 2.1.11. We may omit the representations when they are clear from the context, writing "computability" rather than "δ-computability."

Fact 2.1.12. If $\phi :\subseteq X \to Y$ is computable on D and $x \in D$ is computable, then $\phi(x)$ is computable. This can be made uniform in x: if $x_i \in D$ is uniformly computable (for all $i \in \mathbb{N}$), then $\phi(x_i)$ is uniformly computable.

Remark 2.1.13. From now on, it should be clear from the context how a given definition, result, etc., can be "made uniform."

Notation 2.1.14. When $\phi :\subseteq X \to Y$ is computable on D, we may write: "$\phi(x)$ is computable *uniformly from a representation* of x" (for all $x \in D$).

2.1.2 Names and numbered sets

Under assumptions which are quite reasonable in the context of computable analysis, a set admits a canonical representation.

Definition 2.1.15. A *numbered set* is a pair $\langle \mathcal{O}, (o_n)_{n \in \mathbb{N}} \rangle$ where $\mathcal{O}$ is a countable set and $(o_n)_{n \in \mathbb{N}}$ enumerates $\mathcal{O}$. In that case, $(o_n)_{n \in \mathbb{N}}$ is said to be a *numbering* of $\mathcal{O}$, and $n \in \mathbb{N}$ is said to be a *code* for $o = o_n \in \mathcal{O}$.

Definition 2.1.16. A numbered set $\langle \mathcal{O}, (o_n)_{n \in \mathbb{N}} \rangle$ and a surjective function $\rho :\subseteq \mathcal{O}^{\mathbb{N}} \to X$ induce a *canonical representation* of X, denoted by δ_ρ, via

$$\delta_\rho((n_i)_{i \in \mathbb{N}}) = \rho((o_{n_i})_{i \in \mathbb{N}}) \tag{2.1.4}$$

whenever the right-hand side is defined. ρ is said to be a *description* of X with respect to $\langle \mathcal{O}, (o_n)_{n \in \mathbb{N}} \rangle$, and the elements of $\mathcal{O}$ are *elementary objects*.

Remark 2.1.17. In this dissertation, all the representations will be induced canonically by a numbered set of elementary objects and a description.

Now fix a numbered set $\langle \mathcal{O}, (o_n)_{n\in\mathbb{N}} \rangle$, a description $\rho :\subseteq \mathcal{O}^{\mathbb{N}} \to X$, and the associated canonical representation $\delta_\rho :\subseteq \mathbb{N}^{\mathbb{N}} \to X$.

Definition 2.1.18. Let $x \in X$. If $\rho(\sigma) = x$ for some $\sigma \in \mathcal{O}^{\mathbb{N}}$, σ is said to be a *ρ-name* for x. If we write $\sigma = (o_{n_i})_{i\in\mathbb{N}}$, then $(n_i)_{i\in\mathbb{N}} \subseteq \mathbb{N}$ is said to *index* σ.

Notation 2.1.19. We may omit the numbered set and description when they are clear from the context, writing, e.g., "name" rather than "ρ-name."

Definition 2.1.20. A name σ is said to be *ρ-computable* if it is indexed by a computable sequence $(n_i)_{i\in\mathbb{N}}$. A point $x \in X$ is said to be *ρ-computable* if it has a ρ-computable name σ. Both definitions can be made uniform.

Fact 2.1.21. A point $x \in X$ is δ_ρ-computable (Definition 2.1.8) if and only if x is ρ-computable. This can be made uniform in x.

Notation 2.1.22. It is usually safe to identify an elementary object with a code for that object; see Remark 2.1.6. In that case, a sequence $(n_i)_{i\in\mathbb{N}}$ of codes indexing a name $(o_{n_i})_{i\in\mathbb{N}}$ would simply be identified with that name.

Notation 2.1.23. If $\phi(x)$ is computable uniformly from a representation of x (Notation 2.1.14) and the representation is canonical for a numbered set, we may write: "$\phi(x)$ is computable *uniformly from a name* for x."

2.2 Computable analysis

This section is mostly review material. Part 2.2.1 is essential, as it covers elementary computable analysis. Part 2.2.2, on effective compactness, and Part 2.2.3, on continuous function spaces, are somewhat technical.

In any case, for a more detailed introduction to computable analysis, including proofs, the reader may consult Weihrauch [73] or, for a slightly different style of presentation, Pour-El and Richards [62]. We assume the reader is familiar with classical analysis as in, say, Rudin [66].

2.2.1 Metric spaces

Before we consider computable metric spaces in general, we should explain what it means for a real number to be computable.

We define a canonical representation (see Definition 2.1.16) of $\mathbb{R}$, using the rational numbers as elementary objects and certain Cauchy sequences as names. First, let $(q_i^\circ)_{i\in\mathbb{N}}$ be a standard one-to-one enumeration of $\mathbb{Q}$.

Definition 2.2.1. A sequence $\sigma = (\sigma_i)_{i\in\mathbb{N}}$ of rational numbers is said to be a *fast Cauchy sequence* if $j > i$ implies $|\sigma_i - \sigma_j| < 2^{-i}$ for all $i, j \in \mathbb{N}$.

Fact 2.2.2. Let $\mathcal{Q} \subseteq \mathbb{Q}^{\mathbb{N}}$ be the set of fast Cauchy sequences. The function $\rho_{\lim}(\sigma) = \lim \sigma$ with domain $\mathcal{Q}$ is a description of $\mathbb{R}$ with respect to the numbered set $\langle \mathbb{Q}, (q_i^\circ)_{i\in\mathbb{N}} \rangle$, so it induces a canonical representation $\delta_{\lim}$ of $\mathbb{R}$.

We identify each rational number with its code (see Notation 2.1.22). As a special case of Definition 2.1.18, we have

Definition 2.2.3. A *Cauchy name*, or just a *name*, for a real number x is a sequence $\sigma = (\sigma_i)_{i\in\mathbb{N}}$ of rational numbers such that

(i) σ converges to x, i.e., $\lim_{i\to\infty} |\sigma_i - x| = 0$, and

(ii) σ is fast Cauchy, i.e., $j > i$ implies $|\sigma_i - \sigma_j| < 2^{-i}$ for all $i, j \in \mathbb{N}$.

And as a special case of Definition 2.1.20, we have

Definition 2.2.4. $x \in \mathbb{R}$ is said to be *computable* if it has a computable Cauchy name. The elements of a sequence $(x_i)_{i\in\mathbb{N}}$ are said to be *uniformly computable* if there exists a computable rational (double) sequence $(\sigma_{i,j})_{i,j\in\mathbb{N}}$ such that for all $i \in \mathbb{N}$, $(\sigma_{i,j})_{j\in\mathbb{N}}$ is a Cauchy name for $x_i \in \mathbb{R}$.

Fact 2.2.5. The set of computable real numbers is (i) countable, (ii) dense in $\mathbb{R}$ with the usual metric, and (iii) an algebraically closed subfield of $\mathbb{R}$.

Here it is worth digressing briefly to mention semicomputability:

Definition 2.2.6. $x \in \mathbb{R}$ is said to be *lower semicomputable* if it is the supremum of a computable sequence of rationals, and *upper semicomputable* if it is the infimum of a computable sequence of rationals.

Fact 2.2.7. A real number x is computable if and only if x is both upper and lower semicomputable.

To return to the main subject, knowing what it means to compute a real number, we can generalize to points in a complete, separable metric space:

Definition 2.2.8. A *computable metric space* is a triple $\langle X, d, \theta \rangle$ such that

(i) $\langle X, d \rangle$ is a complete, separable metric space,

(ii) $\theta = (s_n)_{n\in\mathbb{N}}$ enumerates a (countable) dense subset S of X, and

(iii) the real number $d(s_n, s_m)$ is computable uniformly from $n, m \in \mathbb{N}$.

In that case, the elements of S are called the *ideal points* for $\langle X, d, \theta \rangle$.

Generalizing Definition 2.2.1, we have

Definition 2.2.9. A sequence x in a metric space $\langle X, d \rangle$ is said to be a *fast Cauchy sequence* if $j > i$ implies $d(x_i, x_j) < 2^{-i}$ for all $i, j \in \mathbb{N}$.

Now let $\mathbb{X} = \langle X, d, \theta \rangle$ be a computable metric space.

Fact 2.2.10. Let $\mathcal{S}$ be the set of fast Cauchy sequences of ideal points. The function $\rho_S(\sigma) = \lim \sigma$ with domain $\mathcal{S}$ is a description of X with respect to the numbered set $\langle S, \theta \rangle$, so it induces a canonical representation δ_S of X.

Notation 2.2.11. We may identify the enumeration $\theta = (s_n)_{n \in \mathbb{N}}$ with the corresponding set $S = \{s_n : n \in \mathbb{N}\}$, and write $\mathbb{X}$ as $\langle X, d, S \rangle$ rather than $\langle X, d, \theta \rangle$. In practice, this should not be confusing, because the numberings we use are fairly obvious, if not actually canonical.

Notation 2.2.12. We may identify $\mathbb{X}$ with the underlying set X.

As another special case of Definition 2.1.18, we have

Definition 2.2.13. A *Cauchy name*, or just a *name*, for a point $x \in X$ is a sequence $\alpha = (s_{n_i})_{i \in \mathbb{N}}$ of ideal points such that

(i) α converges to x in $\langle X, d \rangle$, i.e., $d(s_{n_i}, x) \to 0$, and

(ii) α is fast Cauchy, i.e., $j > i$ implies $d(s_{n_i}, s_{n_j}) < 2^{-i}$ for all $i, j \in \mathbb{N}$.

And as another special case of Definition 2.1.20, we have

Definition 2.2.14. A point $x \in X$ is said to be *computable* if it has a computable Cauchy name. This definition can be made uniform.

Fact 2.2.15. The set of computable points is countable and dense in $\langle X, d \rangle$.

Examples 2.2.16. $\mathbb{R}^n$ is a computable metric space under the Euclidean metric with $\mathbb{Q}^n$ for ideal points, for each $n \in \mathbb{N}$. The computable points are exactly the n-tuples of computable real numbers.

Similarly, $[0, \infty)^n$ and $[0, 1]^n$ are computable metric spaces.

Fact 2.2.17. $d(x, y)$ is computable, uniformly from Cauchy names for x and y, for all $x, y \in X$. In particular, if x and y are computable points, then $d(x, y)$ is a computable real number.

For completeness and as an example of elementary computable analysis, we include a proof of the following useful fact:

Fact 2.2.18. Let A be a closed subset of X. Let $(a_n)_{n\in\mathbb{N}} \subseteq A$ be a sequence of uniformly computable points dense in A. Then $\mathbb{A} = \langle A, d, (a_n)_{n\in\mathbb{N}}\rangle$ is a computable metric space. Moreover, a Cauchy name for x in $\mathbb{A}$ is computable uniformly from a Cauchy name for x in $\mathbb{X}$, and vice versa; in particular, a point $x \in A$ is computable in $\mathbb{A}$ if and only if x is computable in $\mathbb{X}$.

Proof. It is easy to see from Definition 2.2.8 that $\mathbb{A}$ is a computable metric space: (i) a closed subspace of a complete, separable metric space is complete and separable; (ii) by assumption, $(a_n)_{n\in\mathbb{N}}$ enumerates a dense subset of A; and (iii) $d(a_n, a_m)$ is computable uniformly from $n, m \in \mathbb{N}$ by Fact 2.2.17.

From $\mathbb{A}$ to $\mathbb{X}$: Suppose we are given a Cauchy name $(a_{n_i})_{i\in\mathbb{N}}$ for x as a point in $\mathbb{A}$. Since a_n is computable uniformly from n as a point in $\mathbb{X}$, we can compute from $i \in \mathbb{N}$ an ideal point $s_{m_i} \in S$ such that $d(s_{m_i}, a_{n_i}) < 2^{-i-1}$. Then $(s_{m_{i+1}})_{i\in\mathbb{N}}$ is a Cauchy name for x as a point in $\mathbb{X}$, and it is in fact computable uniformly from the given name for x.

From $\mathbb{X}$ to $\mathbb{A}$: Suppose $x \in A$ and we are given a Cauchy name for x as a point in $\mathbb{X}$. Given $i \in \mathbb{N}$, search for an ideal point a_{n_i} for $\mathbb{A}$ such that $d(x, a_{n_i}) < 2^{-i-1}$. The search is computable by Fact 2.2.17, and it has to terminate because $(a_n)_{n\in\mathbb{N}}$ is dense in A. Then $(a_{n_{i+1}})_{i\in\mathbb{N}}$ is a Cauchy name for x as a point in $\mathbb{A}$, computable uniformly from the given name for x. □

We can treat the topology of X as a represented set, using certain open balls as elementary objects. $\tau(X)$ denotes the family of open subsets of X.

Definition 2.2.19. The *basic open balls* in X are $\varnothing$ and sets of the form

$$B(a, r) = \{x \in X : d(a, x) < r\} \tag{2.2.1}$$

where $a \in S$ and $r \in \mathbb{Q}_{>0} = \mathbb{Q} \cap (0, \infty)$. r is the *radius* and a is the *center*.

Fact 2.2.20. S and $\mathbb{Q}_{>0}$ are numbered sets, so the family $\mathcal{T}$ of basic open balls has a canonical numbering. The union operator $\rho_\tau(\sigma) = \bigcup\sigma$ with domain $\mathcal{T}^{\mathbb{N}}$ is a description of $\tau(X)$ with respect to $\mathcal{T}$, so it induces a canonical representation δ_τ of the topology.

Definition 2.2.21. With respect to the representation δ_τ, a *name* for an open set $U \neq \varnothing$ is essentially a sequence $(a_i, r_i)_{i\in\mathbb{N}} \subseteq S \times \mathbb{Q}_{>0}$ such that

$$U = \bigcup_{i\in\mathbb{N}} B(a_i, r_i). \tag{2.2.2}$$

Definition 2.2.22. When an open set U is computable in $\tau(X)$, we say that U is *effectively open* or Σ^0_1. The complement of an effectively open set is said to be *effectively closed* or Π^0_1. Sets that are both Σ^0_1 and Π^0_1 are said to be Δ^0_1. These definitions can all be made uniform.

Definition 2.2.23. The *basic closed balls* in X are $\varnothing$ and sets of the form

$$\overline{B}(a,r) = \{x \in X : d(a,x) \leq r\} \tag{2.2.3}$$

where $a \in S$ and $r \in \mathbb{Q}_{>0}$.

Fact 2.2.24. Finite unions and intersections of basic open balls are Σ^0_1 sets, and finite unions and intersections of basic closed balls are Π^0_1 sets.

Definition 2.2.25. An intersection $V = \bigcap_{i\in\mathbb{N}} U_i$ of uniformly Σ^0_1 sets U_i is said to be *effectively* G_δ or Π^0_2. That is, a Π^0_2 set has the form

$$V = \bigcap_{i\in\mathbb{N}} \bigcup_{j\in\mathbb{N}} B(a_{i,j}, r_{i,j}) \tag{2.2.4}$$

where $(a_{i,j}, r_{i,j})_{i,j\in\mathbb{N}} \subseteq S \times \mathbb{Q}_{>0}$ is a computable sequence.

The family of basic open balls serves as an effective basis for X in the sense of the following definition. In computable measure theory, we may need to switch to a better-behaved basis, necessitating the general definition:

Definition 2.2.26. Let $(x_i)_{i\in\mathbb{N}}$ and $(\epsilon_i)_{i\in\mathbb{N}}$ be uniformly computable in X and $(0,\infty)$, respectively. The family $\{B(a_i,\epsilon_i) : i \in \mathbb{N}\}$ of open balls is said to form an *effective basis* for X if for all open $U \neq \varnothing$ there exists a sequence $(i_k)_{k\in\mathbb{N}}$ computable uniformly from a name for U in $\tau(X)$ such that

$$U = \bigcup_{k\in\mathbb{N}} B(x_{i_k}, \epsilon_{i_k}). \tag{2.2.5}$$

Now let $\mathbb{X} = \langle X, d_X, S_X\rangle$ and $\mathbb{Y} = \langle Y, d_Y, S_Y\rangle$ be two computable metric spaces. Definition 2.1.10 becomes

Definition 2.2.27. A function $f :\subseteq X \to Y$ is said to be *computable* on a set $D \subseteq X$ if there exists a computable operator $\Phi :\subseteq S_X^{\mathbb{N}} \to S_Y^{\mathbb{N}}$ such that $\Phi(\sigma)$ is a Cauchy name of $f(x)$ for any Cauchy name σ of any point $x \in D$.

Fact 2.2.28. Every computable function is continuous on its domain.

Fact 2.2.29. $f :\subseteq X \to Y$ is computable if and only if for all open $U \subseteq Y$,

$$f^{-1}(U) = V \cap \mathrm{dom}(f) \tag{2.2.6}$$

for some open set $V \subseteq X$ computable uniformly from a name for U (where $\mathrm{dom}(f)$ denotes the domain of f). This can be made uniform.

Definition 2.2.30. A function $f : X \to [0, \infty] = [0, \infty) \cup \{\infty\}$ is said to be *lower semicomputable* if f is the pointwise supremum of some sequence $(f_i)_{i \in \mathbb{N}}$ of uniformly computable functions $f_i : X \to [0, \infty)$.

We handle functions of more than one variable using product spaces:

Definition 2.2.31. The *product space*

$$\mathbb{X} \times \mathbb{Y} = \langle X \times Y, d_{X \times Y}, S_X \times S_Y \rangle \tag{2.2.7}$$

is a computable metric space, where $X \times Y$ and $S_X \times S_Y$ are the usual Cartesian products (the latter enumerated in the canonical way) and

$$d_{X \times Y}(\langle x_1, y_1 \rangle, \langle x_2, y_2 \rangle) = \max\{d_X(x_1, x_2), d_Y(y_1, y_2)\}. \tag{2.2.8}$$

Fact 2.2.32. Addition, subtraction, and multiplication are computable of type $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$. Division $\langle x, y \rangle \mapsto x/y$ is computable on $\mathbb{R} \times (\mathbb{R} \setminus \{0\})$.

For many more elementary results in computable analysis, the reader may consult Pour-El and Richards [62] or Weihrauch [73].

Having defined computable functions between computable metric spaces, we can explain what it means for a vector space to be computable:

Definition 2.2.33. Suppose X is a real vector space with operations $+$ (vector addition) and $\cdot$ (scalar multiplication). The triple $\langle \mathbb{X}, +, \cdot \rangle$ is said to be a *computable vector space* if vector addition is computable of type $X \times X \to X$ and scalar multiplication is computable of type $\mathbb{R} \times X \to X$.

Definition 2.2.34. Suppose $\langle \mathbb{X}, +, \cdot \rangle$ is a normed space. Denote the norm by $\| \cdot \|$. The 4-tuple $\langle \mathbb{X}, \| \cdot \|, +, \cdot \rangle$ is said to be a *computable Banach space* if

(i) $\langle \mathbb{X}, +, \cdot \rangle$ is a computable vector space,

(ii) the norm $\| \cdot \|$ is computable of type $X \to [0, \infty)$, and of course

(iii) the norm actually induces the metric d_X on X.

Notation 2.2.35. To simplify our notation, when $\mathbb{X} = \langle X, d_X, S_X \rangle$, we may denote the computable Banach space $\langle \mathbb{X}, \| \cdot \|, +, \cdot \rangle$ by $\langle X, \| \cdot \|, S_X \rangle$, as d_X is implicit in $\| \cdot \|$ and the operations $+$ and $\cdot$ are usually clear from the context.

We can also explain what it means for two computable metric spaces to be isomorphic in a computable way:

Definition 2.2.36. An *isomorphism* between $\mathbb{X}$ and $\mathbb{Y}$ is a pair $\langle f, g\rangle$ where $f : X \to Y$ and $g : Y \to X$ are continuous bijections and $f^{-1} = g$. An *isometry* is an isomorphism such that $d_Y(f(x), f(y)) = d_X(x, y)$ for all $x, y \in X$.

Definition 2.2.37. An isomorphism (respectively, an isometry) $\langle f, g\rangle$ is said to be *computable* if f and g are computable functions. $\mathbb{X}$ and $\mathbb{Y}$ are said to be *computably isomorphic* (*computably isometric*) if there exists a computable isomorphism (a computable isometry) between them.

Definition 2.2.38. Let d_1 be another metric on $\mathbb{X} = \langle X, d, S\rangle$. d_1 is said to be *computably related* to d in $\mathbb{X}$ if

(i) $d_1(x, y)$ is computable uniformly from names for x and y, and

(ii) there exists a computable sequence $(i_n)_{n\in\mathbb{N}}$ such that for all $x, y \in X$, for all $n \in \mathbb{N}$, $d_1(x, y) < 2^{-i_n}$ implies $d(x, y) < 2^{-n}$.

Two computably related metrics are essentially interchangeable from a computable analysis standpoint [14]:

Fact 2.2.39. If d_1 is computably related to d in $\mathbb{X}$, then $\langle X, d_1, S\rangle$ is a computable metric space computably isomorphic to $\langle X, d, S\rangle$.

Fact 2.2.40. Let $(a_n)_{n\in\mathbb{N}}$ be a dense sequence of uniformly computable points. Then $\langle X, d, (a_n)_{n\in\mathbb{N}}\rangle$ and $\langle X, d, S\rangle$ are computably isometric.

Proof. This follows from Fact 2.2.18 and the identity isometry $\langle \mathrm{id}, \mathrm{id}\rangle$. □

Notation 2.2.41. The sets S and $S_1 = \{a_n : n \in \mathbb{N}\}$ from Fact 2.2.40 are said to be *computably equivalent* sets of ideal points for $\langle X, d\rangle$.

We list a few important computable metric and Banach spaces:

Examples 2.2.42. For all $n \in \mathbb{N}$, the *Euclidean space* $\langle \mathbb{R}^n, |\cdot|, \mathbb{Q}^n\rangle$ is a computable Banach space, where $|\cdot|$ denotes the n-dimensional Euclidean norm. We denote this space simply by $\mathbb{R}^n$. Similarly, the *unit cube* $[0, 1]^n$ and the *unit torus* $\mathbb{T}^n = (\mathbb{R}/\mathbb{Z})^n$ are computable metric spaces.

Our next example is quite important in computable measure theory.

Definition 2.2.43. Let A be any set. A^* denotes the set of strings (of any length) of elements of A. That is, $A^* = \bigcup_{n\in\mathbb{N}} A^n$. As usual, $A^{\mathbb{N}}$ denotes the family of A-valued sequences. For all $\sigma \in A^*$, we define the *cylinder set*

$$[\sigma] = \{\alpha \in A^{\mathbb{N}} : \sigma \sqsubseteq \alpha\} \tag{2.2.9}$$

where $\sigma \sqsubseteq \alpha$ means σ is a prefix of α. The topology generated by the cylinder sets is called the *cylinder topology.*

Example 2.2.44. The *Cantor space* $\{0,1\}^{\mathbb{N}}$ is a computable metric space with respect to the *first-difference metric*

$$d(\alpha,\beta) = 2^{-\min\{n\in\mathbb{N} \,:\, \alpha_n\neq\beta_n\}} \quad \text{for } \alpha \neq \beta, \tag{2.2.10}$$

which induces the cylinder topology. For ideal points, take $\{0,1\}^*$.

Remark 2.2.45. The *Hilbert cube* $[0,1]^{\mathbb{N}}$ is a computable metric space in a similar way. For ideal points, take the eventually-zero sequences of rationals.

There is a more general way to construct a computable power space:

Definition 2.2.46. Suppose $\mathbb{X} = \langle X, d, S\rangle$ is a computable metric space. Let $\widetilde{d} = \min\{1, d\}$. Then $\mathbb{X}^{\mathbb{N}}$ denotes the computable metric space whose underlying set is $X^{\mathbb{N}}$, equipped with the *product metric*

$$d_\Pi(\alpha,\beta) = \sum_{n\in\mathbb{N}} 2^{-(n+1)}\widetilde{d}(\alpha_n,\beta_n), \tag{2.2.11}$$

and for ideal points the eventually-constant S-valued sequences.

We have one more important example of a computable Banach space, which we will revisit later. Let $a < b$ be computable real numbers.

Definition 2.2.47. Let $\mathcal{P}$ enumerate the continuous functions $\phi : [a,b] \to \mathbb{R}$ for which there exist rationals $a = x_0 < x_1 < \cdots < x_n = b$ such that

(i) $\phi(x_0), \phi(x_1), \ldots, \phi(x_n)$ are rational, and

(ii) ϕ is linear on intervals of the form (x_i, x_{i+1}) for all $i < n$.

Example 2.2.48. $\mathbf{C}[a,b]$ denotes the set of continuous $\mathbb{R}$-valued functions on $[a,b]$. $\mathbf{C}[a,b]$ is a computable Banach space under the *sup-norm*

$$\|f\|_\infty = \sup\{|f(t)| : a \le t \le b\}, \tag{2.2.12}$$

which induces the *uniform metric* d_∞. For ideal points, take $\mathcal{P}$.

Of course, there are many computably equivalent sets of ideal points.

An important result in computable analysis effectivizes the Weierstrass approximation theorem [62]:

Fact 2.2.49. A function $f : [a,b] \to \mathbb{R}$ is computable if and only if f is a computable point in the space $\langle \mathbf{C}[a,b], \|\cdot\|_\infty, \mathcal{P}\rangle$. This can be made uniform.

This concludes our review of elementary computable analysis.

2.2.2 Compactness

We turn to an effective theory of compactness, which has been developed by Mori, Tsujii, and Yasugi [56], Yasugi, Mori, and Tsujii [80], Blanck [11], Brattka [14], Brattka and Weihrauch [17], and Brattka and Presser [16].[1]

It seems fair to say that the facts stated below, unless otherwise noted, belong to the folklore of effective compactness. For the sake of completeness, we provide direct proofs. For the relevant concepts from classical analysis and general topology, see, for example, Munkres [59].

As usual, $\mathbb{X} = \langle X, d, S\rangle$ denotes a computable metric space.

Definition 2.2.50. Let $A \subseteq X$. A *basic open cover* of A is a finite set

$$\mathcal{B} = \{B(a_0, r_0), \ldots, B(a_n, r_n)\} \tag{2.2.13}$$

where $n \in \mathbb{N}$, $A \subseteq \bigcup \mathcal{B}$, and for all $i \leq n$, $B(a_i, r_i)$ is a basic open ball. The cover $\mathcal{B}$ may of course be coded as $\langle a_0, r_0, \ldots, a_n, r_n\rangle$.

Remark 2.2.51. $U = \bigcup \mathcal{B}$ is effectively open, uniformly from a code for $\mathcal{B}$.

Definition 2.2.52. A compact set A is said to be *effectively compact* if the set of codes of basic open covers of A is computably enumerable. A compact space X is *effectively compact* if it is effectively compact as a set.

Examples 2.2.53. $[0,1]^n$, $\mathbb{T}^n$, $\{0,1\}^{\mathbb{N}}$, and $[0,1]^{\mathbb{N}}$ are effectively compact. More generally, if X is effectively compact, then X^n and $X^{\mathbb{N}}$ are too.

We collect two facts (and a definition) about effective compactness.

Fact 2.2.54. An effectively compact set K is effectively closed.

Proof. If $K \neq X$, by enumerating (computably) covers of K by basic open balls, then taking their closure, it is easy to write $K^c = X \setminus K$ as a union of uniformly Σ^0_1 sets. (And the case that $K = X$ is trivial.) □

Definition 2.2.55. $A \subseteq X$ is said to be *effectively totally bounded* if we can compute, from any $\epsilon > 0$, a basic open cover of A by balls of radius ϵ.

Fact 2.2.56. Let $K \subseteq X$ be compact. K is effectively compact if and only if K is effectively totally bounded. This can be made uniform.

[1] The terminology of effective compactness varies in the literature.

Proof. ($\Rightarrow$) If K is effectively compact, then K is certainly totally bounded. Therefore, given $\epsilon > 0$, if we enumerate (the codes of) basic open covers

$$\mathcal{B} = \{B(a_0, r_0), \ldots, B(a_n, r_n)\}$$

of K, we will eventually find one such that $r_i \leq \epsilon$ for all $i \leq n$, which yields a basic open ϵ-cover (i.e., cover by balls of radius ϵ) of K, as required.

($\Leftarrow$) Suppose K is effectively totally bounded. We can enumerate the codes of all finite sets $\mathcal{D} = \{B(a_0, r_0), \ldots, B(a_n, r_n)\}$ of basic open balls. By hypothesis, given $\epsilon > 0$, we can compute a basic open ϵ-cover

$$\{B(e_0, \epsilon), \ldots, B(e_m, \epsilon)\}$$

of K. From a code for $\mathcal{D}$, we can verify computably that

$$(\forall j \leq m)\ (\exists i \leq n) \quad d(e_j, a_i) + \epsilon < r_i. \tag{2.2.14}$$

Therefore, we can enumerate (the codes of) all finite sets $\mathcal{D}$ for which there exists an ϵ such that (2.2.14) holds. The claim is that this enumerates exactly the basic open covers of K.

Fix $\mathcal{D}$ and ϵ. For all $y \in K$, $d(y, e_j) < \epsilon$ for some $j \leq m$. If (2.2.14) holds,

$$d(y, a_i) \leq d(y, e_j) + d(e_j, a_i) < \epsilon + d(e_j, a_i) < r_i$$

for some $i \leq n$. Hence $y \in B(a_i, r_i)$, proving that $\mathcal{D}$ covers K.

Conversely, let $\mathcal{D}$ cover K. Let $\mathcal{B}$ be the family of open balls $B(x, \epsilon)$ such that $x \in K$, $\epsilon > 0$, and $d(x, a_i) + 2\epsilon < r_i$ for some $i \leq n$. $\mathcal{B}$ is an open cover of K, so it has a finite subcover $\mathcal{B}_0 = \{B(x_k, \epsilon_k)\}$.

Let $\delta = \min_k\{\epsilon_k\}$. Let $\epsilon < \delta$. $\mathcal{B}_0$ is a cover, so for all $j \leq m$, there exists a k such that $d(e_j, x_k) < \epsilon_k$. By definition of $\mathcal{B}_0$, there exists an $i \leq n$ such that $d(x_k, a_i) + 2\epsilon_k < r_i$. Hence

$$d(e_j, a_i) + \epsilon \leq d(e_j, x_k) + d(x_k, a_i) + \epsilon < \epsilon_k + d(x_k, a_i) + \epsilon_k < r_i,$$

which gives us (2.2.14). This proves the claim and completes the proof. □

In computable analysis, it is necessary in general to assume that compact sets satisfy a stronger property than effective compactness.

Definition 2.2.57. For non-empty $A \subseteq X$,

$$x \mapsto d(x, A) = \inf\{d(x, y) : y \in A\} \tag{2.2.15}$$

is the *distance function* $X \to [0, \infty)$ associated with A (and d).

Note that (classically) if A is compact, then the infimum is attained.

Definition 2.2.58. A non-empty, effectively compact $K \subseteq X$ is said to be *effectively located* in X if its distance function $d(\cdot, K)$ is computable.

The effectively located sets turn out to be the computable points in a certain computable metric space:

Definition 2.2.59. $\mathcal{H}(X)$ denotes the family of non-empty compact subsets of X. $\mathcal{H}(X)$ is a computable metric space under the *Hausdorff metric*

$$d_H(K_1, K_2) = \max\{\sup_{x \in K_1}\{d(x, K_2)\}, \sup_{y \in K_2}\{d(K_1, y)\}\}. \tag{2.2.16}$$

Both suprema are attained. For ideal points, take the family $\mathcal{S}$ of non-empty finite subsets of S (the enumerated set of ideal points for X).

Fact 2.2.60. Uniformly from a name for $K \in \mathcal{H}(X)$, we can

(i) compute the distance function associated with K,

(ii) enumerate the set of codes of basic open covers of K,

(iii) compute the diameter of K, $\operatorname{diam}(K) = \sup\{d(x, y) : x, y \in K\}$, and

(iv) compute a sequence $(x_j)_{j \in \mathbb{N}}$ such that $\overline{\{x_j : j \in \mathbb{N}\}} = K$.

Proof. The following is uniform from a name $(A_n)_{n \in \mathbb{N}}$ for K.

We begin with claim (i). For all $x \in X$, by the triangle inequality,

$$(\forall n \in \mathbb{N}) \quad |d(x, A_n) - d(x, K)| \leq d(A_n, K) < 2^{-n}.$$

A_n is finite and made up of ideal points, so $d(x, A_n)$ is computable uniformly from $n \in \mathbb{N}$ and a name for $x \in X$. This proves claim (i).

For claim (ii), by Fact 2.2.56, it suffices to show that K is effectively totally bounded. For all $i \in \mathbb{N}$, write $A_i = \{e_0, \ldots, e_m\}$. Then

$$\{B(e_0, 2^{-i}), \ldots, B(e_m, 2^{-i})\}$$

is a basic open cover of K, as required. This proves claim (ii).

For claim (iii), we appeal again to total boundedness. If we cover K with open balls $B(e_0, \epsilon), \ldots, B(e_m, \epsilon)$ such that $d(e_i, K) < \epsilon$ for all i, then $\operatorname{diam}(K)$ is within 2ϵ of $\max_{i,j} d(e_i, e_j)$. This proves claim (iii).

For claim (iv), fix $n \in \mathbb{N}$ and $a \in A_n$. Compute a sequence $(a_i)_{i \in \mathbb{N}} \subseteq S$ as follows. Let $a_0 = a \in A_n$ and proceed by induction. Suppose we have

already computed $a_0, \ldots, a_k$ such that $a_i \in A_{n+i}$ for all $i \leq k$ and $d(a_i, a_{i+1}) < 2^{-(n+i)+1}$ for all $i < k$. If A_{n+k+1} includes at least one point a' such that $d(a_k, a') < 2^{-(n+k)+1}$, we can compute such a point and set $a_{k+1} = a'$.

Since $a_k \in A_{n+k}$, there exists a $z \in K$ such that $d(a_k, z) < 2^{-(n+k)}$. Hence there exists an $a' \in A_{n+k+1}$ such that $d(z, a') < 2^{-(n+k)-1}$. It follows that $d(a_k, a') < 2^{-(n+k)+1}$, as required to compute the sequence $(a_i)_{i\in\mathbb{N}}$.

It is easy to see that $(a_i)_{i\in\mathbb{N}}$ converges computably. Ket x denote its limit, which satisfies $d(a_i, x) < 2^{-(n+i)+2}$ for all $i \in \mathbb{N}$. In particular, $d(a, x) < 2^{-n+2}$. Since $\lim_{i\to\infty} d(a_i, K) = 0$, x belongs to $\overline{K} = K$. Thus we assign, to each $n \in \mathbb{N}$ and each $a \in A_n$, a computable point $x = x(n, a) \in K$. Enumerate all of these as one sequence $(x_j)_{j\in\mathbb{N}} \subseteq K$ of uniformly computable points.

We just need to show that $\{x_j : j \in \mathbb{N}\}$ is dense in K. Fix $y \in K$ and $n \in \mathbb{N}$. There exists an $a \in A_n$ such that $d(y, a) < 2^{-n}$, and there exists a $j \in \mathbb{N}$ such that $d(a, x_j) < 2^{-n+2}$, so that $d(y, x_j) < 2^{-n+3}$. This proves claim (iv), which completes the proof. □

Fact 2.2.61. Let K be a non-empty compact set. K is computable in $\mathcal{H}(X)$ if and only if K is effectively located.

Proof. If K is computable in $\mathcal{H}(X)$, then K is effectively located by Fact 2.2.60 (ii). Conversely, suppose K is effectively located. Fix $i \in \mathbb{N}$. We can enumerate the codes of all finite sets of ideal points $A = \{a_0, \ldots, a_n\} \subseteq S$. By hypothesis, given A, we can verify computably that

(i) $\sup_{x\in A} d(x, K) < 2^{-i}$, and that

(ii) $B(a_0, 2^{-i}), \ldots, B(a_n, 2^{-i})$ is a basic open cover of K.

Together, (i) and (ii) imply that $d_H(A, K) < 2^{-i}$. Classically, for all $i \in \mathbb{N}$, there exists an A satisfying both conditions, which completes the proof. □

Definition 2.2.62. $A \subseteq X$ is said to be *effectively separable* if there exists a sequence $(x_j)_{j\in\mathbb{N}}$ of uniformly computable points such that $A = \overline{\{x_j : j \in \mathbb{N}\}}$.

Corollary 2.2.63. If K is effectively located, then it is effectively separable. Hence K is a computable metric subspace of X: for ideal points, compute any dense sequence in K.

That brings us to the effective version of local compactness. $\mathbb{X} = \langle X, d, S\rangle$ continues to denote a computable metric space.

Definition 2.2.64. X is said to be *effectively locally compact* if there exists a sequence $(K_i)_{i\in\mathbb{N}}$ of non-empty compact sets such that

(i) $X = \bigcup_{i\in\mathbb{N}} K_i$ and $K_i \subseteq K_{i+1}^\circ$ (the interior of K_{i+1}) for all $i \in \mathbb{N}$,

(ii) K_i is effectively located, uniformly from i, for all $i \in \mathbb{N}$, and

(iii) K_i° is effectively open, uniformly from i, for all $i \in \mathbb{N}$.

In that case, $(K_i)_{i\in\mathbb{N}}$ is said to be an *effective exhausting sequence* for X.

Brattka [14] deduces

Fact 2.2.65. Let $(K_i)_{i\in\mathbb{N}}$ be an effective exhausting sequence. Uniformly from a name for $x \in X$, we can compute an $i \in \mathbb{N}$ such that $x \in K_i^\circ$.

Example 2.2.66. For each $n \in \mathbb{N}$, $\mathbb{R}^n$ is effectively locally compact with an effective exhausting sequence defined by $K_i = [-i, i]^n$ for all $i \in \mathbb{N}$.

It is convenient and sometimes necessary to impose extra conditions on the effective exhausting sequence in relation to a given metric d' on X.

Definition 2.2.67. A metric d' on X is said to be *adapted to* $(K_i)_{i\in\mathbb{N}}$ if $d' \leq 1$ and for all $i \in \mathbb{N}$, if $x \in K_i$ and $d'(x, y) < 1$, then $y \in K_{i+1}^\circ$.

Example 2.2.68. In $\mathbb{R}^n$, the metric $d'(x, y) = \min\{1, |x - y|\}$ is adapted to the effective exhausting sequence $(K_i)_{i\in\mathbb{N}}$ from Example 2.2.66.

In Example 2.2.68, the adapted metric was computably related to the original metric. Brattka [14] shows that we can achieve this in general:

Fact 2.2.69. Let $(K_i)_{i\in\mathbb{N}}$ be an effective exhausting sequence. X has an adapted metric d' that is computably related to d. Moreover, the Hausdorff metric d'_H associated with d' is computably related to d_H.

Classically, every locally compact separable metric space X admits a metrizable one-point compactification X^* (see, for example, Kechris [48]). Mandelkern [54] shows that X^* is constructively metrizable, and we will adapt his method to show that X^* is computably metrizable.

Definition 2.2.70. For non-empty $A \subseteq X$ and $\epsilon > 0$, define the open set

$$A^\epsilon = \{x \in X : (\exists a \in A)\ d(x, a) < \epsilon\} = \{x \in X : d(x, A) < \epsilon\}.$$

Remark 2.2.71. If $K \in \mathcal{H}(X)$, then by Fact 2.2.60, K^ϵ is effectively open uniformly from names for ϵ and K.

Lemma 2.2.72. Let $(K_i)_{i\in\mathbb{N}}$ be an effective exhausting sequence. Given $i \in \mathbb{N}$, we can compute an $\epsilon > 0$ such that $K_i^\epsilon \subseteq K_{i+1}^\circ$.

Proof. Classically, in any metric space, if K is compact, F is closed, and $K \cap F = \varnothing$, then K and F are separated:

$$d(K, F) = \inf\{d(x, y) : x \in K, y \in F\} > 0.$$

Fixing $i \in \mathbb{N}$ and setting $K = K_i$ and $F = X \setminus K^\circ_{i+1}$, we have $K^\epsilon_i \subseteq K^\circ_{i+1}$ for all sufficiently small ϵ. We just need to find one such ϵ.

Write $K^\circ_{i+1} = \bigcup_{j\in\mathbb{N}} B(a_j, r_j)$. Enumerate the basic open covers of K_i. Given a cover $\{B(e_0, \epsilon_0), \dots, B(e_n, \epsilon_n)\}$, we can verify computably that

$$(\forall i \le n)\ (\exists j \in \mathbb{N}) \quad d(e_i, a_j) + 2\epsilon_i < r_j.$$

Classically, such a cover exists. Set $\epsilon = \min\{\epsilon_0, \dots, \epsilon_n\}$. If $d(x, y) < \epsilon$ and $y \in K_i$, then there exists an $i \in \mathbb{N}$ such that $d(y, e_i) < \epsilon_i$. Hence

$$(\exists j \in \mathbb{N}) \quad d(x, a_j) \le d(x, y) + d(y, e_i) + d(e_i, a_j) < \epsilon + \epsilon_i + d(e_i, a_j) < r_j.$$

So $x \in K^\circ_{i+1}$, which proves that $K^\epsilon_i \subseteq K^\circ_{i+1}$. □

Now we can effectivize the one-point compactification:

Theorem 2.2.73. Let $(K_i)_{i\in\mathbb{N}}$ be an effective exhausting sequence. There exists an effectively compact space $\langle X^*, d^*, S^*\rangle$ which is an *effective one-point compactification* of X in the following sense:

(i) There exists a point $* \in X^*$ and a one-to-one computable function $\theta : X \to X^*$ such that $\theta(X) = X^* \setminus \{*\}$.

(ii) It is also the case that θ^{-1} is computable on $\theta(K)$, uniformly from a name for K, for all $K \in \mathcal{H}(X)$.

(iii) If X is *not* compact, then $\theta(X)$ is dense in X^*. Otherwise, $*$ is isolated.

Proof. The proof is adapted from the constructive proof of Theorem 2 in Mandelkern [54]. By Lemma 2.2.72, we can compute a decreasing sequence $(\gamma_i)_{i\in\mathbb{N}}$ of positive rationals such that for all $i \in \mathbb{N}$, $\gamma_i < 2^{-i-1}$ and $K^{\gamma_i}_i \subseteq K^\circ_{i+1}$.

Using $(\gamma_i)_{i\in\mathbb{N}}$, define a function $h : X \to (0, 1/2)$ as follows:

$$h(x) = \sup_{i\in\mathbb{N}}\{\gamma_i - d(x, K_i)\}. \tag{2.2.17}$$

By Fact 2.2.65, uniformly from a name for x, we can compute an $n \in \mathbb{N}$ such that $d(x, K_i) = 0$ for all $i \ge n$. In particular, the supremum in (2.2.17) is the maximum of finitely many terms. Using Fact 2.2.60 (i), it is easy to see that $h(x)$ is computable, uniformly from a name for x, for all $x \in X$.

Note that $h(x) \geq \gamma_i$ for $x \in K_i$ and $h(x) < \gamma_i$ otherwise. Next, we show that $|h(x) - h(y)| \leq d(x,y)$ for all $x, y \in X$. Fix x, y and $\epsilon > 0$. For some $i \in \mathbb{N}$, $\gamma_i - d(x, K_i) > h(x) - \epsilon$. By the definition of h, $\gamma_i - d(y, K_i) \leq h(y)$, so

$$h(x) - h(y) < \gamma_i - d(x, K_i) + \epsilon - (\gamma_i - d(y, K_i)) \leq d(x,y) + \epsilon,$$

which proves that $|h(x) - h(y)| \leq d(x,y)$.

$*$ is a new symbol: a formal "point at infinity." We define a metric d^* on $X_1 = X \cup \{*\}$ as follows: for all $x, y \in X$,

$$d^*(x,y) = \min\{d(x,y), h(x) + h(y)\};$$

for $x \in X$, $d^*(x,*) = h(x)$; and of course $d^*(*,*) = 0$. It is not difficult to prove (classically) that d^* is a metric. Note that $d^* \leq 1$, $d^* \leq d$ on $X \times X$, and d^* is computable of type $X \times X \to [0, \infty)$.

Let X^* be the completion of $\langle X_1, d^* \rangle$. Extend d^* to X^*. Let $\theta : X_1 \to X^*$ be the canonical inclusion mapping; denote $\theta(*)$ by $*$. Then $\theta(S) \cup \{*\}$ is dense in X^*, so in particular X^* is separable. Writing $S = \{s_n : n \in \mathbb{N}\}$, we can easily enumerate $S^* = \theta(S) \cup \{*\}$ as

$$(u_n)_{n \in \mathbb{N}} = (*, \theta(s_0), \theta(s_1), \ldots).$$

It is clear that $d^*(u_n, u_m)$ is computable uniformly from $n, m \in \mathbb{N}$.

The compact space $\langle X^*, d^*, S^* \rangle$ is therefore a computable metric space. By Fact 2.2.56, X^* is effectively compact if it is effectively totally bounded.

Given $\epsilon > 0$, compute an $i \in \mathbb{N}$ such that $\gamma_i < \epsilon$ and compute a basic open cover $\{B(a_0, \epsilon), \ldots, B(a_n, \epsilon)\}$ of K_i. Let

$$\sigma = \langle \theta(a_0), \ldots, \theta(a_n), * \rangle.$$

For all $x \in E$, either $x \in K_i$, in which case $d^*(x, x_j) < \epsilon$ for some $j \leq n$, or $x \in X \setminus K_i$, in which case $d^*(x,*) = h(x) < \gamma_i < \epsilon$. Hence σ covers X^*, which proves that X^* is effectively totally bounded.

The restriction of $\theta : X_1 \to X^*$ to X is also denoted by θ. For all $x, y \in X$, $d^*(\theta(x), \theta(y)) = d^*(x,y) \leq d(x,y)$. Since θ maps S to S^*, it follows that θ is computable of type $X \to X^*$. Next, we will show that $\theta(X) = X^* \setminus \{*\}$.

For all $x \in X$, $d^*(\theta(x), *) = h(x) > 0$, so $\theta(X) \subseteq X^* \setminus \{*\}$. For the reverse inclusion, fix $y \in X^*$ such that $d^*(y,*) > 0$. Choose a $j \in \mathbb{N}$ so that $d^*(y,*) > \gamma_j$. Write y as the limit of a sequence $(y_n)_{n \in \mathbb{N}}$ in X^*. For all sufficiently large $n \in \mathbb{N}$, $d^*(\theta(y_n), *) > \gamma_j$. Hence $(y_n)_{n \in \mathbb{N}} \subseteq X$ without loss of generality. For all $n \in \mathbb{N}$, $h(y_n) > \gamma_j$ so $y_n \in K_j$. Since $(y_n)_{n \in \mathbb{N}}$ is Cauchy under d^* and h is bounded away from zero on $(y_n)_{n \in \mathbb{N}}$, that sequence is

Cauchy under d. Hence there exists a point $x \in K_j$ such that $y_n \to x$ with respect to d, which implies that $y = \lim \theta(y_n) = \theta(x)$ belongs to $\theta(X)$. Therefore, $\theta(X) \supseteq X^* \setminus \{*\}$, proving their equality.

Finally, we will show that θ^{-1} is computable on $\theta(K)$, uniformly from a name for K, for all $K \in \mathcal{H}(X)$. Using our name for K, compute a point $y \in K$ such that $d^*(*, y) \leq 2d^*(*, K)$. Since $y \in X^* \setminus \{*\}$, $d^*(*, y) > 0$. For $x \in \theta^{-1}(K)$, $\theta(x) \in K$, so $d^*(*, \theta(x)) = d^*(*, x) = h(x)$. Hence h is bounded below on $\theta^{-1}(K)$, which implies that $\theta^{-1} : K \to X$ is effectively uniformly continuous. Since it maps $\theta(S)$ to S, θ^{-1} is computable on K, uniformly from the name for K.

Classically, it is well known that $\theta(X)$ is dense in X^* if X is not compact, and $*$ is isolated in X^* otherwise. This completes the proof. □

2.2.3 Continuous function spaces

Effective local compactness gives us two generalizations of the computable metric space $\mathbf{C}[a, b]$. The first of these was developed by Brattka [14].

For the relevant concepts from classical functional analysis, Conway [23] or another standard text should be more than sufficient.

Let $\langle X, d, S_X \rangle$ be effectively locally compact with effective exhausting sequence $(K_i)_{i \in \mathbb{N}}$, and let $\langle Y, \| \cdot \|, S_Y \rangle$ be a computable Banach space.

Definition 2.2.74. $\mathbf{C}(X, Y)$ denotes the set of (total) continuous functions $f : X \to Y$. For all $i \in \mathbb{N}$ and $f \in \mathbf{C}(X, Y)$, define the semi-norm

$$\|f\|_{K_i} = \sup\{\|f(x)\| : x \in K_i\}. \tag{2.2.18}$$

$\mathbf{C}(X, Y)$ is a complete, separable metric space under the *Fréchet metric*

$$d_{\mathbf{C}}(f, g) = \sum_{i \in \mathbb{N}} 2^{-i-1} \cdot \frac{\|f - g\|_{K_i}}{1 + \|f - g\|_{K_i}}, \tag{2.2.19}$$

which is bounded by 1 and induces the *compact-open topology*.

Notation 2.2.75. When $Y = \mathbb{R}$, we may write $\mathbf{C}(X)$ instead of $\mathbf{C}(X, \mathbb{R})$.

Next, we choose ideal points for $\mathbf{C}(X, Y)$, starting with the case $Y = \mathbb{R}$.

Definition 2.2.76. Let $\mathcal{F} \subseteq \mathbf{C}(X)$. $\mathcal{F}$ is said to *separate points* if for all $x, y \in X$ such that $x \neq y$, there exists an $f \in \mathcal{F}$ such that $f(x) \neq f(y)$.

Definition 2.2.77. Let $\mathcal{F} \subseteq \mathbf{C}(X)$. A *polynomial of $\mathcal{F}$ over $\mathbb{Q}$* is a function $f : X \to \mathbb{R}$ of the form

$$f = p(f_0, \ldots, f_n) \tag{2.2.20}$$

where $n \in \mathbb{N}$, $p \in \mathbb{Q}[x_0, \ldots, x_n]$, and $f_0, \ldots, f_n \in \mathcal{F}$. We may denote the set of all polynomials of $\mathcal{F}$ over $\mathbb{Q}$ by $\mathrm{Poly}_{\mathbb{Q}}(\mathcal{F})$.

As a consequence of the Stone-Weierstrass theorem, we have

Fact 2.2.78. If $\mathcal{F}$ separates points, then $\mathrm{Poly}_{\mathbb{Q}}(\mathcal{F})$ is dense in $\mathbf{C}(X)$.

Definition 2.2.79. Let $\mathcal{F} \subseteq \mathbf{C}(X)$ and $A \subseteq Y$. A *linear combination of $\mathcal{F}$ over A* is a function $f : X \to Y$ of the form

$$f(x) = \sum_{i \leq k} f_i(x) a_i \tag{2.2.21}$$

where $k \in \mathbb{N}$, $f_0, \ldots, f_k \in \mathcal{F}$, and $a_0, \ldots, a_k \in A$. We may denote the set of all linear combinations of $\mathcal{F}$ over A by $\mathrm{Lin}_A(\mathcal{F})$.

By the Stone-Weierstrass theorem again, we have

Fact 2.2.80. If $\mathcal{F}$ separates points and A is dense in Y, then

$$\mathrm{Lin}_A(\mathrm{Poly}_{\mathbb{Q}}(\mathcal{F})) \tag{2.2.22}$$

(that is, the set of linear combinations of polynomials) is dense in $\mathbf{C}(X, Y)$.

Definition 2.2.81. For $s \in S_X$, let $d_s(x) = d(x, s)$. Then

$$\mathcal{F}_{S_X} = \{d_s : s \in S_X\} \subseteq \mathbf{C}(X) \tag{2.2.23}$$

is a canonical sequence of functions that separates points. Enumerate

$$\mathrm{Lin}_{S_Y}(\mathrm{Poly}_{\mathbb{Q}}(\mathcal{F}_{S_X})) = \{\psi_i : i \in \mathbb{N}\}. \tag{2.2.24}$$

Brattka [14] proves the following five facts.

Fact 2.2.82. $\langle \mathbf{C}(X, Y), d_{\mathbf{C}}, (\psi_i)_{i \in \mathbb{N}} \rangle$ is a computable metric space.

Remark 2.2.83. If X is compact, we can just set $K_i = X$ for all $i \in \mathbb{N}$. Then $d_{\mathbf{C}} = d_\infty/(1 + d_\infty)$ is computably related to d_∞.

Fact 2.2.84. The computable points in $\mathbf{C}(X, Y)$ are exactly the (total) computable functions $X \to Y$. This can be made uniform.

Note that, in light of Remark 2.2.83, Fact 2.2.84 generalizes Fact 2.2.49, the computable Weierstrass approximation theorem.

Fact 2.2.85. Let X_1, X_2 be effectively locally compact, and let $f : X_1 \times X_2 \to Y$. The following are equivalent, and this can be made uniform:

(a) $\langle x_1, x_2\rangle \mapsto f(x_1, x_2)$ is computable of type $X_1 \times X_2 \to Y$

(b) $x_1 \mapsto (x_2 \mapsto f(x_1, x_2))$ is computable of type $X_1 \to \mathbf{C}(X_2, Y)$

Fact 2.2.86. The following elementary function operations are computable.

(i) Evaluation: $\langle f, x\rangle \mapsto f(x)$, which is of type $\mathbf{C}(X,Y) \times X \to Y$

(ii) Addition: $\langle f, g\rangle \mapsto f + g$, of type $\mathbf{C}(X,Y) \times \mathbf{C}(X,Y) \to \mathbf{C}(X,Y)$

(iii) Scalar multiplication: $\langle a, f\rangle \mapsto af$, of type $\mathbf{C}(X) \times \mathbf{C}(X,Y) \to \mathbf{C}(X,Y)$

(iv) Norm: $f \mapsto \|f\|$, of type $\mathbf{C}(X,Y) \to \mathbf{C}(X)$

Fact 2.2.87. The real number $\|f\|_{K_i}$ from Definition 2.2.74 is computable uniformly from $i \in \mathbb{N}$ and a name for $f \in \mathbf{C}(X,Y)$.

At this point, we have managed to generalize from $\mathbf{C}[a,b]$ to the space $\mathbf{C}(X,Y)$, where X is effectively locally compact and Y is a computable Banach space—but to do this, we had to change the metric from d_∞ to $d_{\mathbf{C}}$. Next, we describe a different way to generalize $\mathbf{C}[a,b]$ which allows us to continue using the sup-norm.

$\langle X, d, S_X\rangle$ is still effectively locally compact with effective exhausting sequence $(K_i)_{i\in\mathbb{N}}$, and $Y = \langle Y, \|\cdot\|, S_Y\rangle$ is still a computable Banach space.

Definition 2.2.88. A function $f : X \to Y$ is said to *vanish at infinity* if for all $\epsilon > 0$ there exists a compact $K \subseteq X$ such that $\|f(x)\| < \epsilon$ for all $x \notin K$.

A *vanishing rate* for a function f is a sequence $(i_n)_{n\in\mathbb{N}} \subseteq \mathbb{N}$ such that for all $n \in \mathbb{N}$, $\|f(x)\| < 2^{-n}$ for all $x \notin K_{i_n}$.

Fact 2.2.89. f vanishes at infinity if and only if f has a vanishing rate.

Definition 2.2.90. $\mathbf{C}_0(X,Y)$ denotes the space of continuous functions $f : X \to Y$ vanishing at infinity. $\mathbf{C}_0(X,Y)$ is a separable Banach space with respect to the *sup-norm*

$$\|f\|_\infty = \sup\{\|f(x)\| : x \in X\}, \tag{2.2.25}$$

which induces the *uniform metric* d_∞ and the *uniform topology*.

Notation 2.2.91. We may write $\mathbf{C}_0(X)$ instead of $\mathbf{C}_0(X, \mathbb{R})$.

The uniform topology, though not the metric, extends to $\mathbf{C}(X,Y)$:

Definition 2.2.92. The *uniform topology* on $\mathbf{C}(X,Y)$ is the topology of uniform convergence: $f_n \to f$ if and only if $\sup\{\|f_n(x) - f(x)\| : x \in X\} \to 0$.

Note that $\mathbf{C}_0(X,Y)$ is a closed vector subspace of $\mathbf{C}(X,Y)$ with the uniform topology, and the restricted topology is induced by the sup-norm. Also, if X is compact, then evidently $\mathbf{C}_0(X,Y) = \mathbf{C}(X,Y)$.

Now, the set $\mathbf{C}_b(X,Y)$ of bounded continuous functions $X \to Y$ is also a closed vector subspace of $\mathbf{C}(X,Y)$ under the uniform topology, hence a Banach space under the sup-norm. However, $\mathbf{C}_b(X,Y)$ is not separable in general, as the following example shows.

Example 2.2.93. $\mathbf{C}(\mathbb{R},[0,1])$ is not separable: for all $\alpha \in \{0,1\}^{\mathbb{Z}}$, define a continuous function $f_\alpha : \mathbb{R} \to [0,1]$ such that $f_\alpha(n) = \alpha_n$ for all $n \in \mathbb{Z}$. Then $\|f_\alpha - f_\beta\|_\infty = 1$ whenever $\alpha \neq \beta$, and $\{0,1\}^{\mathbb{Z}}$ is uncountable.

We distinguish a well-behaved dense subset of $\mathbf{C}_0(X,Y)$: the functions with compact support.

Definition 2.2.94. The *support* of a function $f : X \to Y$ is given by

$$\operatorname{supp}(f) = \overline{\{x \in X : f(x) \neq 0\}}. \tag{2.2.26}$$

Definition 2.2.95. $\mathcal{K}(X,Y)$ denotes the set of (total) continuous functions $f : X \to Y$ with compact support. $\mathcal{K}(X,Y)$ is a vector subspace of $\mathbf{C}_0(X,Y)$, and it is dense with respect to the sup-norm.

Notation 2.2.96. We may write $\mathcal{K}(X)$ instead of $\mathcal{K}(X,\mathbb{R})$.

Going forward, we need an auxiliary function that modifies continuous functions to make them vanish at infinity:

Definition 2.2.97. For all $i \in \mathbb{N}$, define a function $\eta_i : X \to \mathbb{R}$ by

$$\eta_i(x) = \max\{0, 1 - \operatorname{dist}(x, K_i)\}. \tag{2.2.27}$$

Note that η_i is computable uniformly from $i \in \mathbb{N}$, and $0 \leq \eta_i \leq 1$.

Remark 2.2.98. By Fact 2.2.69, we can assume without loss of generality that d is adapted to $(K_i)_{i\in\mathbb{N}}$. In that case, $\operatorname{supp}(\eta_i) = K_{i+1}$.

Let $f \in \mathbf{C}(X,Y)$. Evidently, for all $i \in \mathbb{N}$, $\eta_i \cdot f : X \to Y$ is continuous and equals f on K_i, and its support is $\operatorname{supp}(f) \cap K_{i+1}$.

Proposition 2.2.99. The function $\langle f, i\rangle \mapsto \eta_i \cdot f$ is computable of type $\mathbf{C}(X,Y) \times \mathbb{N} \to \mathbf{C}(X,Y)$.

Proof. By Fact 2.2.60 (i) and Fact 2.2.84, $\eta_i \in \mathbf{C}(X)$ is computable uniformly from $i \in \mathbb{N}$, and the result follows from Fact 2.2.86 (iii). □

The next result is fairly obvious, but we will work through the details.

Proposition 2.2.100. If $\{f_j : j \in \mathbb{N}\}$ is dense in $\mathbf{C}(X,Y)$ with respect to $d_{\mathbf{C}}$, then $\{\eta_i \cdot f_j : i,j \in \mathbb{N}\}$ is dense in $\mathbf{C}_0(X,Y)$ with respect to d_∞.

Proof. Let $g \in \mathbf{C}_0(X,Y)$, let $(i_n)_{n\in\mathbb{N}}$ be a vanishing rate for g, and let $n \in \mathbb{N}$. We will identify a pair $\langle i,j\rangle$ such that $\|g - \eta_i \cdot f_j\|_\infty \le 3\cdot 2^{-n}$.

Set $i = i_n$ so that $\|g(x)\| \le 2^{-n}$ for all $x \notin K_i$. Choose a $j \in \mathbb{N}$ such that $d_{\mathbf{C}}(g, f_j) \le 2^{-n-i-2}$. Incidentally, this shows that we can compute $\langle i,j\rangle$ uniformly from a name for $g \in \mathbf{C}(X,Y)$ and a vanishing rate for g.

We claim that $\|g(x) - \eta_i(x)\cdot f_j(x)\| \le 3\cdot 2^{-n}$ for all $x \in X$. There are three cases to consider: $x \in K_i$, $x \in K_{i+1} \smallsetminus K_i$, and $x \notin K_{i+1}$.

First, on K_i, $\eta_i \cdot f_j = f_j$. Since $d_{\mathbf{C}}(g,f_j) \le 2^{-n-i-2}$,

$$\|g - f_j\|_{K_i} \le 2^{i+1}2^{-n-i-2} = 2^{-n-1}.$$

So $\|g(x) - \eta_i(x)\cdot f_j(x)\| \le 2^{-n-1} < 3\cdot 2^{-n}$ for all $x \in K_i$.

Next, for $x \in K_{i+1}\smallsetminus K_i$, $\|g(x)\| \le 2^{-n}$. Since $d_{\mathbf{C}}(g,f_j) \le 2^{-n-i-2}$,

$$\|g - f_j\|_{K_{i+1}} \le 2^{i+2}2^{-n-i-2} = 2^{-n}.$$

It follows that $\|\eta_i(x)\cdot f_j(x)\| \le \|f_j(x)\| \le 2^{-n} + 2^{-n}$, which implies that

$$\|g(x) - \eta_i(x)\cdot f_j(x)\| \le 2^{-n} + 2^{-n} + 2^{-n} = 3\cdot 2^{-n}.$$

Finally, for $x \notin K_{i+1}$, $\eta_i(x)\cdot f_j(x) = 0$ and $\|g(x)\| \le 2^{-n}$. In all cases, $\|g(x) - \eta_i(x)\cdot f_j(x)\| \le 3\cdot 2^{-n}$, as required. □

Now recall the ideal points $\{\psi_i : i \in \mathbb{N}\}$ for $\mathbf{C}(X,Y)$ (Definition 2.2.81). Corollary 2.2.87 and Proposition 2.2.100 imply

Corollary 2.2.101. $\mathbf{C}_0(X,Y)$ is a computable Banach space with respect to the sup-norm: for ideal points, take $\{\eta_i \cdot \psi_j : i,j \in \mathbb{N}\}$.

With our new ideal points, it is easy to prove

Proposition 2.2.102. The evaluation operator $\langle f, x\rangle \mapsto f(x)$ is computable of type $\mathbf{C}_0(X,Y)\times X \to Y$.

Naturally, the space $\langle \mathbf{C}_0(X,Y), d_\infty\rangle$ is closely related to $\mathbf{C}_0(X,Y)$ as a subspace of $\langle \mathbf{C}(X,Y), d_{\mathbf{C}}\rangle$:

Proposition 2.2.103. Let $f \in \mathbf{C}_0(X,Y)$.

(i) Uniformly from a name for f in $\langle \mathbf{C}_0(X,Y), d_\infty \rangle$, we can compute a name for f in $\langle \mathbf{C}(X,Y), d_{\mathbf{C}} \rangle$ and a vanishing rate for f.

(ii) Conversely, a name for f in $\langle \mathbf{C}_0(X,Y), d_\infty \rangle$ is computable uniformly from a name for f in $\langle \mathbf{C}(X,Y), d_{\mathbf{C}} \rangle$ and a vanishing rate for f.

Proof. Suppose we are given a name for f in $\langle \mathbf{C}_0(X,Y), d_\infty \rangle$. Then it is easy to compute a vanishing rate for f, because if $\|f - \eta_i \cdot \psi_j\|_\infty < \epsilon$, then $\|f\| < \epsilon$ outside K_{i+1}. Since moreover $d_{\mathbf{C}}(f, \eta_i \cdot \psi_j) < \epsilon$ provided that $\epsilon < 1$, Proposition 2.2.99 lets us compute a name for f in $\langle \mathbf{C}_0(X,Y), d_{\mathbf{C}} \rangle$.

Conversely, given a name for f in $\langle \mathbf{C}(X,Y), d_{\mathbf{C}} \rangle$ and a vanishing rate for f, Corollary 2.2.87 lets us approximate f in sup-norm. □

Our effective one-point compactification X^* (Theorem 2.2.73) happens to be consistent with our computable function space $\mathbf{C}_0(X,Y)$:

Proposition 2.2.104. If f is a computable point in $\mathbf{C}_0(X,Y)$, then the extension $\widetilde{f}$ of f to X^* by setting $\widetilde{f}(*) = 0$ is computable of type $X^* \to Y$. $\widetilde{f}$ is also a computable point in $\mathbf{C}(X^*, Y)$. This can be made uniform.

Proof. The case that X is compact (and $*$ is isolated) is trivial, so let us assume that X is not compact (and X is dense in X^*).

The extension of f to X^* is clearly computable on X, which leaves only $\widetilde{f}(*)$. Suppose $(x_n)_{n\in\mathbb{N}} \subseteq X$ converges to $*$ computably with respect to d^*. Then it is enough to show that $f(x_n) \to 0$ at a computable rate (in Y).

By definition, $d^*(x_n, *) = h(x_n)$; that is, $h(x_n) \to 0$ computably. By the definition of h, we can compute a sequence $(n_i)_{i\in\mathbb{N}} \subseteq \mathbb{N}$ such that for all $i \in \mathbb{N}$, for all $n \geq n_i$, x_n lies outside K_i. But we can also use Proposition 2.2.103 to compute a vanishing rate (Definition 2.2.88) for f. Combine that vanishing rate with the sequence $(n_i)_{i\in\mathbb{N}}$ to see that $f(x_n) \to 0$ at a computable rate.

Finally, $\widetilde{f}$ is a computable point in $\mathbf{C}(X^*, Y)$ by Fact 2.2.84. □

2.3 Computable measure theory

At this point, we have all the tools from computable analysis we need to study computable measure theory. Again, that theory has been developed, quite recently, by Gács [39], Hoyrup and Rojas [47], Rute [67], Weihrauch [72], Schröder [68], and Bosserhoff [13], among others (see Section 1.1).

Rute [67] provides a very useful overview; unless otherwise noted, our own review is based on that source. We assume the reader is familiar with the principles of classical measure theory as in, say, Cohn [21].

2.3.1 Measure

We begin by defining a metric space whose points are probability measures.

Let $\mathbb{X} = \langle X, d, S\rangle$ be a computable metric space. Of course, $\mathbb{X}$ is a measurable space as well, equipped with the Borel σ-algebra $\mathscr{B}(\mathbb{X})$.

Definition 2.3.1. $\mathcal{M}_1(\mathbb{X})$ denotes the set of probability measures on $\mathbb{X}$. $\mathcal{M}_1(\mathbb{X})$ is a complete, separable metric space under the *Prokhorov metric*

$$\pi(\mu,\nu) = \inf\{\epsilon > 0 : \mu(A) \leq \nu(A^\epsilon) + \epsilon \text{ for each Borel set } A\}. \tag{2.3.1}$$

The topology induced by π is called the *weak topology* on $\mathcal{M}_1(\mathbb{X})$.

Fact 2.3.2. The weak topology is the finest (i.e., largest) topology on $\mathcal{M}_1(\mathbb{X})$ satisfying the following convergence property:

$$\mu_n \to \mu \iff (\forall f \in \mathbf{C}_b(X)) \quad \lim_n \int f d\mu_n = \int f d\mu. \tag{2.3.2}$$

Gács [39] proves

Fact 2.3.3. $\mathcal{M}_1(\mathbb{X})$ is a computable metric space under the Prokhorov metric. For ideal points, take the measures concentrated on finitely many atoms $s_0, \ldots, s_n \in S$, assigning each a rational measure.

Definition 2.3.4. A probability measure μ on $\mathbb{X}$ is said to be *computable* if μ is computable in $\mathcal{M}_1(\mathbb{X})$. A *computable probability space* is a pair $\langle \mathbb{X}, \mu\rangle$ where $\mathbb{X}$ is a computable metric space and $\mu \in \mathcal{M}_1(\mathbb{X})$ is computable.

Remark 2.3.5. Suppose ν is a finite, nontrivial measure on $\mathbb{X}$. Then $\mu = \nu/\nu(X)$ is a probability measure, and we can identify ν with the pair $\langle \mu, \nu(X)\rangle \in \mathcal{M}_1(\mathbb{X}) \times [0,\infty)$. In general, if $\langle \mu, M\rangle$ is a computable point in the product space, we can call $\nu = M\mu$ a *computable measure* on $\mathbb{X}$.

The next two facts effectivize measure and integration on a computable metric space [47]. Recall that $\tau(\mathbb{X})$ denotes the topology of $\mathbb{X}$.

Fact 2.3.6. $\mu(U)$ is lower semicomputable uniformly from names for $\mu \in \mathcal{M}_1(\mathbb{X})$ and $U \in \tau(\mathbb{X})$. $\mu \in \mathcal{M}_1(\mathbb{X})$ is computable if and only if $\mu(U)$ is lower semicomputable, uniformly from a name for U, for all $U \in \tau(\mathbb{X})$.

Fact 2.3.7. Let $f : X \to \mathbb{R}$ be computable and bounded by $M > 0$. Then the integral operator

$$\mu \mapsto \int f d\mu \tag{2.3.3}$$

is computable of type $\mathcal{M}_1(\mathbb{X}) \to \mathbb{R}$, and this can be made uniform in names for f and M. Moreover, μ is computable if and only if $\int f d\mu$ is computable, uniformly from a name for f, for all computable $f : X \to [0,1]$.

When X is effectively locally compact, we can define the continuous function spaces $\mathbf{C}(X, Y)$ and $\mathbf{C}_0(X, Y)$ from Part 2.2.3.

Lemma 2.3.8. Let $(K_i)_{i\in\mathbb{N}}$ be an effective exhausting sequence for $\mathbb{X}$. Then there exists a computable sequence $(i_n)_{n\in\mathbb{N}} \subseteq \mathbb{N}$ such that for all $n \in \mathbb{N}$,

$$\mu(X \setminus K_{i_n}) < 2^{-n}. \tag{2.3.4}$$

Proof. This follows from Fact 2.3.6 when we note that $\mu(K_i) \uparrow 1$. □

Proposition 2.3.9. Let $\mathbb{X}$ be effectively locally compact. The integral operator $\langle \mu, f\rangle \mapsto \int f d\mu$ is

(i) computable of type $\mathcal{M}_1(\mathbb{X}) \times \mathbf{C}(\mathbb{X}, [0,1]) \to [0,1]$, and

(ii) computable of type $\mathcal{M}_1(\mathbb{X}) \times \mathbf{C}_0(\mathbb{X}) \to \mathbb{R}$.

Moreover, if $\mu \in \mathcal{M}_1(\mathbb{X})$, then the following are equivalent:

(a) μ is computable

(b) $\int f d\mu$ is computable uniformly from a name for $f \in \mathbf{C}(\mathbb{X}, [0,1])$

(c) $\int f d\mu$ is computable uniformly from a name for $f \in \mathbf{C}_0(\mathbb{X})$

Proof. With Lemma 2.3.8, all claims follow easily from Fact 2.3.7. □

We list a few important computable probability spaces:

Example 2.3.10. The unit cube $[0,1]^n$ is a computable probability space under Lebesgue measure, denoted as usual by λ, for each $n \in \mathbb{N}$.

Example 2.3.11. Cantor space $\{0,1\}^{\mathbb{N}}$ (Example 2.2.44) is a computable probability space under the *fair coin measure*

$$P([\sigma]) = 2^{-|\sigma|} \tag{2.3.5}$$

where $\sigma \in \{0,1\}^*$ and $[\sigma]$ is the associated cylinder set (Definition 2.2.43).

Example 2.3.12. $\mathbf{C}[a,b]$ (Example 2.2.48) is a computable probability space under Wiener measure; for details, see Fouché [33].

We have mentioned the need for an effective basis (Definition 2.2.26) for the underlying space that is, in some sense, well behaved.

Definition 2.3.13. Let $\langle \mathbb{X}, \mu \rangle$ be a computable probability space. A set $A \subseteq X$ is said to be *μ-almost decidable*, or just *almost decidable*, if there exist effectively open sets $U, V \subseteq X$ such that $U \subseteq A$, $V \subseteq A^c$, $U \cup V$ is dense, and $\mu(U \cup V) = 1$. Then A is said to be almost decidable by the pair $\langle U, V \rangle$.

The elements of a sequence $(A_i)_{i \in \mathbb{N}}$ of sets are said to be *uniformly almost decidable* if there exists a sequence of uniformly computable pairs $\langle U_i, V_i \rangle$ such that A_i is almost decidable by $\langle U_i, V_i \rangle$ for all $i \in \mathbb{N}$.

Remark 2.3.14. An almost decidable set has a null boundary.

Hoyrup and Rojas [47] prove an important technical result:

Fact 2.3.15. Write $\mathbb{X} = \langle X, d, (s_i)_{i \in \mathbb{N}} \rangle$. There exists a sequence $(\epsilon_j)_{j \in \mathbb{N}}$ of uniformly computable positive reals dense in $(0, \infty)$ such that $\{B(s_i, \epsilon_j) : i, j \in \mathbb{N}\}$ is an effective basis whose elements are uniformly almost decidable.

Notation 2.3.16. The choice of effective basis in Fact 2.3.15 is not unique, so we fix one and denote it by $\mathrm{Basis}(\mathbb{X}, \mu)$.

The proof of Fact 2.3.15 uses an effective version of the Baire category theorem [56, 15] which is worth stating separately:

Fact 2.3.17. Every dense Π^0_2 set in a computable metric space $\mathbb{X}$ contains a dense sequence of uniformly computable points. This can be made uniform.

Definition 2.3.18. A *cell* for $\langle \mathbb{X}, \mu \rangle$ is $\varnothing$ or a set of the form

$$C = A_0 \cap \cdots \cap A_n \cap B_0^c \cap \cdots \cap B_m^c \tag{2.3.6}$$

where $A_0, \ldots, A_n, B_0, \ldots, B_m \in \mathrm{Basis}(\mathbb{X}, \mu)$.

Evidently, a cell may be encoded as a string. Rute [67] proves

Fact 2.3.19. The measure of a cell C is computable from (a code for) C.

Because it will be useful later, we add an easy

Lemma 2.3.20. If $\mu(C) > 0$, then we can compute a point in C from C.

Proof. The boundary of each cell is μ-null, so the interior of C has positive measure; in particular, C° is a non-empty open set, which therefore includes an ideal point. Enumerate ideal points s until $s \in A_i$ for all $i \leq n$ and $s \notin \overline{B_i}$ for all $i \leq m$, which can be verified computably and implies $s \in C^\circ$. □

2.3.2 Integration

Let $\langle \mathbb{X}, \mu \rangle$ be a computable probability space. The facts stated below, unless otherwise noted, are proved in Rute [67].

Definition 2.3.21. $L^0(\mathbb{X}, \mu)$ denotes the family of μ-a.e. equivalence classes $[f]_\sim$ of (Borel) measurable functions $f : X \to \mathbb{R}$. $L^0(\mathbb{X}, \mu)$ is a complete, separable metric space under the *Ky-Fan metric*

$$d_\mu([f]_\sim, [g]_\sim) = \int \min\{|f - g|, 1\} d\mu. \tag{2.3.7}$$

Notation 2.3.22. We will identify a measurable function f with its μ-a.e. equivalence class $[f]_\sim$. This should not be too confusing.

Use Fact 2.3.15 to fix an effective basis of uniformly almost decidable sets, which we denote by $\mathrm{Basis}(\mathbb{X}, \mu)$. Recall the definition (2.3.18) of a cell.

Definition 2.3.23. A *test function* for $\langle \mathbb{X}, \mu \rangle$ is a function of the form

$$g = q_0 \mathbf{1}_{C_0} + \cdots + q_k \mathbf{1}_{C_k} \tag{2.3.8}$$

where $q_0, \ldots, q_k \in \mathbb{Q}$ and $C_0, \ldots, C_k$ are pairwise disjoint cells.

Fact 2.3.24. $L^0(\mathbb{X}, \mu)$ is a computable metric space under d_μ with (μ-a.e. equivalence classes of) test functions for ideal points.

Definition 2.3.25. When f is computable as a point in $L^0(\mathbb{X}, \mu)$, we say that f is $L^0(\mu)$*-computable*, or just L^0*-computable*.[2]

We proceed to the spaces of integrable and square-integrable functions. Again, the test functions from Definition 2.3.23 serve as ideal points.

Definition 2.3.26. $L^1(\mathbb{X}, \mu)$ denotes the family of L^1-functions $f : X \to \mathbb{R}$. $L^1(\mathbb{X}, \mu)$ is a separable Banach space under the L^1*-norm*

$$\|f\|_{L^1} = \int |f| d\mu. \tag{2.3.9}$$

Definition 2.3.27. $L^2(\mathbb{X}, \mu)$ denotes the family of L^2-functions $f : X \to \mathbb{R}$. $L^2(\mathbb{X}, \mu)$ is a separable Banach space under the L^2*-norm*

$$\|f\|_{L^2} = \left(\int |f|^2 d\mu \right)^{1/2}. \tag{2.3.10}$$

[2] In the literature, L^0-computability is more often called "effective measurability."

Fact 2.3.28. $L^1(\mathbb{X}, \mu)$ is a computable Banach space under the L^1-norm, and $L^2(\mathbb{X}, \mu)$ is a computable Banach space under the L^2-norm, with test functions (Definition 2.3.23) for ideal points in both cases.

Definition 2.3.29. When f is computable as a point in $L^1(\mathbb{X}, \mu)$, we say that f is $L^1(\mu)$-*computable*. When f is computable as a point in $L^2(\mathbb{X}, \mu)$, we say that f is $L^2(\mu)$-*computable*. As usual, we may omit the measure μ.

Fact 2.3.30. L^2-computability implies L^1-computability, which implies L^0-computability. This can be made uniform.

We can also effectivize the space of measurable sets:

Definition 2.3.31. $\mathscr{B}(\mathbb{X}, \mu)$ denotes the family of (μ-a.e. equivalence classes of Borel) measurable subsets of X. $\mathscr{B}(\mathbb{X}, \mu)$ is a complete, separable metric space under the *symmetric difference metric*

$$d_{\ominus}(A, B) = \mu(A \ominus B), \tag{2.3.11}$$

where of course $\ominus$ denotes the symmetric difference of sets.

Definition 2.3.32. A *test set* for $\langle \mathbb{X}, \mu \rangle$ is a set of the form

$$A = C_0 \cup \cdots \cup C_k \tag{2.3.12}$$

where $C_0, \ldots, C_k$ are pairwise disjoint cells.

Fact 2.3.33. $\mathscr{B}(\mathbb{X}, \mu)$ is a computable metric space under the symmetric difference metric with test sets for ideal points.

Definition 2.3.34. When A is computable as a point in $\mathscr{B}(\mathbb{X}, \mu)$, we say that A is $L^0(\mu)$-*computable*, or just L^0-*computable*.

The next step is to generalize L^0-computability to functions taking values in an arbitrary computable metric space (e.g., the space of càdlàg functions). Let $\mathbb{Y} = \langle Y, d, S_Y \rangle$ be a computable metric space.

Definition 2.3.35. $L^0(\mathbb{X}, \mu; \mathbb{Y})$ denotes the family of measurable functions $f : X \to Y$. $L^0(\mathbb{X}, \mu; \mathbb{Y})$ is complete and separable under the *Ky-Fan metric*

$$d_\mu(f, g) = \int \min\{d(f, g), 1\} d\mu. \tag{2.3.13}$$

Definition 2.3.36. A *test function* for $\langle \mathbb{X}, \mu \rangle$ and $\mathbb{Y}$ has the form

$$g(x) = s_i \quad \text{if } x \in C_i,\ i = 0, \ldots, k, \tag{2.3.14}$$

where $s_0, \ldots, s_k \in S_Y$ and $C_0, \ldots, C_k$ are pairwise disjoint cells that cover X up to a null set, i.e., such that $\mu(C_0 \cup \cdots \cup C_k) = 1$.

Fact 2.3.37. $L^0(\mathbb{X}, \mu; \mathbb{Y})$ is a computable metric space under d_μ with test functions (Definition 2.3.36) for ideal points.

Definition 2.3.38. When f is computable as a point in $L^0(\mathbb{X}, \mu; \mathbb{Y})$, we say that f is $L^0(\mu)$-*computable*, or just L^0-*computable*.

Remark 2.3.39. If $Y = \{0, 1\}$, we can identify $L^0(\mathbb{X}, \mu; \mathbb{Y})$ with the space $\mathscr{B}(\mathbb{X}, \mu)$ of measurable sets, since $A \mapsto \mathbf{1}_A$ is a bijective isometry.

We give an example of a class of L^0-computable functions which is useful in its own right.

Definition 2.3.40. A function $f : X \to Y$ is said to be μ-*a.e. computable*, or just *a.e. computable*, if f is computable on a Π^0_2 set V such that $\mu(V) = 1$.

Remark 2.3.41. Evidently, total computable functions are a.e. computable.

Hoyrup and Rojas [47] prove the next result:

Fact 2.3.42. Every partial computable function defined on a set of measure one can be extended to an a.e. computable function (i.e., to a Π^0_2 domain).

Fact 2.3.43. Every a.e. computable function is L^0-computable (uniformly).

The notion of an image (or "push-forward") measure is essential:

Definition 2.3.44. Let $f : X \to Y$ be measurable. The *image measure* of μ along f is the probability measure $\mu_f = f(\mu)$ on $\mathbb{Y}$ given by

$$\mu_f(A) = \mu(f^{-1}(A)). \tag{2.3.15}$$

Remark 2.3.45. In probability, the image measure of a random variable (i.e., measurable function) f is typically called the *distribution* or *law* of f.

Fact 2.3.46. If f is $L^0(\mu)$-computable, then μ_f is computable (uniformly).

The next result calls to mind the Skorokhod representation theorem from probability theory, which we will effectivize in Part 2.3.3.

Fact 2.3.47. A measure μ on $\mathbb{X}$ is computable if and only if μ is the image of the fair coin measure P along some P-a.e. computable function $\{0, 1\}^{\mathbb{N}} \to X$.

We will often want to compose measurable functions, for which we need

Fact 2.3.48. Let $f : X \to Y$ be $L^0(\mu)$-computable and let $g : Y \to Z$ be $L^0(\mu_f)$-computable (note the measure). Then $g \circ f$ is $L^0(\mu)$-computable uniformly from names for f and g.

There are many computably equivalent ways to choose test functions (that is, ideal points for the L^0, L^1, and L^2 spaces defined above). Rute [67] lists several, including one due to Gács [39]: a set $\mathcal{G}$ of bounded, computable, Lipschitz continuous test functions, defined below.

Because $\mathcal{G}$ does not depend on μ, it is an appropriate choice for effective probability: $\mathbb{X}$ may represent the state space of a stochastic process and take many different measures (e.g., initial distributions), and we would rather use the same test functions in all cases.

Definition 2.3.49. Let $\mathcal{G}_0$ be the set of functions $X \to \mathbb{R}$ of the form

$$g_{u,r,\epsilon}(x) = \max\{0, 1 - \max\{0, d(x,u) - r\}/\epsilon\}. \tag{2.3.16}$$

where $u \in S_Y$, $r \in \mathbb{Q}_{>0}$, and $\epsilon = 1/n$ for some $n \in \mathbb{N}$.

Computably enumerate, as $(g_i)_{i\in\mathbb{N}}$, the smallest set $\mathcal{G}$ of functions $X \to \mathbb{R}$ containing $\mathcal{G}_0$ and the constant function 1 and closed under finite maxima and minima and rational linear combinations.

Remark 2.3.50. The function $g_{u,r,\epsilon}$ is continuous, identically 1 on $B(u,r)$, and identically 0 outside $B(u, r+\epsilon)$, taking intermediate values in between. The function g_i is bounded, computable, and Lipschitz continuous.

Gács [39] proves

Fact 2.3.51. All bounded continuous functions can be obtained as the limit of a pointwise increasing sequence of functions in $\mathcal{G}$.

And we already mentioned

Fact 2.3.52. Let μ be a computable probability measure on $\mathbb{X}$. $\mathcal{G}$ is a computably equivalent set of test functions for $L^0(\mathbb{X}, \mu)$.

We list four other important facts, which are proved in Rute [67]:

Fact 2.3.53. The integral operator $f \to \int f d\mu$ is computable from $L^1(\mathbb{X}, \mu)$ to $\mathbb{R}$. More generally, the integral operator $\langle f, A\rangle \to \int_A f d\mu$ is computable from $L^1(\mathbb{X}, \mu) \times \mathscr{B}(\mathbb{X}, \mu)$ to $\mathbb{R}$. Moreover, the valuation operator $A \to \mu(A)$ is computable of type $\mathscr{B}(\mathbb{X}, \mu) \to \mathbb{R}$.

Fact 2.3.54. $\|\cdot\|_{L^1}$-convergence implies d_μ-convergence, and a convergence rate for the latter is computable from a convergence rate for the former.

Fact 2.3.55. f is L^1-computable if and only if f is L^0-computable and $\|f\|_{L^1} < \infty$ is computable. f is L^2-computable if and only if f is L^0-computable and $\|f\|_{L^2} < \infty$ is computable. If $0 \le f \le 1$, then f is L^2-computable if and only if f is L^0-computable. This can be made uniform.

Fact 2.3.56. If f and g are L^1-computable and $0 \le g \le 1$, then $f \cdot g$ is L^1-computable. This can be made uniform.

Finally, we add

Proposition 2.3.57. If f is L^0-computable, g is L^1-computable, and $|f| \le g$ almost everywhere, then f is L^1-computable. This can be made uniform.

Proof. The following is uniform in names for $f \in L^0(\mathbb{X}, \mu)$ and $g \in L^1(\mathbb{X}, \mu)$.

Split f into its positive and negative parts. Considering each separately, we may assume, without loss of generality, that $f \ge 0$.

Let $f_n = \min\{\max\{f, n\}, n+1\} - n$ for each $n \in \mathbb{N}$, so $f = \sum_{n\in\mathbb{N}} f_n$. f_n is L^0-computable (uniformly from n). Moreover, since $0 \le f_n \le 1$, f_n is L^1-computable (again, uniformly) by Fact 2.3.55.

Similarly, let $g_n = \min\{\max\{g, n\}, n+1\} - n$. Then $\|g - (g_0 + \cdots + g_n)\|_{L^1} \to 0$ by the Lebesgue monotone convergence theorem. g_n is L^1-computable, so given $\epsilon > 0$, we can compute an $n \in \mathbb{N}$ such that $\|g - (g_0 + \cdots + g_n)\|_{L^1} < \epsilon$.

Using the fact that $f_n \le g_n$ almost everywhere, we get

$$\begin{aligned} \|f - (f_0 + \cdots + f_n)\|_{L^1} &= \|f_{n+1} + f_{n+2} + \cdots\|_{L^1} \\ &\le \|g_{n+1} + g_{n+2} + \cdots\|_{L^1} < \epsilon. \end{aligned}$$

Evidently, $f_0 + \cdots + f_n$ is L^1-computable, which completes the proof. □

2.3.3 Convergence in probability

As usual, let $\langle \mathbb{X}, \mu \rangle$ be a computable probability space and let $\mathbb{Y} = \langle Y, d, S_Y \rangle$ be a computable metric space. Let f and $(f_i)_{i\in\mathbb{N}}$ be measurable functions $X \to Y$. The facts below, unless otherwise noted, are proved in Rute [67].

Definition 2.3.58. $(f_i)_{i\in\mathbb{N}}$ is said to *converge almost surely* to f (with respect to μ) if for all $\epsilon > 0$ there exists an $n \in \mathbb{N}$ such that

$$\mu\{x \in X : \sup_{i \ge n}\{d(f_i(x), f(x))\} \ge \epsilon\} < \epsilon. \tag{2.3.17}$$

Definition 2.3.59. $(f_i)_{i\in\mathbb{N}}$ is said to converge *effectively almost surely* to f if $n = n(\epsilon)$ is computable from ϵ in Definition 2.3.58.

Definition 2.3.60. $(f_i)_{i\in\mathbb{N}}$ is said to *converge in probability* to f if for all $\epsilon > 0$ there exists an $n \in \mathbb{N}$ such that

$$(\forall i \ge n) \quad \mu\{x \in X : d(f_i(x), f(x)) \ge \epsilon\} < \epsilon. \tag{2.3.18}$$

Definition 2.3.61. $(f_i)_{i\in\mathbb{N}}$ is said to converge *effectively in probability* to f if $n = n(\epsilon)$ is computable from ϵ in Definition 2.3.60.

Recall the Ky-Fan metric d_μ (Definition 2.3.35).

Fact 2.3.62. Convergence in probability is equivalent to convergence with respect to d_μ, and either convergence rate is computable from the other. Almost-sure convergence implies convergence in d_μ, and a convergence rate for the latter is computable from a convergence rate for the former.

Fact 2.3.63. If $(f_i)_{i\in\mathbb{N}}$ converges in d_μ at a *geometric rate*, meaning that

$$(\forall i \in \mathbb{N}) \quad d_\mu(f_i, f) < 2^{-i}, \tag{2.3.19}$$

then $(f_i)_{i\in\mathbb{N}}$ converges effectively almost surely.

Proposition 2.3.64. f is L^0-computable if and only if $\{f \le a\}$ is L^0-computable uniformly from $a \in \mathbb{Q}$ and $\mu\{|f| \le a\} \to 1$ computably as $a \to \infty$.

Proof. Suppose f is L^0-computable. Clearly, $\mathbf{1}_{(-\infty,a]}$ is $L^0(\mu_f)$-computable (uniformly from a) of type $X \to \{0,1\}$. Hence $\{f \le a\}$ is L^0-computable by Fact 2.3.48. Similarly, $\{|f| \le a\}$ is L^0-computable (uniformly), so $\mu\{|f| \le a\}$ is computable by Fact 2.3.53. Since $\mu\{|f| \le a\}$ converges monotonically to 1 as $a \to \infty$, we can compute a rate of convergence.

Conversely, fix $\epsilon > 0$. For $z \in \mathbb{Z}$, let $a_z = z\epsilon$ and let $f_z = a_z \mathbf{1}_{\{a_z < f \le a_{z+1}\}}$. $\{a_z < f \le a_{z+1}\} = \{f \le a_{z+1}\} \setminus \{f \le a_z\}$ is L^0-computable (uniformly from z), and so is f_z. Hence $g_n = \sum_{|z|\le n} f_z$ is L^0-computable. By hypothesis, we can compute an $n = n(\epsilon)$ such that $\mu\{|f| > a_n\} < \epsilon$. On the complement of that set, we have $|g_n - f| < \epsilon$ by construction. We have shown that $(g_n)_{n\in\mathbb{N}}$ converges to f effectively in probability. By Fact 2.3.62, f is L^0-computable. □

Remark 2.3.65. We can replace $\{f \le a\}$ by $\{f < a\}$ in Proposition 2.3.64.

Fact 2.3.47 called to mind an important result from probability theory, known as the Skorokhod representation theorem[3] [70]:

Fact 2.3.66. Let M be a complete, separable metric space. Let $(P_n)_{n\in\mathbb{N}}$ be a sequence of probability measures on M converging weakly to P_0. There exist M-valued random variables $(Z_n)_{n\in\mathbb{N}}$ on $[0,1]$ (with Lebesgue measure) such that Z_n has distribution P_n for all n and $Z_n \to Z_0$ almost surely.

[3]This has nothing to do with Skorokhod space, the subject of Chapter 3.

As an application of computable measure theory, we prove an effective Skorokhod representation theorem:

Theorem 2.3.67. Let $\mathbb{X}$ be a computable metric space. Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence of uniformly computable probability measures on $\mathbb{X}$ such that $\pi(\mu_n, \mu_0) < 2^{-n}$ for all n. There exist uniformly a.e. computable X-valued random variables $(Z_n)_{n\in\mathbb{N}}$ on $[0,1]$ (with Lebesgue measure) such that Z_n has distribution μ_n for all n and $Z_n \to Z_0$ effectively almost surely.

Remark 2.3.68. Cantor space with the fair coin measure is measure-theoretically identical to the unit interval with Lebesgue measure, as the binary expansion defines a measure-preserving computable surjection from the former to the latter with an a.e. computable inverse. From a computable measure-theoretic standpoint, these spaces are essentially interchangeable.

To prove Theorem 2.3.67, we first need to generalize Fact 2.3.15. Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence of uniformly computable probability measures on a computable metric space $\mathbb{X} = \langle X, d, (s_i)_{i\in\mathbb{N}}\rangle$.

Definition 2.3.69. The elements of a sequence $(A_i)_{i\in\mathbb{N}}$ of measurable sets are said to be *uniformly almost decidable for* $(\mu_n)_{n\in\mathbb{N}}$ if there exist sequences $(U_i)_{i\in\mathbb{N}}$ and $(V_i)_{i\in\mathbb{N}}$ of uniformly Σ^0_1 sets such that for all $i, n \in \mathbb{N}$, A_i is μ_n-almost decidable by $\langle U_i, V_i\rangle$.

Lemma 2.3.70. There exists a sequence $(r_j)_{j\in\mathbb{N}}$ of uniformly computable positive reals dense in $(0,\infty)$ such that $\{B(s_i, r_j) : i, j \in \mathbb{N}\}$ is an effective basis whose elements are uniformly almost decidable for $(\mu_n)_{n\in\mathbb{N}}$.

Proof. This proof is a uniform version of the proof of (what we call) Fact 2.3.15 in Hoyrup and Rojas [47]. For all $n, i, j \in \mathbb{N}$, let

$$U_{n,i,j} = \{r > 0 : \mu_n(\overline{B}(s_i, r)) < \mu_n(B(s_i, r)) + 1/j\}.$$

$U_{n,i,j}$ is Σ^0_1 (uniformly) as a subset of $[0,\infty)$. Note that for all $r > 0$,

$$\mu_n(\overline{B}(s_i, r)) - \mu_n(B(s_i, r)) = \mu_n(S_{i,r})$$

where $S_{i,r}$ denotes the sphere $\overline{B}(s_i, r) \setminus B(s_i, r)$. For $r \neq r'$, $S_{i,r}$ and $S_{i,r'}$ are disjoint. μ_n is finite, so $\mu_n(S_{i,r}) \geq 1/j$ for only finitely many r. In particular, $U_{n,i,j}$ is dense in $[0,\infty)$.

For all $i, j \in \mathbb{N}$, let $V_{i,j} = [0,\infty) \setminus \{d(s_i, s_j)\}$. Then $V_{i,j}$ is dense and Σ^0_1 uniformly from i and j. The intersection

$$A = \bigcap_{n,i,j\in\mathbb{N}} U_{n,i,j} \cap \bigcap_{i,j\in\mathbb{N}} V_{i,j}$$

is Π^0_2 in $[0,\infty)$. By the Baire category theorem, A is dense in $[0,\infty)$. By the effective Baire category theorem (Fact 2.3.17), A contains a sequence (r_k) of uniformly computable reals dense in $[0,\infty)$. For all $n,i,k \in \mathbb{N}$, $B(s_i,r_k)$ is almost decidable with respect to μ_n by $U_{\langle i,k\rangle} = B(s_i,r_k)$ and $V_{\langle i,k\rangle} = X \smallsetminus \overline{B}(s_i,r_k)$, which are Σ^0_1 uniformly from i and k.

Finally, let $U = \bigcup_{k\in\mathbb{N}} B(s_{i_k},q_{i_k})$ be an open subset of X. Uniformly from $k \in \mathbb{N}$ and a name for U, $A_k = \{r \in A : r < q_{i_k}\}$ is computably enumerable as $(t_{k,j})_{j\in\mathbb{N}}$. A is dense, so $B(s_{i_k},q_{i_k}) = \bigcup_{j\in\mathbb{N}} B(s_{i_k},t_{k,j})$ for all $k \in \mathbb{N}$. It follows that $U = \bigcup_{k,j\in\mathbb{N}} B(s_{i_k},t_{k,j})$, and this is uniform in a name for U. We have shown that $\{B(s_i,r_k)\}$ is an effective basis. □

Notation 2.3.71. The choice of "uniform" effective basis in Fact 2.3.15 is not unique, so we fix one and denote it by $\mathrm{Basis}(\mathbb{X},(\mu_n)_{n\in\mathbb{N}})$.

Definition 2.3.72. As in Definition 2.3.18, a *cell* is $\varnothing$ or a set of the form

$$C = A_0 \cap \cdots \cap A_n \cap B_0^c \cap \cdots \cap B_m^c, \tag{2.3.20}$$

where $A_0,\dots,A_n,B_0,\dots,B_m \in \mathrm{Basis}(\mathbb{X},(\mu_n)_{n\in\mathbb{N}})$.

Again, cells may be encoded as strings. The proof of (what we call) Fact 2.3.19 in Rute [67] also proves

Lemma 2.3.73. $\mu_n(C)$ is computable uniformly from n and C.

The proof of Lemma 2.3.20 also proves:

Lemma 2.3.74. If $\mu_n(C) > 0$, we can compute a point in C from n and C.

We also need an effective version of part of the portmanteau theorem.

Lemma 2.3.75. Suppose (μ_n) is uniformly computable and $\pi(\mu_n,\mu_0) < 2^{-n}$ for all $n \in \mathbb{N}$. For all cells C and $n \in \mathbb{N}$, there exists an $m = m(n,C) \in \mathbb{N}$ such that $|\mu_i(C) - \mu_0(C)| < 2^{-n}$ for all $i \geq m$.

Proof. First, suppose $A \in \mathrm{Basis}(\mathbb{X},(\mu_n)_{n\in\mathbb{N}})$ and write $A = B(s_i,r_k)$. A is a continuity set for μ_0, so $\mu_0(B(s_i,r)) \downarrow \mu_0(A)$ as $r \downarrow r_k$. Using this fact and Lemma 2.3.73, uniformly from $\epsilon > 0$ and (a code for) A, we can compute a $j = j(\epsilon,i,k) \in \mathbb{N}$ such that $r_j > r_k$ and

$$\mu_0(B(s_i,r_j)) < \mu_0(B(s_i,r_k)) + \epsilon.$$

Letting $\delta = r_j - r_k > 0$, we have $\mu_0(A^\delta) < \mu_0(A) + \epsilon$.

Similarly, $\mu_0(B(s_i, r')) \uparrow \mu_0(A)$ as $r' \uparrow r_k$. Hence, uniformly from ϵ and A, we can compute a $j_1 \in \mathbb{N}$ such that $r_{j_1} < r_k$ and

$$\mu_0(B(s_i, r_{j_1})) > \mu_0(B(s_i, r_k)) - \epsilon.$$

Letting $\delta_1 = r_k - r_{j_1} > 0$, we have $\mu_0((X \setminus A)^{\delta_1}) < \mu_0(X \setminus A) + \epsilon$.

Generalizing to a finite collection of basic open balls $A_0, \ldots, A_m$, we see that, uniformly from ϵ and a cell C, we can compute a $\delta = \delta(\epsilon, C)$ such that $\mu_0(C^\delta) < \mu_0(C) + \epsilon$ and $\mu_0((X \setminus C)^\delta) < \mu_0(X \setminus C) + \epsilon$. Replacing δ by $\min\{\delta, \epsilon\}$, we have $\delta < \epsilon$ without loss of generality.

Now suppose $\pi(\mu_i, \mu_0) < \delta$. We have

$$\mu_i(C) < \mu_0(C^\delta) + \delta < \mu_0(C) + \epsilon + \delta < \mu_0(C) + 2\epsilon$$

and similarly

$$\mu_i(X \setminus C) < \mu_0(X \setminus C) + 2\epsilon.$$

Hence $|\mu_i(C) - \mu_0(C)| < 2\epsilon$ for all sufficiently large i, as required. □

Now we can prove the effective Skorokhod representation theorem.

Proof of Theorem 2.3.67. We effectivize a classical proof of the Skorokhod representation theorem (Fact 2.3.66) presented by Watanabe and Ikeda [71].

Write $\mathbb{X} = \langle X, d, S \rangle$ and $S = (s_i)_{i \in \mathbb{N}}$. Apply Lemma 2.3.70 to fix an effective basis $\{B(s_i, r_k) : i, k \in \mathbb{N}\}$ whose elements are uniformly almost decidable for $(\mu_n)_{n \in \mathbb{N}}$. $(r_k)_{k \in \mathbb{N}}$ is uniformly computable and dense in $[0, \infty)$, and each open ball $B(s_i, r_k)$ has a μ_n-null boundary for all n.

Compute a subsequence $(r_{m_k})_{k \in \mathbb{N}}$ such that for all k, $2^{-(k+1)} < r_{m_k} < 2^{-k}$. This allows us to compute, from i and k, the cell

$$D_i^k = B\left(s_i, r_{m_k}\right) \setminus \bigcup_{j<i} B\left(s_j, r_{m_k}\right).$$

In turn, we can compute, from $i_1, i_2, \ldots, i_k$, the cell

$$S_{i_1, i_2, \ldots, i_k} = D_{i_1}^1 \cap D_{i_2}^2 \cdots \cap D_{i_k}^k.$$

These cells have the following properties:

- If $i_j \neq i'_j$ for some j, then $S_{i_1, i_2, \ldots, i_k} \cap S_{i'_1, i'_2, \ldots, i'_k} = \varnothing$
- $\bigcup_{i_k \in \mathbb{N}} S_{i_1, i_2, \ldots, i_{k-1}, i_k} = S_{i_1, i_2, \ldots, i_{k-1}}$ and $\bigcup_{i \in \mathbb{N}} S_i = X$
- The diameter of $S_{i_1, i_2, \ldots, i_k}$ is less than 2^{-k}

- $\mu_n(\partial S_{i_1,i_2,\ldots,i_k}) = 0$ for all n

Because μ_n is uniformly computable, $\mu_n(S_{i_1,i_2,\ldots,i_k})$ is computable uniformly from n and $i_1, i_2, \ldots, i_k$ by Lemma 2.3.73.

Next, we compute, from n and $i_1, i_2, \ldots, i_k$, the endpoints of an open subinterval $\Delta^n_{i_1,i_2,\ldots,i_k}$ of $[0,1]$ with the following properties:

- The length of $\Delta^n_{i_1,i_2,\ldots,i_k}$ is $\mu_n(S_{i_1,i_2,\ldots,i_k})$
- If $i_k \neq i'_k$, then $\Delta^n_{i_1,i_2,\ldots,i_k}$ and $\Delta^n_{i_1,i_2,\ldots,i'_k}$ are disjoint
- Whenever $i_k < i'_k$, $\Delta^n_{i_1,i_2,\ldots,i_k}$ lies to the left of $\Delta^n_{i_1,i_2,\ldots,i'_k}$
- The union of $\Delta^n_{i_1,i_2,\ldots,i_{k-1},i_k}$ over all i_k is $\Delta^n_{i_1,i_2,\ldots,i_{k-1}}$, so in particular

$$\Delta^n_{i_1,i_2,\ldots,i_{k-1},i_k} \subseteq \Delta^n_{i_1,i_2,\ldots,i_{k-1}}$$

Fix $n \in \mathbb{N}$. If $\mu_n(S_{i_1,\ldots,i_k}) > 0$, we can apply Lemma 2.3.74 to compute, from $i_1, \ldots, i_k$, a point $x^n_{i_1,\ldots,i_k}$ in the interior of the cell $S_{i_1,\ldots,i_k}$. Otherwise, $x^n_{i_1,\ldots,i_k}$ is undefined. We will now define the a.e. computable random variable Z_n on $[0,1]$. For all $k \in \mathbb{N}$ and $t \in [0,1]$, set

$$Z^k_n(t) = x^n_{i_1,\ldots,i_k} \text{ for } t \in \Delta^n_{i_1,\ldots,i_k}$$

if the computable point $x^n_{i_1,\ldots,i_k}$, exists; otherwise, $Z^k_n(t)$ is undefined on $\Delta^n_{i_1,\ldots,i_k}$. Z^k_n is also undefined at the endpoints of those intervals; the set of endpoints has Lebesgue measure zero.

λ denotes Lebesgue measure. $x^n_{i_1,\ldots,i_k}$ is undefined only if $\mu_n(S_{i_1,\ldots,i_k}) = 0$, in which case $\lambda(\Delta^n_{i_1,i_2,\ldots,i_k}) = 0$. Therefore, Z^k_n is a.e. computable. A countable union of null sets has measure zero, so there exists a set $H_n \subseteq [0,1]$ such that (i) $\lambda(H_n) = 0$ and (ii) for all k, Z^k_n is computable on $[0,1] \setminus H_n$.

It is not difficult to see that our construction guarantees that for all $k' > k \in \mathbb{N}$, for all $t \notin H_n$, $d(Z^k_n(t), Z^{k'}_n(t)) \leq 2^{-k}$. Set $Z_n(t) = \lim_k Z^k_n(t)$. Then Z_n is a.e. computable. Again, a countable union of null sets has measure zero, so there exists a set $H \subseteq [0,1]$ such that (a) $\lambda(H) = 0$ and (b) for all n, Z_n is computable on $[0,1] \setminus H$.

By the classical proof of Fact 2.3.66, the distribution of Z_n is in fact μ_n for all n. We just need to show that $Z_n \to Z_0$ effectively almost everywhere.

Fix $\epsilon > 0$. Compute a k such that $2^{-k} < \epsilon$. Intervals of the form $\Delta^0_{i_1,\ldots,i_k}$ cover $[0,1]$ up to a null set, so compute a finite set $\{c_1, \ldots, c_m\}$ of sequences $c_j = \langle i^j_1, \ldots, i^j_k \rangle$ of length k such that

$$\lambda(\Delta^0_{c_1} \cup \cdots \cup \Delta^0_{c_m}) > 1 - \epsilon/2, \quad \text{i.e., } \lambda(J) < \epsilon/2, \tag{2.3.21}$$

where $J = J(\epsilon)$ denotes the complement of $\Delta^0_{c_1} \cup \cdots \cup \Delta^0_{c_m}$ in $[0,1]$.

There are only finitely many cells $S_{c_1}, \ldots, S_{c_m}$, so Lemma 2.3.75 says we can compute an $N \in \mathbb{N}$ such that for all $n \geq N$,

$$\max_{1 \leq j \leq m} \left| \mu_n(S_{c_j}) - \mu_0(S_{c_j}) \right|$$

is arbitrarily small, which implies $\max_{1 \leq j \leq m} \lambda(\Delta^0_{c_j} \ominus \Delta^n_{c_j})$ can be made small, where $\ominus$ denotes symmetric difference. $\Delta^0_{c_j}$ and $\Delta^n_{c_j}$ are intervals, so we have shown that their endpoints can be made arbitrarily close; the endpoints of $\Delta^0_{c_j}$ are fixed for each j, so we can compute an $N = N(\epsilon)$ such that for all $n \geq N$, for $j = 1, \ldots, m$, the left and right endpoints of $\Delta^n_{c_j}$ are within $\epsilon/8m$ of the left and right endpoints, respectively, of $\Delta^0_{c_j}$.

For $j = 1, \ldots, m$, let F_j and G_j be the real intervals of length $\epsilon/4m$ centered at the left and right endpoints, respectively, of $\Delta^0_{c_j}$. Hence, for all $n \geq N$, the endpoints of $\Delta^n_{c_j}$ belong to $(F_j \cup G_j) \cap [0,1]$, and so the points $t \in \Delta^0_{c_j} \setminus (F_j \cup G_j)$ belong to $\Delta^n_{c_j}$ as well as $\Delta^0_{c_j}$.

Let $K = K(\epsilon) = (F_1 \cup G_1 \cup \cdots \cup F_m \cup G_m) \cap [0,1]$. Then

$$\lambda(K) \leq 2m \cdot \epsilon/4m = \epsilon/2.$$

Finally, $\lambda(J \cup K) < \epsilon/2 + \epsilon/2 = \epsilon$, and for all $t \notin J \cup K$, for all $n \geq N$, our construction guarantees that

$$d(Z_0(t), Z_n(t)) < 2^{-k} < \epsilon,$$

since $Z_0(t), Z_n(t) \in S_{c_j}$ for some $j = 1, \ldots, m$, and the diameter of each cell S_{c_j} is less than 2^{-k}. Therefore, $Z_n \to Z_0$ effectively almost everywhere, which completes the proof of Theorem 2.3.67. $\square$

Chapter 3

Effective Skorokhod space

A function, defined on some subset of the real line and taking values in a metric space, is said to be càdlàg[1] if it is everywhere continuous on the right with limits on the left. Essentially, càdlàg functions are continuous except where they jump, which is countably often at most. In this chapter, we introduce a computable metric space whose points are càdlàg functions on a closed and bounded interval. This space is fundamental to our effective theory of Lévy and Feller processes.

How is the function space relevant? A stochastic process X is typically described as a time-indexed family $\{X_t\}$ of random variables $\omega \mapsto X_t(\omega)$, but it can also be treated as a single random variable $\omega \mapsto (t \mapsto X_t(\omega))$ taking values in some space of functions; the values of this "random function" are the sample paths of the process. In computable analysis, a sample path presentation is more useful, assuming the function space is reasonably well behaved.[2] In fact, the sample paths of Lévy and Feller processes are càdlàg (to be precise, up to a modification), so they live in Skorokhod space.

In Section 3.1, we review the classical theory of càdlàg functions. In Section 3.2, we show that those functions form a computable metric space. In Section 3.3, we establish the computable properties of càdlàg functions that will be needed in the following chapters.

[1] From the French: "continue à droite, limite à gauche." Also known as *rcll*, for "right-continuous, left-limited." Rogers and Williams [65] use the term *R-function*.

[2] As an example of a function space that is *not* well behaved, the Daniell-Kolmogorov extension theorem gives a canonical process with sample paths in the space of *all* functions from, say, $[0, 1]$ to $\mathbb{R}$, but that space (with the product topology) is not even separable. See Theorem 1.3.2 in Revuz and Yor [63] or Theorem II.30.1 in Rogers and Williams [65].

3.1 Skorokhod space

This section is review, based on the well-known Billingsley [7]. We describe a topology on the space of right-continuous, left-limited functions from a closed and bounded interval to a complete, separable metric space.

3.1.1 Càdlàg functions

We begin with a standard bit of notation (and a convention). Fix a complete, separable metric space $\langle E, d\rangle$ and a real number $T > 0$.

Notation 3.1.1. Let $f : [0,T] \to E$.

(i) For $0 < t \le T$, $f(t-)$ denotes the left-hand limit $\lim_{s\uparrow t} f(s)$, if it exists.

(ii) For $0 \le t < T$, $f(t+)$ denotes the right-hand limit $\lim_{s\downarrow t} f(s)$, if it exists.

By convention, $f(0-) = f(0)$ and $f(T+) = f(T)$.

The reader will be quite familiar with spaces of continuous functions.

Definition 3.1.2. $\mathbf{C}_E = \mathbf{C}_E[0,T]$ denotes the set of continuous functions from $[0,T]$ to E. $\mathbf{C}_E$ is complete and separable under the *uniform metric*

$$d_\infty(f,g) = \sup\{d(f(t),g(t)) : 0 \le t \le T\}. \tag{3.1.1}$$

Our interest, again, is in a larger class of functions:

Definition 3.1.3. A function f from $[0,T]$ to E is said to be *càdlàg* if f is everywhere continuous on the right with limits on the left; that is,

(i) for $0 < t \le T$, $f(t-)$ exists, and

(ii) for $0 \le t < T$, $f(t+)$ exists and equals $f(t)$.

The set of càdlàg functions from $[0,T]$ to E is denoted by $\mathbf{D}_E = \mathbf{D}_E[0,T]$.

Notation 3.1.4. When $E = \mathbb{R}$, the most common case, we may omit E from the notation for the set of càdlàg functions and write simply $\mathbf{D} = \mathbf{D}[0,T]$.

Remark 3.1.5. Evidently, continuous functions are càdlàg, and the reverse inclusion does not hold for nontrivial E: $\mathbf{C}_E[0,T] \subsetneq \mathbf{D}_E[0,T]$.

We need some notation to describe the jumps (discontinuities) of a càdlàg function f, and we can be more descriptive when f is vector-valued:

Definition 3.1.6. In general, for $f \in \mathbf{D}_E[0,T]$ and $0 \le t \le T$, we let

$$\Delta f(t) = d(f(t), f(t-)), \tag{3.1.2}$$

which we can call the *size of the jump* at t.

If $E = V$ is a vector space, hence a Banach space, then for $f \in \mathbf{D}_V[0,T]$, we can also define its *jump function* $\mathcal{J}_f : [0,T] \to V$ by

$$\mathcal{J}_f(t) = f(t) - f(t-). \tag{3.1.3}$$

The following auxiliary function is occasionally useful, as it measures how much a given function varies on a given set:

Definition 3.1.7. Let $f : [0,T] \to E$ be any function, not necessarily càdlàg. Let $P \subseteq [0,T]$ be any set. Let

$$\eta_f(P) = \sup\{d(f(s), f(t)) : s, t \in P\}. \tag{3.1.4}$$

For instance, we can use η to define the modulus of continuity:

Definition 3.1.8. Let $f : [0,T] \to E$ be any function. The *modulus of continuity* of f is given by

$$\pi_{\mathbf{C}}(f, \delta) = \sup\{\eta_f[t, t+\delta] : 0 \le t \le T - \delta\}. \tag{3.1.5}$$

Fact 3.1.9. f is continuous if and only if $\lim_{\delta \to 0} \pi_{\mathbf{C}}(f, \delta) = 0$.

Continuous functions on $[0,T]$ are uniformly continuous, of course, and there is a notion of uniformity for càdlàg functions too, as follows.

Definition 3.1.10. A *partition* σ of $[0,T]$ is a string $\langle \sigma_i \rangle_{i \le n}$ of real numbers such that $0 = \sigma_0 < \sigma_1 < \cdots < \sigma_{n-1} < \sigma_n = T$. Let $|\sigma| = n$.

Intervals of the form $[\sigma_i, \sigma_{i+1})$ are called *subintervals* of the partition. The *mesh* of the partition is the length of the longest subinterval:

$$\mathrm{mesh}(\sigma) = \max_{i < |\sigma|} \{\sigma_{i+1} - \sigma_i\}. \tag{3.1.6}$$

That brings us to uniformity for càdlàg functions:

Fact 3.1.11. Let $f \in \mathbf{D}_E$ and $\epsilon > 0$. There exists a partition σ such that

$$\max_{i < |\sigma|} \eta_f[\sigma_i, \sigma_{i+1}) < \epsilon. \tag{3.1.7}$$

Fact 3.1.11 is quite useful. For example, it implies

Fact 3.1.12. Let $f \in \mathbf{D}_E$.

(i) For all $M > 0$, $\{t : \Delta f(t) > M\}$ is finite. Hence $\{t : \Delta f(t) \neq 0\}$ is countable: càdlàg functions jump at most countably often.

(ii) The range of f has compact closure. In particular, the uniform metric d_∞ extends from $\mathbf{C}_E$ to $\mathbf{D}_E$.

(iii) f can be approximated in the uniform metric d_∞ by piecewise constant functions. In particular, f is Borel measurable.

As we know, the modulus of continuity $\pi_{\mathbf{C}}$ characterizes the continuous functions. The càdlàg functions are characterized by a different modulus, as follows; naturally, we call it the "càdlàg modulus."

Definition 3.1.13. A partition σ is said to be δ-*sparse* if $\mathrm{mesh}(\sigma) > \delta$.

Definition 3.1.14. Let $f : [0,T] \to E$ be any function (not necessarily càdlàg). The *càdlàg modulus* of f is given by

$$\pi_{\mathbf{D}}(f,\delta) = \inf\{\max_{i<|\sigma|} \eta_f[\sigma_i, \sigma_{i+1}) : \sigma \text{ is } \delta\text{-sparse}\}. \tag{3.1.8}$$

Remark 3.1.15. If $\delta_0 < \delta_1$, then $\pi_{\mathbf{D}}(f,\delta_0) \leq \pi_{\mathbf{D}}(f,\delta_1)$.

The càdlàg modulus $\pi_{\mathbf{D}}$ does in fact characterize the càdlàg functions:

Fact 3.1.16. A function f is càdlàg if and only if $\lim_{\delta\to 0} \pi_{\mathbf{D}}(f,\delta) = 0$.

We may need to compare the two moduli, $\pi_{\mathbf{C}}$ and $\pi_{\mathbf{D}}$. It is easy to prove

Fact 3.1.17. Let $\delta < T/2$. Then

(i) $\pi_{\mathbf{D}}(f,\delta) \leq \pi_{\mathbf{C}}(f,2\delta)$ for all $f : [0,T] \to E$, and

(ii) $\pi_{\mathbf{C}}(f,\delta) \leq 2\pi_{\mathbf{D}}(f,\delta)$ for all continuous $f : [0,T] \to E$.

Hence the moduli are essentially equivalent for continuous functions.

3.1.2 Skorokhod topology

In the last part, we introduced $\mathbf{D}_E[0,T]$, the set of càdlàg functions from $[0,T]$ to E. Now we give it a metric that induces a separable topology.

We know that the uniform metric d_∞ extends from $\mathbf{C}_E$ to $\mathbf{D}_E$, and it is easy to see that $\mathbf{D}_E$ is complete under d_∞. However, the uniform topology is too large: $\langle \mathbf{D}_E, d_\infty \rangle$ is not separable for nontrivial E, as follows.

Example 3.1.18. Let $x, y \in E$ be distinct points and let $t_0 \geq 0$. Define a càdlàg function $h_{t_0} = h_{t_0}^{x,y}$ by the formula

$$h_{t_0}(t) = \begin{cases} x & \text{if } t < t_0, \\ y & \text{if } t \geq t_0. \end{cases} \tag{3.1.9}$$

If $t_1 < t_2 \leq T$, then $d_\infty(h_{t_1}, h_{t_2}) = d(x, y)$. It follows that no countable subset of $\mathbf{D}_E$ can be dense with respect to d_∞.

Example 3.1.18 will be the basis for several other examples.

We can express uniform convergence in terms of graphs, which motivates the definition of a more suitable topology on $\mathbf{D}_E$.

Definition 3.1.19. Let $f : [0, T] \to E$ be any function. The *graph* of f is the set of ordered pairs

$$\{\langle t, x\rangle : f(t) = x\} \subseteq [0, T] \times E. \tag{3.1.10}$$

In the pair $\langle t, x\rangle$, t is called the *abscissa* and x the *ordinate*.

Essentially, two functions f and g are near one another in the uniform topology if the graph of f can be moved onto the graph of g by a uniformly small change in the ordinates, leaving the abscissas fixed. We can define another topology on $\mathbf{D}_E$, called the Skorokhod topology, in such a way that two functions f and g are near one another if the graph of f can be moved onto the graph of g by a uniformly small change in both the ordinates and the abscissas. Thus two functions can be near one another in the Skorokhod topology even if they do not jump at the same times—interpreting $[0, T]$ as a time interval, obviously. We will now make this precise.

The Skorokhod topology is defined in terms of time changes:

Definition 3.1.20. A *time change* ϕ is a continuous, strictly increasing bijection of $[0, T]$ onto $[0, T]$. The set of time changes is denoted by $\Lambda = \Lambda_T$.

Fact 3.1.21. We list some basic facts about time changes.

(i) Time changes are uniformly continuous, and they fix 0 and T.

(ii) The identity function $\mathrm{id} : [0, T] \to [0, T]$ is a time change.

(iii) For all $\phi, \psi \in \Lambda$, $\phi \circ \psi \in \Lambda$ and $\phi^{-1} \in \Lambda$. In other words, the time changes form a group under composition.

Definition 3.1.22. The *Skorokhod metric* d_J° on $\mathbf{D}_E$ is given by

$$d_J^\circ(f,g) = \inf_{\phi\in\Lambda} \max\left\{\|\phi - \mathrm{id}\|_\infty, d_\infty(f, g\circ\phi)\right\}. \tag{3.1.11}$$

The *Skorokhod topology*, denoted by J_1, is the topology induced by d_J°.

We collect some useful facts about the Skorokhod metric and topology.

Fact 3.1.23. For any time change ϕ and càdlàg functions f and g,

$$\|\phi - \mathrm{id}\|_\infty = \|\mathrm{id} - \phi^{-1}\|_\infty \quad \text{and} \quad d_\infty(f, g\circ\phi) = d_\infty(f\circ\phi^{-1}, g). \tag{3.1.12}$$

Fact 3.1.24. For any time change ϕ and partition σ, $\phi(\sigma) = \langle\phi(\sigma_i)\rangle$ is a partition (of equal length). Moreover,

$$|\mathrm{mesh}(\sigma) - \mathrm{mesh}(\phi(\sigma))| \le 2\|\phi - \mathrm{id}\|_\infty. \tag{3.1.13}$$

Fact 3.1.25. Let f and $(f_n)_{n\in\mathbb{N}}$ be càdlàg. Then $d_J^\circ(f_n, f) \to 0$ if and only if there exists a sequence $(\phi_n)_{n\in\mathbb{N}}$ of time changes such that

$$\|\phi_n - \mathrm{id}\|_\infty \to 0 \quad \text{and} \quad d_\infty(f_n\circ\phi_n, f) \to 0. \tag{3.1.14}$$

Fact 3.1.26. The Skorokhod topology is coarser than the uniform topology: if $d_\infty(f_n, f) \to 0$, then $d_J^\circ(f_n, f) \to 0$, but the converse does not hold.

Example 3.1.27. Recall Example 3.1.18. Let $x \ne y \in E$ and $t_0 < T$. Let

$$f = h_{t_0}^{x,y} \quad \text{and} \quad f_n = h_{t_0+1/n}^{x,y} \quad \text{for all } n \in \mathbb{N}. \tag{3.1.15}$$

Evidently, $d_J^\circ(f_n, f) \to 0$, but $d_\infty(f_n, f) = d(x,y)$ for all n.

Fact 3.1.28. If $d_J^\circ(f_n, f) \to 0$, then $d(f_n(t), f(t)) \to 0$ for all points of continuity t for f. If moreover f is continuous, then $d_\infty(f_n, f) \to 0$. The Skorokhod topology relativized to $\mathbf{C}_E$ coincides with the uniform topology.

Finally, we identify a countable dense subset of $\mathbf{D}_E$, which will be used extensively when we effectivize the Skorokhod topology.

Definition 3.1.29. Recall that E is separable. Fix a countable dense subset S of E. For all $n \in \mathbb{N}$, let $\mathscr{S}_n$ be the set of S-valued càdlàg step functions on $[0,T]$ that are constant on each of the intervals $[i2^{-n}T, (i+1)2^{-n}T)$ for $i = 0, \ldots, 2^n - 1$. The elements of $\mathscr{S} = \bigcup_{n\in\mathbb{N}} \mathscr{S}_n$ are called *simple* functions.

Fact 3.1.30. A càdlàg function f can be approximated in d_J° by simple functions. Hence the Skorokhod topology is separable: $\mathscr{S}$ is dense.

3.1.3 Billingsley metric

Both the Skorokhod metric d_J° and the topology J_1 it induces have unusual and inconvenient features. For instance, under J_1, if V is a (nontrivial) Banach space, then $\mathbf{D}_V$ is neither complete nor a topological group under pointwise addition, as the following example shows.

Example 3.1.31. Recall Example 3.1.18. Let $v_1 \neq v_2 \in V$ and $t_0 < T$. Let

$$f = h_{t_0}^{v_1,v_2} \quad \text{and} \quad f_n = h_{t_0+1/n}^{v_1,v_2} \quad \text{for all } n \in \mathbb{N}. \tag{3.1.16}$$

We know that $d_J^\circ(f_n, f) \to 0$. But consider the sequence of differences

$$(f_n - f)(s) = \begin{cases} 0 & \text{if } s < t_0, \\ v_1 - v_2 & \text{if } t_0 \le s < t_0 + 1/n, \\ 0 & \text{if } t_0 + 1/n \le s. \end{cases} \tag{3.1.17}$$

Although it is a Cauchy sequence with respect to d_J°, it does not converge.

A similar example shows that $\mathbf{D}_E$ is not complete under d_J° for any (nontrivial) complete, separable metric space E.

On the other hand, completeness is not a topological property, and there is a topologically equivalent metric d_J, which we call the Billingsley metric, under which $\mathbf{D}_E$ is complete. The definition of this new metric hinges on a different notion of what it might mean for a change in the abscissas to be "uniformly small," as follows.

Definition 3.1.32. The *deviation* $\|\phi\|^\circ$ of a time change ϕ is given by

$$\|\phi\|^\circ = \sup\left\{\left|\ln \frac{\phi t - \phi s}{t - s}\right| : 0 \le s < t \le T\right\}. \tag{3.1.18}$$

Remark 3.1.33. Since $(\phi t - \phi s)/(t - s) > 0$ is the slope of a chord of the graph of ϕ, the deviation measures the most those slopes deviate from 1.

Fact 3.1.34. For all $\phi, \phi_1, \phi_2 \in \Lambda$,

(i) $\|\phi^{-1}\|^\circ = \|\phi\|^\circ \ge 0 = \|\mathrm{id}\|^\circ$, and

(ii) $\|\phi_1 \circ \phi_2\|^\circ \le \|\phi_1\|^\circ + \|\phi_2\|^\circ$.

Definition 3.1.35. The *Billingsley metric* d_J on $\mathbf{D}_E$ is given by

$$d_J(f, g) = \inf_{\phi \in \Lambda} \max\{\|\phi\|^\circ, d_\infty(f, g \circ \phi)\}. \tag{3.1.19}$$

Remark 3.1.36. Compare Definition 3.1.22. The only difference is that where the Skorokhod metric uses the sup-norm to measure how much a time change differs from identity, the Billingsley metric uses the deviation.

The analogue of Fact 3.1.25 for the Billingsley metric is

Fact 3.1.37. Let f and $(f_n)_{n\in\mathbb{N}}$ be càdlàg. Then $d_J(f_n, f) \to 0$ if and only if there exists a sequence $(\phi_n)_{n\in\mathbb{N}}$ of time changes such that

$$\|\phi_n\|^\circ \to 0 \quad \text{and} \quad d_\infty(f_n \circ \phi_n, f) \to 0. \tag{3.1.20}$$

We need to be able to compare the Skorokhod and Billingsley metrics. It is not difficult to prove

Fact 3.1.38. For all $\phi \in \Lambda$, $\|\phi - \mathrm{id}\|_\infty \le T(e^{\|\phi\|^\circ} - 1)$.

Fact 3.1.38 and the elementary inequality $u \le e^u - 1$ for all $u \in \mathbb{R}$ imply

Fact 3.1.39. For all $f, g \in \mathbf{D}_E$, $d_J^\circ(f, g) \le \max\{1, T\}(e^{d_J(f,g)} - 1)$.

Then Fact 3.1.39 implies that if $d_J(f_n, f) \to 0$, then $d_J^\circ(f_n, f) \to 0$. The converse follows from

Fact 3.1.40. For all $f, g \in \mathbf{D}_E$, if $d_J^\circ(f, g) < \delta^2$ and $\delta < \min\{T, 1/4\}$, then

$$d_J(f, g) \le 4\delta + \pi_{\mathbf{D}}(f, \delta), \tag{3.1.21}$$

where $\pi_{\mathbf{D}}$ is the càdlàg modulus (see Definition 3.1.14 and Fact 3.1.16).

Facts 3.1.39 and 3.1.40 together imply

Fact 3.1.41. d_J and d_J° are topologically equivalent.

Incidentally, Fact 3.1.38 and the elementary inequality $e^u - 1 \le 2u$ for all $0 \le u \le 1/2$ give us a convenient estimate:

Fact 3.1.42. If $\|\phi\|^\circ \le 1/2$, then $\|\phi - \mathrm{id}\|_\infty \le 2T\|\phi\|^\circ$.

It remains only to show that the Billingsley metric d_J is complete. First, we associate to each càdlàg function f and partition σ an auxiliary càdlàg function $A_\sigma f$ which is useful in certain proofs.

Definition 3.1.43. For a given partition σ, the operator $A_\sigma : \mathbf{D}_E \to \mathbf{D}_E$ is defined as follows: $A_\sigma f$ takes the constant value $f(\sigma_i)$ on the interval $[\sigma_i, \sigma_{i+1})$ for $i < |\sigma|$, and $A_\sigma f(T) = f(T)$.

Fact 3.1.44. If $\mathrm{mesh}(\sigma) \leq \delta$, then $d_J^\circ(f, A_\sigma f) \leq \max\{\delta, \pi_{\mathbf{D}}(f, \delta)\}$.

We already know that the Skorokhod topology is separable. Using Fact 3.1.44, we can show that $\mathbf{D}_E$ is complete under the new metric. To sum up:

Fact 3.1.45. The space $\mathbf{D}_E$ is separable and complete with respect to the Billingsley metric d_J. The simple functions $\mathscr{S}$ are dense in $\mathbf{D}_E$.

Definition 3.1.46. The space $\langle \mathbf{D}_E, d_J \rangle$ is called *Skorokhod space*.

3.2 Effective Skorokhod space

Previously, we defined Skorokhod space $\langle \mathbf{D}_E, d_J \rangle$, where $\mathbf{D}_E = \mathbf{D}_E[0,T]$ is the set of càdlàg functions from $[0,T]$ to a complete, separable metric space E, and d_J is the Billingsley metric, which makes $\mathbf{D}_E$ complete and separable. Now we turn Skorokhod space into a computable metric space by identifying a suitable set of ideal points.

3.2.1 Simple functions

Fix a computable metric space $\langle E, d, S \rangle$ and $T > 0$. All the results in this section and the next are uniform from a name for T, although it is generally safe to assume that T is a whole number, or at the very least rational. In any case, we begin by identifying a set of ideal points for $\langle \mathbf{D}_E, d_J \rangle$.

We were given a countable dense subset S of E. Recall Definition 3.1.29. For all $n \in \mathbb{N}$, $\mathscr{S}_n$ denotes the set of S-valued càdlàg step functions on $[0,T]$ that are constant on each of the intervals $[i2^{-n}T, (i+1)2^{-n}T)$ for $i = 0, \ldots, 2^n - 1$. The elements of $\mathscr{S} = \bigcup_{n\in\mathbb{N}} \mathscr{S}_n$ are called *simple* functions, and $\mathscr{S}$ is dense in $\mathbf{D}_E$. The simple functions will be our ideal points.

Remark 3.2.1. A simple function can be coded as a string in S.

In order to use the simple functions as ideal points for $\mathbf{D}_E$, we must be able to compute $d_J(f,g)$ uniformly from codes for simple functions f, g. Because simple functions have a special form, we will be able to prove, in Proposition 3.2.7 below, that we can compute, from codes for f and g, a code for a time change ψ which is optimal in the sense that the infimum

$$d_J(f,g) = \inf_{\phi\in\Lambda} \max\{\|\phi\|^\circ, d_\infty(f, g\circ\phi)\}$$

is actually attained for $\phi = \psi$. Essentially, this is because ψ itself must have a special form. We will now make this precise.

Notation 3.2.2. For all $n \in \mathbb{N}$, for $i = 0, \ldots, 2^n$, let $r_{n,i} = i2^{-n}T$. In this context, we will call a real number of the form $r_{n,i}$ a *breakpoint*.

Definition 3.2.3. For all $n \in \mathbb{N}$, Γ_n denotes the set of time changes ψ for which there exists a partition $\sigma = \langle r_{n,i(k)} \rangle$ made up of breakpoints such that

(i) for all $k \leq |\sigma|$, there exists a j such that $\psi(r_{n,i(k)}) = r_{n,j}$, and

(ii) for all $k < |\sigma|$, ψ is linear on the interval $(r_{n,i(k)}, r_{n,i(k+1)})$.

Evidently, Γ_n is finite. Note also that $\Gamma_n \subseteq \Gamma_m$ for all $m > n$. The elements of $\Gamma = \bigcup_{n \in \mathbb{N}} \Gamma_n$ are called *basic time changes*.

Remark 3.2.4. A basic time change can be coded as a string of rationals. It is also evidently computable as a function from $[0,T]$ to $[0,T]$.

It is not difficult to prove the following two lemmas. The first belongs to computability theory; the second, to elementary geometry.

Lemma 3.2.5. Given codes for $f, g \in \mathscr{S}$ and $\psi \in \Gamma$, we can compute the rationals $d_\infty(f, g \circ \psi)$ and $\|\psi - \mathrm{id}\|_\infty$, the real number $\|\psi\|^\circ$, and the rationals $\min_{s<t}\{(\psi t - \psi s)/(t-s)\}$ and $\max_{s<t}\{(\psi t - \psi s)/(t-s)\}$.

Lemma 3.2.6. Let $\phi_1, \phi_2 \in \Lambda$ and $0 \leq s < t \leq T$. If $\phi_1 = \phi_2$ outside (s,t), ϕ_1 is linear on (s,t), and ϕ_2 is *not* linear on that interval, then

$$\|\phi_1\|^\circ \leq \|\phi_2\|^\circ \quad \text{and} \quad \|\phi_1 - \mathrm{id}\|_\infty \leq \|\phi_2 - \mathrm{id}\|_\infty. \tag{3.2.1}$$

With Lemmas 3.2.5 and 3.2.6, we can prove

Proposition 3.2.7. Given codes for $f, g \in \mathscr{S}$, we can compute (i) the code of a basic time change ψ_0 such that

$$d_J(f,g) = \max\{\|\psi_0\|^\circ, d_\infty(f, g \circ \psi_0)\}, \tag{3.2.2}$$

so that ψ_0 may be considered optimal with respect to d_J, and (ii) the code of a basic time change ψ_1 such that

$$d_J^\circ(f,g) = \max\{\|\psi_1 - \mathrm{id}\|_\infty, d_\infty(f, g \circ \psi_1)\}, \tag{3.2.3}$$

so that ψ_1 may be considered optimal with respect to d_J°.

Proof. We will prove claim (i), constructing a time change ψ that is optimal with respect to d_J. The proof of claim (ii) is exactly analogous.

Given codes for $f, g \in \mathscr{S}$, compute an $n \in \mathbb{N}$ such that $f, g \in \mathscr{S}_n$. We claim that there is no time change $\phi \notin \Gamma_n$ such that for all $\psi \in \Gamma_n$,

$$\max\{\|\phi\|^\circ, d_\infty(f, g \circ \phi)\} < \max\{\|\psi\|^\circ, d_\infty(f, g \circ \psi)\}.$$

If this claim is true, then the result follows from Lemma 3.2.5: we compute the code of each element ψ of Γ_n, a finite set, then compute $\|\psi\|^\circ \in \mathbb{R}$ and $d_\infty(f, g \circ \psi) \in \mathbb{Q}$, and on that basis choose an optimal ψ.

Suppose the claim is false, i.e., that such a time change ϕ exists. Let $a_{n,i} = \phi^{-1}(r_{n,i})$ for $i = 0, \ldots, 2^n$. Because f is constant on each interval of the form $(r_{n,i}, r_{n,i+1})$, the behavior of ϕ on the interval $(a_{n,i}, a_{n,i+1})$ cannot affect the value of $d_\infty(f, g \circ \phi)$. By Lemma 3.2.6, we may assume that ϕ is linear on $(a_{n,i}, a_{n,i+1})$ for all i, as this can only improve $\|\phi\|^\circ$.

Next, note that $\|\phi^{-1}\|^\circ = \|\phi\|^\circ$ and $d_\infty(f, g \circ \phi) = d_\infty(f \circ \phi^{-1}, g)$. Let $b_{n,i} = \phi(r_{n,i})$ for all i. Because z is constant on each interval of the form $(r_{n,i}, r_{n,i+1})$, the behavior of ϕ^{-1} on the interval $(b_{n,i}, b_{n,i+1})$ cannot affect the value of $d_\infty(f \circ \phi^{-1}, g)$. By Lemma 3.2.6 again, we may assume that ϕ^{-1} is linear on $(b_{n,i}, b_{n,i+1})$ for all i. That is, ϕ is linear on intervals of the form $(r_{n,i}, r_{n,i+1})$. Note that this change cannot undo our progress by somehow making ϕ nonlinear on intervals of the form $(a_{n,i}, a_{n,i+1})$.

In short, ϕ, or an improvement on ϕ, is linear on intervals of the form

$$(\phi^{-1}(r_{n,i}), \phi^{-1}(r_{n,i+1})) \quad \text{and} \quad (r_{n,i}, r_{n,i+1}) \quad \text{for all } i.$$

Therefore, if the graph of ϕ is nonlinear in a neighborhood of some point $\langle t, \phi t\rangle$, both t and ϕt must be breakpoints. In other words, ϕ maps some of the breakpoints to breakpoints and is linear in between; that is, ϕ belongs to Γ_n after all. This completes the proof. □

Together, Lemmas 3.2.5 and 3.2.7 imply

Proposition 3.2.8. The real numbers $d_J(f, g)$ and $d_J^\circ(f, g)$ are computable uniformly from codes for simple functions f and g.

$\mathscr{S}$ admits a canonical enumeration (given that the ideal points S of E have already been enumerated). Hence Proposition 3.2.8 implies that $\mathbf{D}_E$ is a computable metric space under d_J with $\mathscr{S}$ for ideal points.

Notation 3.2.9. From now on, by *Skorokhod space* (Definition 3.1.46) we mean the computable metric space $\langle \mathbf{D}_E, d_J, \mathscr{S}\rangle$, ideal points included.

The next two definitions are special cases of Definitions 2.2.13 and 2.2.14.

Definition 3.2.10. A *Cauchy name*, or just a *name*, for a càdlàg function f is a sequence $(f_n)_{n\in\mathbb{N}}$ of simple functions such that

(i) $d_J(f_n, f) \to 0$, and

(ii) $d_J(f_n, f_m) < 2^{-n}$ for all $m > n$.

Definition 3.2.11. A càdlàg function f is said to be *computably càdlàg* if it has a computable Cauchy name. This definition can be made uniform.

3.2.2 Time changes and partitions

If $(f_n)_{n\in\mathbb{N}}$ is a name for a càdlàg function f, then there must be a sequence $(\phi_n)_{n\in\mathbb{N}}$ of time changes such that $\|\phi_n\|^\circ \to 0$ and $d_\infty(f_n \circ \phi_n, f) \to 0$. Here we consider the computability of $(\phi_n)_{n\in\mathbb{N}}$ and of other useful time changes. For clarity, we state

Definition 3.2.12. Quite simply, a time change ϕ is said to be *computable* if ϕ is computable as a function from $[0,T]$ to $[0,T]$.

Proposition 3.2.13. Uniformly from a name $(f_n)_{n\in\mathbb{N}}$ for $f \in \mathbf{D}_E$, we can compute a sequence $(\phi_n)_{n\in\mathbb{N}}$ of time changes such that for all $n \in \mathbb{N}$,

$$\|\phi_n\|^\circ < 2^{-n+1} \quad \text{and} \quad d_\infty(f_n \circ \phi_n, f) < 2^{-n+1}.$$

Proof. The following is uniform from $(f_n)_{n\in\mathbb{N}}$.

Since $d_J(f_n, f_m) < 2^{-n}$ for all $m > n$, Lemma 3.2.7 says that, given n, we can compute (the code of) a basic time change ψ_n such that $\|\psi_n\|^\circ < 2^{-n}$ and $d_\infty(f_n, f_{n+1} \circ \psi_n) < 2^{-n}$.

For all $n, m \in \mathbb{N}$, compute the basic time change

$$\rho_n^m = \psi_{n+m} \circ \psi_{n+m-1} \circ \cdots \circ \psi_{n+1} \circ \psi_n.$$

Note that $\rho_n^{m+1} = \psi_{n+m+1} \circ \rho_n^m$. It follows from this and Fact 3.1.42 that

$$\|\rho_n^{m+1} - \rho_n^m\|_\infty = \|\psi_{n+m+1} - \mathrm{id}\|_\infty \le 2T\|\psi_{n+m+1}\|^\circ < 2^{-(n+m)}T.$$

Hence, for all n, $(\rho_n^m)_{m\in\mathbb{N}}$ yields a computable Cauchy name in $\mathbf{C}[0,T]$. Its limit τ_n is computable on $[0,T]$, uniformly from n, and given by

$$\tau_n = \lim_m (\psi_{n+m} \circ \psi_{n+m-1} \circ \cdots \circ \psi_{n+1} \circ \psi_n).$$

Evidently, τ_n is non-decreasing and fixes 0 and T, so if we can just show that τ_n is strictly increasing, then τ_n must be a time change. Fact 3.1.34 (ii) implies that $\|\tau_n\|^\circ < 2^{-n+1}$. Since, in particular, $\|\tau_n\|^\circ < \infty$, τ_n must be strictly increasing, which makes it a time change.

Time changes are continuous, so

$$\tau_{n+1} \circ \psi_n = \lim_m (\psi_{n+m} \circ \psi_{n+m-1} \circ \cdots \circ \psi_{n+1} \circ \psi_n) = \tau_n.$$

Equivalently, $\tau_{n+1}^{-1} = \psi_n \circ \tau_n^{-1}$. Letting $\phi_n = \tau_n^{-1}$ for all n, we have

$$d_\infty(f_n \circ \phi_n, f_{n+1} \circ \phi_{n+1}) = d_\infty(f_n, f_{n+1} \circ \psi_n) < 2^{-n}.$$

So $(f_n \circ \phi_n)_{n\in\mathbb{N}}$ is Cauchy in $\mathbf{D}_E$ with respect to d_∞. $\mathbf{D}_E$ is complete in the uniform topology, so $d_\infty(f_n \circ \phi_n, g) \to 0$ for some $g \in \mathbf{D}_E$. Since $\|\phi_n\|^\circ \to 0$ as well, $d_J(f_n, g) \to 0$ (Fact 3.1.37). Limits are unique, so $g = f$.

In fact, we also found a rate of convergence: $d_\infty(f_n \circ \phi_n, f) < 2^{-n+1}$. Finally, ϕ_n is computable uniformly from n, and $\|\phi_n\|^\circ = \|\tau_n\|^\circ < 2^{-n+1}$, which completes the proof. □

In light of Proposition 3.2.13, without loss of generality, whenever we compute from a name $(f_n)_{n\in\mathbb{N}}$ for some $f \in \mathbf{D}_E$, we also have a sequence $(\phi_n)_{n\in\mathbb{N}}$ of time changes such that $\|\phi_n\|^\circ < 2^{-n}$ and $d_\infty(f_n \circ \phi_n, f) < 2^{-n}$. This fact is often useful. For example, we can use it to effectivize Fact 3.1.11 quite easily, as follows. Recall the auxiliary function η_f (Definition 3.1.7), which we used to define the càdlàg modulus $\pi_{\mathbf{D}}$ (Definition 3.1.14).

Proposition 3.2.14. Given $n \in \mathbb{N}$, uniformly from a name for $f \in \mathbf{D}_E$, we can compute a partition σ and rationals r_i such that for all $i < |\sigma|$, for all $t \in [\sigma_i, \sigma_{i+1})$, we have $d(f(t), r_i) < 2^{-n-1}$. In particular, $\eta_f[\sigma_i, \sigma_{i+1}) < 2^{-n}$.

Naturally, to "compute a partition" $\sigma = \langle t_0, \ldots, t_k \rangle$ is simply to compute $k = |\sigma|$ and $t_0, \ldots, t_k \in \mathbb{R}$.

Proof. Using Proposition 3.2.13, compute a time change ϕ and (the code of) a simple function g such that $d_\infty(f, g \circ \phi) < 2^{-n-1}$. g is constant on intervals of the form $[i2^{-k}T, (i+1)2^{-k}T)$ for some k computable from g. For $i = 0, \ldots, 2^k$, let $s_i = i2^{-k}T$ and compute the real number $t_i = \phi^{-1}(s_i)$ and the rational number $r_i = g(q_i)$. For $i = 0, \ldots, 2^k - 1$ and $t \in [t_i, t_{i+1})$,

$$d(f(t), r_i) \le d_\infty(f, g \circ \phi) < 2^{-n-1}.$$

Let $\sigma = \langle t_0, \ldots, t_{2^k} \rangle$. We have $\eta_f[t_i, t_{i+1}) < 2^{-n}$ by the triangle inequality. □

Proposition 3.2.14, in turn, lets us effectivize Fact 3.1.16:

Corollary 3.2.15. Given $n \in \mathbb{N}$, uniformly from a name for $f \in \mathbf{D}_E$, we can compute a positive rational δ and a δ-sparse partition σ such that $\eta_f[\sigma_i, \sigma_{i+1}) < 2^{-n}$ for all $i < |\sigma|$. In particular, $\pi_{\mathbf{D}}(f, \delta) < 2^{-n}$.

Proof. Compute σ as in Proposition 3.2.14 and choose any $\delta < \mathrm{mesh}(\sigma)$. □

3.3 Computably càdlàg functions

Computable functions (between computable metric spaces) are necessarily continuous, so computably càdlàg functions that are not continuous cannot be computable as functions. In this section, we investigate what we can and cannot compute from a name for a càdlàg function.

3.3.1 Partial computability

Here we establish basic computability properties of Skorokhod space $\mathbf{D}_E = \mathbf{D}_E[0,T]$, which will be used extensively in the sequel. The key result is

Theorem 3.3.1. If t is a point of continuity for a càdlàg function f, i.e., if $\Delta f(t) = 0$, then $f(t)$ is computable uniformly from names for f and t. Also, $f(T)$ is computable uniformly from a name for f.

Proof. First of all, $f(T)$ is computable uniformly from a name for f because time changes fix T, so $d(f(T), g(T)) \le d_J(f, g)$ for all càdlàg g.

As for the other claim, let t be a point of continuity for f. The following is uniform from a name $(f_n)_{n\in\mathbb{N}}$ for f and a name for t. Let $\epsilon < 1/2$ be a positive rational. We will compute an ideal point $a \in S$ such that $d(a, f(t)) < 5\epsilon$.

First, we compute a sequence $(g_k)_{k\in\mathbb{N}}$ of simple functions. For all k, let $\delta_k = 2^{-k}\epsilon$. Then use Proposition 3.2.13 to compute a simple function g_k and a time change ϕ_k such that $\|\phi_k\|^\circ < \delta_k/T$ and $d_\infty(f, g_k \circ \phi_k) < \delta_k$.

By Fact 3.1.42, $\|\phi_k - \mathrm{id}\|_\infty < 2\delta_k$. Let $s_k = \phi_k(t)$, so

$$d(f(t), g_k(s_k)) < \delta_k \quad \text{and} \quad |s_k - t| < 2\delta_k.$$

Compute a rational t_k such that $|t - t_k| < \delta_k$. Then $|s_k - t_k| < 3\delta_k$. Let J_k be the rational interval $(t_k - 3\delta_k, t_k + 3\delta_k) \cap [0, T]$, so that $s_k \in J_k$.

Next, compute $\eta_{g_k}(J_k)$ (Definition 3.1.7). Suppose $\eta_{g_k}(J_k) < 4\epsilon$. Then $f(t)$ is within δ_k of $g_k(s_k)$, s_k belongs to J_k, g_k varies by less than 4ϵ on J_k, and $\delta_k \le \epsilon$. Hence $f(t)$ is within 5ϵ of $a = g_k(s_k)$.

On the other hand, if $\eta_{g_k}(J_k) \geq 4\epsilon$, we can increment k and try again with $\delta_{k+1} = \delta_k/2$. It suffices that this search terminate, i.e., that for large enough k, $\eta_{g_k}(J_k) < 4\epsilon$. f is continuous at t, so for sufficiently large k,

$$(\forall 0 \leq r \leq T) \quad |t-r| < 6\delta_k \implies d(f(t), f(r)) < \epsilon.$$

Let $u \in J_k$ and let $r = \phi_k^{-1}(u)$. It follows that $d(f(r), g_k(u)) < \delta_k$ and $|u-r| < 2\delta_k$. We also have $|t-t_k| < \delta_k$ and $|t_k-u| < 3\delta_k$, so $|t-r| < 6\delta_k$. Hence $d(f(t), f(r)) < \epsilon$. We have

$$d(f(t), g_k(u)) \leq d(f(t), f(r)) + d(f(r), g(u)) < \epsilon + \delta_k \leq 2\epsilon$$

for all $u \in J_k$. $f(t)$ is fixed, so $\eta_{g_k}(J_k) < 2 \cdot 2\epsilon = 4\epsilon$, as required. □

Corollary 3.3.2. Let $f : [0,T] \to E$ be continuous. Then f is computably càdlàg if and only if f is computable.

Using Fact 2.2.84 (generalized Fact 2.2.49), we can strengthen Corollary 3.3.2 for functions taking values in a computable Banach space V:

Corollary 3.3.3. Let $f : [0,T] \to V$ be continuous. Then a name for f in $\mathbf{D}_V[0,T]$ is computable from a name for f in $\mathbf{C}_V[0,T]$, and vice versa.

The set of jump times is countable, so Theorem 3.3.1 also implies

Corollary 3.3.4. Uniformly from a name for f in $\mathbf{D}_E[0,T]$, the function $f : [0,T] \to E$ is a.e. computable (with respect to Lebesgue measure λ). Hence the inclusion mapping $\mathbf{D}_E[0,T] \to L^0([0,T], \lambda; E)$ is computable.

Moreover, the proof of Theorem 3.3.1 may be adapted to prove

Proposition 3.3.5. The set of points of continuity for a càdlàg function f is Π^0_2 (i.e., the set of jump times is Σ^0_2) uniformly from a name for f.

Proof. The following is uniform in a name for $f \in \mathbf{D}_E$.

Suppose we have a name for t. As in the proof of Theorem 3.3.1, given a positive rational $\epsilon < 1/2$, for all $k \in \mathbb{N}$, we can compute the following:

- $\delta_k = 2^{-k}\epsilon$,
- a simple function g_k and a time change ϕ_k such that $d_\infty(f, g_k \circ \phi_k) < \delta_k$ and $\|\phi_k\|^\circ < \delta_k/T$, which in turn implies $\|\phi_k - \mathrm{id}\|_\infty < 2\delta_k$,
- $s_k = \phi_k(t)$,

- a rational t_k such that $|t - t_k| < \delta_k$,
- the rational interval $J_k = (t_k - 3\delta_k, t_k + 3\delta_k) \cap [0, T]$ containing s_k, and
- the real number $\eta_{g_k}(J_k)$ (Definition 3.1.7).

Recall that if $\Delta f(t) = 0$, then $\eta_{g_k}(J_k) < \epsilon$ for sufficiently large k. Let

$$C = \{t \in [0,T] : (\forall \epsilon)\ (\exists k)\ (\eta_{g_k}(J_k) < \epsilon)\}.$$

Evidently, C is Π^0_2. We need to show that if $\Delta f(t) > 0$, then $t \notin C$.

Suppose $\Delta f(t) > 0$ and fix $\epsilon < \Delta f(t)/3$. For each k, compute the list of objects mentioned above, then let $I_k = (t_k - \delta_k, t_k + \delta_k) \cap [0, T]$. Because $\|\phi_k - \mathrm{id}\|_\infty < 2\delta_k$, $\phi_k^{-1}(I_k) \subseteq J_k$. It follows that

$$\eta_{g_k}(J_k) \geq \eta_{g_k}(\phi^{-1}(I_k)) > \eta_f(I_k) - 2\delta_k.$$

Moreover, since $|t - t_k| < \delta_k$, I_k includes t as well as all points $s < t$ that are sufficiently close to t. Hence $\eta_f(I_k) \geq d(f(t), f(t-)) = \Delta f(t)$.

Since $\Delta f(t) > 3\epsilon$ and $\delta_k \leq \epsilon$, $\eta_{g_k}(J_k) > 3\epsilon - 2\delta_k \geq \epsilon$. This holds for all k, so $t \notin C$, which completes the proof. □

There is an alternative characterization of computably càdlàg functions that will prove quite useful:

Theorem 3.3.6. Uniformly from a name for $f \in \mathbf{D}_E[0,T]$, we can compute

(i) the endpoint $f(T)$,

(ii) a sequence $(\sigma^n)_{n\in\mathbb{N}}$ of partitions such that for all n,

$$\eta_f[\sigma^n_k, \sigma^n_{k+1}) < 2^{-n} \quad \text{for all } k < |\sigma^n|, \text{ and} \tag{3.3.1}$$

(iii) a sequence $(t_j)_{j\in\mathbb{N}}$ of reals dense in $[0,T]$ such that $f(t_j)$ is computable uniformly from j.

Conversely, from names for (i), (ii), and (iii), we can compute f.

Proof. First, suppose we have a name for f. Theorem 3.3.1 gives us (i) and Proposition 3.2.14 gives us (ii). Proposition 3.3.5 states that the set C of points of continuity for f is a dense Π^0_2 set. By Fact 2.3.17 (computable Baire category theorem), C contains a dense sequence of uniformly computable points; this and Theorem 3.3.1 give us (iii).

Conversely, suppose we have names for (i), (ii), and (iii). Fix $n \in \mathbb{N}$, $\sigma = \sigma_n$, and the sequence $(t_j)_{j\in\mathbb{N}}$. Let $\langle r_k, f(r_k)\rangle$ be a string of uniformly

computable pairs such that $\sigma_k \le r_k < \sigma_{k+1}$ for all $k < |\sigma|$. Define a càdlàg function g by $g(t) = f(r_k)$ if $\sigma_k \le t < \sigma_{k+1}$ for some k, and $g(T) = f(T)$.

Since σ is fixed, it is easy to see that g is computably càdlàg. Moreover,

$$d_J(f,g) \le d_\infty(f,g) \le \max_k \eta_f[\sigma_k^n, \sigma_{k+1}^n) < 2^{-n}.$$

Thus we can compute f, which completes the proof. □

Theorem 3.3.6 is an easy way to show that we can concatenate, truncate, and perform other basic operations on càdlàg functions in a computable way, though endpoints will always present minor difficulties.

Corollary 3.3.7. Let $T_1, T_2 > 0$. Let $f \in \mathbf{D}_E[0,T_1]$ and $g \in \mathbf{D}_E[0,T_2]$. The concatenated function $f^\frown g$ defined by

$$f^\frown g(t) = \begin{cases} f(t) & \text{if } 0 \le t < T_1, \\ g(t-T_1) & \text{if } T_1 \le t \le T_1 + T_2 \end{cases} \tag{3.3.2}$$

is computable in $\mathbf{D}_E[0, T_1+T_2]$ uniformly from names for f, g, T_1, T_2.

Corollary 3.3.8. Let $T_2 > T_1 > 0$. Let $f \in \mathbf{D}_E[0,T_2]$. The truncated function g defined by $g(t) = f(t)$ for $0 \le t < T_1$ and $g(T_1) = f(T_1-)$ is computable in $\mathbf{D}_E[0,T_1]$ uniformly from names for f, T_1, T_2.

Corollary 3.3.9. Let $T_1, T_2 > 0$. Let $f \in \mathbf{D}_E[0,T_1]$. The càdlàg function g defined by $g(t) = f(t \cdot T_1/T_2)$ for $0 \le t \le T_2$ is computable in $\mathbf{D}_E[0,T_2]$ uniformly from names for f, T_1, and T_2.

The next result, which assembles a piecewise constant càdlàg function, follows quite easily from Theorem 3.3.6 or directly from the definitions.

Corollary 3.3.10. Let $a_0, a_1, \ldots, a_n \in E$. Let $0 = t_0 < t_1 < \cdots < t_n = T$. The càdlàg function f defined by $f(t) = a_i$ if $t \in [t_i, t_{i+1})$ for some $i < n$ and $f(t_n) = a_n$ is computable in $\mathbf{D}_E[0,T]$ from names for $t_0, a_0, \ldots, t_n, a_n$.

3.3.2 Jump complexity

We collect some useful facts about the discontinuities of càdlàg functions.

The results are stated for càdlàg functions taking values in a computable Banach space $\langle V, \|\cdot\|, S\rangle$ because it allows us to define the jump function $\mathcal{J}_f(t) = f(t) - f(t-)$. However, the results can typically be generalized in a straightforward way to càdlàg functions taking values in an arbitrary computable metric space E.

We begin with an auxiliary result about approximating a single jump:

Lemma 3.3.11. Let $f, g \in \mathbf{D}_V[0,T]$. If $d_J(f,g) < \epsilon \le 1/2$, then

$$\|\mathcal{J}_f(t) - \mathcal{J}_g(s)\| < 2\epsilon \tag{3.3.3}$$

for some s such that $|t-s| < 2T\epsilon$. So if $\|\mathcal{J}_f(t)\| > K$, $\|\mathcal{J}_g(s)\| > K - 2\epsilon$.

Proof. By hypothesis, there exist a $\delta < \epsilon$ and a time change ϕ such that $\|\phi\|^\circ < \delta$ and $\|f - g\circ\phi\|_\infty < \delta$. By Fact 3.1.42, $\|\phi - \mathrm{id}\|_\infty < 2T\delta$.

Let $h = g\circ\phi$, so that $\|f-h\|_\infty < \delta$. Let $\gamma > 0$. For $r < t$, if $|r-t|$ is small enough, then $\|f(r) - f(t-)\| < \gamma$ and $\|h(r) - h(t-)\| < \gamma$. It follows that $\|f(t-) - h(t-)\| < \delta + 2\gamma$. γ was arbitrary, so $\|f(t-) - h(t-)\| \le \delta$.

Since $\mathcal{J}_f(t) = f(t) - f(t-)$ and $\mathcal{J}_h(t) = h(t) - h(t-)$, we have

$$\|\mathcal{J}_h(t) - \mathcal{J}_f(t)\| \le \|f(t) - h(t)\| + \|f(t-) - h(t-)\| \le 2\delta < 2\epsilon.$$

Let $s = \phi^{-1}(t)$, so that $\mathcal{J}_g(s) = \mathcal{J}_h(t)$. Since $|t-s| < 2T\delta < 2T\epsilon$ and the last claim in the lemma is obvious, this completes the proof. □

Lemma 3.3.11 generalizes to multiple jumps, which we state as

Proposition 3.3.12. Let $f, g \in \mathbf{D}_V[0,T]$. Let $d_J(f,g) < \epsilon \le 1/2$ and let $0 \le t_0 < \cdots < t_n \le T$ be distinct times. Then there exist distinct times $0 \le s_0 < \cdots < s_n \le T$ such that for all $i \le n$,

$$|t_i - s_i| < 2T\epsilon \quad \text{and} \quad \|\mathcal{J}_f(t_i) - \mathcal{J}_g(s_i)\| < 2\epsilon. \tag{3.3.4}$$

In particular, if $\|\mathcal{J}_f(t_i)\| > K_i$, $\|\mathcal{J}_g(s_i)\| > K_i - 2\epsilon$.

Proof. This is a straightforward generalization of the proof of Lemma 3.3.11. Regarding the distinctness of the times s_i, just note that when we set $s_i = \phi^{-1}(t_i)$, we get $s_i \ne s_j$ whenever $t_i \ne t_j$, for any time change ϕ. □

As stated, Lemma 3.3.11 and Proposition 3.3.12 have no computable content. Still, we use them to prove the key fact about jump computability:

Theorem 3.3.13. Uniformly from a name for $f \in \mathbf{D}_V[0,T]$, we can compute sequences $(t_i)_{i\in\mathbb{N}} \subseteq [0,T]$, $(v_i)_{i\in\mathbb{N}} \subseteq V$, and $(n_j)_{j\in\mathbb{N}} \subseteq \mathbb{N}$ such that

(i) for all i, $\mathcal{J}_f(t_i) = v_i \ne 0$,

(ii) each jump of f appears exactly once as a pair $\langle t_i, v_i\rangle$, and

(iii) for all i and j, if $i > n_j$, then $\|v_i\| < 2^{-j}$.

Before we prove Theorem 3.3.13, we state two (and a half) corollaries.

Corollary 3.3.14. If f is computably càdlàg, any given jump $\langle t, \mathcal{J}_f(t)\rangle$ of f is computable. (Note that this is not stated uniformly.)

Proof. Fix i in the sequences $(t_i)_{i\in\mathbb{N}}$, $(v_i)_{i\in\mathbb{N}}$ from Theorem 3.3.13. □

Corollary 3.3.15. Given $j \in \mathbb{N}$, uniformly from a name for $f \in \mathbf{D}_V[0,T]$, we can compute a Π^0_1 null set $A \subseteq [0,T]$ such that $\|\mathcal{J}_f(t)\| < 2^{-j}$ for all $t \notin A$.

Proof. Compute $(t_i)_{i\in\mathbb{N}}$, $(v_i)_{i\in\mathbb{N}}$, and n_j from Theorem 3.3.13. For all $n \in \mathbb{N}$, compute closed rational intervals $I_0, \ldots, I_{n_j-1} \subseteq [0,T]$ of length $2^{-n}/n_j$ and containing $t_0, \ldots, t_{n_j-1}$ (i.e., the first n_j jump times).

Let $B_n = [0,T] \setminus \bigcup_{k<n_j} I_k$, a finite union of open rational intervals. Then $B = \bigcup_{n\in\mathbb{N}} B_n$ is Σ^0_1. Of course, it follows that $A = [0,T] \setminus B$ is Π^0_1, and A is evidently a null set. This completes the proof. □

In addition to Corollaries 3.3.14 and 3.3.15, note that Theorem 3.3.13 affirms Proposition 3.3.5: the image of a one-to-one enumeration is Σ^0_2 (in a uniform way), and in this case the image of $(t_i)_{i\in\mathbb{N}}$ is the set of jump times.

Proof of Theorem 3.3.13. The following is uniform from a name for $f \in \mathbf{D}_V[0,T]$. We proceed in stages, indexed by $m \in \mathbb{N}$.

Let m be even and let $n \geq m+2$. Compute a simple function g such that $d_J(f,g) < 2^{-n}$. Compute the list of jumps $\langle t, \mathcal{J}_g(t)\rangle$ of g such that $\|\mathcal{J}_g(t)\| > 2^{-m}$, say $\langle s_0, w_0\rangle, \ldots, \langle s_k, w_k\rangle$.

By Proposition 3.3.12, each s_i is an approximation (to within $2^{-n+1}T$) of a distinct jump time t of f, and the corresponding w_i is an approximation (to within 2^{-n+1}) of $v = \mathcal{J}_f(t)$. Any jump $\langle t, v\rangle$ of f so approximated must satisfy $\|v\| > 2^{-m} - 2^{-n+1} \geq 2^{-m}/2$. Moreover, each jump $\langle t, v\rangle$ of f such that $\|v\| > 2^{-m} + 2^{-n+1}$ corresponds to exactly one pair $\langle s_i, w_i\rangle$.

Enumerate $s_0, \ldots, s_k$ as approximations of distinct jump times t of f such that $\|\mathcal{J}_f(t)\| > 2^{-m}/2$, and enumerate $w_0, \ldots, w_k$ as approximations of the corresponding values $\mathcal{J}_f(t)$. In short, the jumps we are enumerating by this stage are distinct, include all jumps $\langle t, v\rangle$ of f such that $\|v\| \geq 2 \cdot 2^{-m} > 2^{-m} + 2^{-n+1}$, and include only jumps $\langle t, v\rangle$ of f such that $\|v\| > 2^{-m}/2$.

Having completed stage m, we add 2 to m and repeat. At this stage, the jumps we are enumerating include all jumps $\langle t, v\rangle$ of f such that $\|v\| \geq 2^{-m}/2$. In particular, they include every jump that we partially enumerated in the previous stage, and every jump such that $2^{-m}/2 < \|v\| < 2 \cdot 2^{-m}$, which we may or may not have begun to enumerate in the previous stage.

The enumeration just described establishes the theorem: each jump $\langle t, v\rangle$ of f satisfies $\|v\| > 2^{-m}$ for some least even $m \in \mathbb{N}$, so it shows up in either

stage m or $m+2$. In any case, each jump is enumerated exactly once, and we can easily compute the sequence $(n_j)_{j\in\mathbb{N}}$, as required. □

Proposition 3.3.12 (jump approximation) can be used to prove a number of facts about the arithmetical complexity of sets in $\mathbf{D}_V = \mathbf{D}_V[0,T]$.

Proposition 3.3.16. Let $B \subseteq V$ be open. Let A be the set of $f \in \mathbf{D}_V$ that jump at least n times by vectors belonging to $B \setminus \{0\}$; that is, such that

$$|\{t : \mathcal{J}_f(t) \in B \setminus \{0\}\}| \geq n. \tag{3.3.5}$$

Then A is Σ^0_1 uniformly from n and a name for $B \in \tau(V)$.

Remark 3.3.17. Proposition 3.3.16 applies to sets of the following forms, which are Σ^0_1 uniformly from names for $a_0, a_1, a_2, a_3 > 0$:

$$B_1 = \{v : a_0 < \|v\| < a_1\}, \quad B_2 = \{v : \|v\| > a_2\}, \quad B_3 = \{v : \|v\| < a_3\}.$$

Proof of Proposition 3.3.16. The proof is a straightforward generalization of the following special case, when $n = 1$, B is an open ball, and $0 \notin B$.

Let A be the set of $f \in \mathbf{D}_V$ with at least one jump $v \in B_r(v_0)$, where $0 < r < \|v_0\|$. We will show that A is Σ^0_1 uniformly from names for r and v_0.

Suppose $f \in A$. Then $\|v - v_0\| < r - 4\epsilon$ for some positive rational $\epsilon < 1/2$. By Proposition 3.3.12, jump approximation (actually, in this special case, Lemma 3.3.11 suffices), any simple function g such that $d_J(f,g) < \epsilon$ will have a jump v_1 such that $\|v_1 - v_0\| < r - 2\epsilon$. Such a g exists.

Conversely, suppose g is a simple function with a jump v_1 such that $\|v_1 - v_0\| < r - 2\epsilon$ for some positive rational $\epsilon < 1/2$, and $d_J(f,g) < \epsilon$. By Proposition 3.3.12, f has a jump v such that $\|v - v_0\| < r - \epsilon$.

Given names for r, v_0, and ϵ, we can enumerate the simple functions g with at least one jump v_1 such that $\|v_1 - v_0\| < r - 2\epsilon$. It follows that

$$A = \{f : (\exists \epsilon)\ (\exists g)\ (\exists s)\ (d_J(f,g) < \epsilon \text{ and } \|\mathcal{J}_g(s) - v_0\| < 2\epsilon)\}$$

is Σ^0_1 uniformly from names for r and v_0, as required. □

With a little work, many other facts about arithmetical complexity in $\mathbf{D}_V$ follow from Proposition 3.3.12 plus the fact (3.1.12) that càdlàg functions have only finitely many jumps whose norm exceeds any given $K > 0$.

Proposition 3.3.18. Let $v_0 \in V \setminus \{0\}$. Let G be the set of $f \in \mathbf{D}_V$ with fewer than n jumps equal to v_0; that is, such that

$$|\{t : \mathcal{J}_f(t) = v_0\}| < n. \tag{3.3.6}$$

G is Σ^0_1 uniformly from n and a name for v_0.

Moreover, let $r > 0$ and let F be the set of $f \in \mathbf{D}_V$ with fewer than n jumps of norm equal to r. F is Σ^0_1 uniformly from n and a name for r.

Proof. Consider the first claim. The proof is a straightforward generalization of the following special case, when $n = 1$. The second claim is similar.

Let G be the set of $f \in \mathbf{D}_V$ with no jumps equal to v_0. We will show that G is Σ^0_1 uniformly from a name for v_0.

Suppose $f \in G$. f has only finitely many jumps v such that, say, $\|v\| > \|v_0\|/2$ (Fact 3.1.12), so there exists a positive rational $\epsilon < 1/2$ such that f has no jumps $v \in B_{4\epsilon}(v_0)$. By Proposition 3.3.12, any simple function g such that $d_J(f,g) < \epsilon$ must have no jumps $v_1 \in B_{2\epsilon}(v_0)$. Such a g exists.

Conversely, suppose there exists a positive rational $\epsilon < 1/2$ and a simple function g such that $d_J(f,g) < \epsilon$ and g has no jumps $v_1 \in B_{2\epsilon}(v_0)$. By Proposition 3.3.12, f has no jumps equal to v_0.

Given v_0 and ϵ, we can enumerate the simple functions g with no jumps v_1 such that $\|v_1 - v_0\| < 2\epsilon$. It follows that the set

$$G = \{f : (\exists \epsilon)\ (\exists g)\ (d_J(f,g) < \epsilon \text{ and } (\forall s)\ \|\mathcal{J}_g(s) - v_0\| \geq 2\epsilon)\}$$

is Σ^0_1 uniformly from a name for v_0, as required. □

Together, Proposition 3.3.12 and Theorem 3.3.13 (jump enumeration) can be used to evaluate the complexity of sets defined not only by the number and size of jumps, but also by the times when they occur.

Proposition 3.3.19. Uniformly from a name for $t \in [0,T]$, the set of $f \in \mathbf{D}_V$ such that $\|\mathcal{J}_f(t)\| = 0$ (i.e., such that t is a continuity point of f) is Π^0_2.

Proof. By Proposition 3.3.12, $\|\mathcal{J}_f(t)\| = 0$ if and only if for all positive rational ϵ there exists a simple function g and a rational interval I containing t such that $d_J(f,g) < \epsilon$ and $\|\mathcal{J}_g(s)\| = 0$ for all $s \in I$. This is Π^0_2 uniformly in a name for t, as required. □

Remark 3.3.20. Proposition 3.3.19 exchanges the roles of $f \in \mathbf{D}_V[0,T]$ and $t \in [0,T]$ in Proposition 3.3.5.

Just one more example of this sort of jump complexity fact before we proceed to effective measurability in Skorokhod space:

Proposition 3.3.21. Uniformly from names for $t \in [0,T]$ and an open set $B \subseteq V$, the set of $f \in \mathbf{D}_V$ such that $\mathcal{J}_f(t) \in B \setminus \{0\}$ is Π^0_2.

Proof. The following is uniform from names for t and B. ϵ and δ range over the positive rationals.

Suppose $v = \mathcal{J}_x(t) \in B \setminus \{0\}$. Fixing ϵ, for small enough δ, by Proposition 3.3.12, any simple function g such that $d_J(f,g) < \delta$ will have a jump $\langle s, v_1 \rangle$ such that $|s-t| < \epsilon$ and $v \in B(v_1, 2\delta)$, where $B(v_1, 2\delta) \subseteq B$.

Conversely, suppose that for all ϵ, there exist a δ and a simple function g such that $d_J(f,g) < \delta < \epsilon$ and g has a jump $\langle s, v_1 \rangle$ such that $|s-t| < \epsilon$ and $B(v_1, 2\delta) \subseteq B$. By Proposition 3.3.12 and the fact that f has only finitely many jumps larger than any given $K > 0$, f must jump at exactly time t. If we set $v = \mathcal{J}_f(t)$, we have $v \in B(v_1, 2\delta) \subseteq B$.

The following property of f is Π^0_2: for all ϵ, there exist a δ and a simple function g such that $d_J(f,g) < \delta < \epsilon$ and g jumps by v_1 at s where $|s-t| < \epsilon$ and $B(v_1, 2\delta) \subseteq B$. This completes the proof. □

3.3.3 Effective measurability

As we said in the introduction to this chapter, a stochastic process on a closed and bounded time interval whose sample paths are càdlàg, or at least almost surely càdlàg, may be treated as a random variable taking values in Skorokhod space. (This may not be obvious: see Fact 3.3.24.) Here we develop the basic tools needed to effectivize such a "random function." They will be used in Chapters 4 and 5 for Lévy and Feller processes, respectively.

We begin with some essential terminology:

Notation 3.3.22. For all $0 \le t \le T$, $\pi_t : \mathbf{D}_E \to E$ denotes the *projection*

$$\pi_t : f \mapsto f(t). \tag{3.3.7}$$

If $X : \Omega \to \mathbf{D}_E$, then X_t or $X(t) : \Omega \to E$ denotes the *coordinate process*

$$\pi_t \circ X : \omega \mapsto X(\omega)(t). \tag{3.3.8}$$

Our first two facts about measurability in Skorokhod space are proved in Billingsley [7]. As usual, by "measurable," we mean Borel measurable.

Fact 3.3.23. The projection $\pi_t : \mathbf{D}_E \to E$ is measurable for all $0 \le t \le T$.

Fact 3.3.24. Let $\langle \Omega, \mathscr{F} \rangle$ be a measurable space and let $X : \Omega \to \mathbf{D}_E$. Then X is measurable if and only if $X_t : \Omega \to E$ is measurable for all $0 \le t \le T$.

We used Proposition 3.3.12 (jump approximation) and Theorem 3.3.13 (jump enumeration) to establish a number of facts about jump complexity in Skorokhod space. If we know that jumps of a given size do not occur, or at least almost surely (a.s.) do not occur, then we can say more.

Proposition 3.3.25. Let $\langle \Omega, \mathbb{P} \rangle$ be a computable probability space. Let $X : \Omega \to \mathbf{D}_V[0,T]$ be $L^0(\mathbb{P})$-computable. Suppose there exist $r_1, r_2 > 0$ such that for $\mathbb{P}$-almost every (a.e.) $\omega \in \Omega$, for all $0 \leq t \leq T$,

$$\|\mathcal{J}_{X(\omega)}(t)\| \notin \{r_1, r_2\}. \tag{3.3.9}$$

Let $\Psi(x) = \langle t_1, \ldots, t_k \rangle$, where $0 < t_1 < \cdots < t_k \leq T$ are exactly the times when $r_1 < \|\mathcal{J}_x(t)\| < r_2$. Then $\Psi : \mathbf{D}_V[0,T] \to [0,T]^*$ is $\mathbb{P}_X$-a.s. computable uniformly from names for r_1 and r_2. This also holds when $r_2 = \infty$.

Proof. By hypothesis, $\mathbb{P}_X$-a.s., $f \in \mathbf{D}_V[0,T]$ has no jump v such that $\|v\| = r_1$ or r_2. But if $\|v\| \neq r_1, r_2$, then we can decide $r_1 < \|v\| < r_2$ computably, given a name for v. Apply Theorem 3.3.13 to complete the proof. □

The next result is also not difficult to prove, but it is quite important. It extends Corollary 3.3.10, assembling a random càdlàg function.

Theorem 3.3.26. Again, let $\langle \Omega, \mathbb{P} \rangle$ be a computable probability space. Let $F_0, \ldots, F_n : \Omega \to E$ and $U_0, \ldots, U_n : \Omega \to \mathbb{R}$ be measurable. Suppose that

$$(\forall \omega \in \Omega) \quad 0 = U_0(\omega) < U_1(\omega) < \cdots < U_n(\omega) = T. \tag{3.3.10}$$

Define a function $X : \Omega \to \mathbf{D}_E[0,T]$ by

$$X_t(\omega) = F_i(\omega) \quad \text{if } t \in [U_i(\omega), U_{i+1}(\omega)) \text{ for some } i < n, \tag{3.3.11}$$

and $X_T(\omega) = F_n(\omega)$. Then X is $L^0(\mathbb{P})$-computable uniformly from names for $\mathbb{P} \in \mathcal{M}_1(\Omega)$, $F_0, \ldots, F_n \in L^0(\Omega, \mathbb{P}; E)$, and $U_0, \ldots, U_n \in L^0(\Omega, \mathbb{P})$.

Proof. Corollary 3.3.10 defined a partial computable function

$$\Phi :\subseteq E^{n+1} \times [0,T]^{n+1} \to \mathbf{D}_E[0,T]$$

such that $\Phi \circ \langle F_0, \ldots, F_n, U_0, \ldots, U_n \rangle = X$ on all of Ω. By Fact 2.3.48, X is uniformly $L^0(\mathbb{P})$-computable, as required. □

Our final result exchanges the roles of $f \in \mathbf{D}_E[0,T]$ and $t \in [0,T]$ in Corollary 3.3.4. (Compare Remark 3.3.20.)

Proposition 3.3.27. Once again, let $\langle \Omega, \mathbb{P} \rangle$ be a computable probability space. Let $X : \Omega \to \mathbf{D}_E[0,T]$ be $L^0(\mathbb{P})$-computable. Suppose that for all $0 \leq t \leq T$, for $\mathbb{P}$-a.e. $\omega \in \Omega$, $\Delta X_t(\omega) = 0$. Then the *evaluation mapping*

$$\mathrm{eval} : [0,T] \to L^0(\mathbf{D}_E[0,T], \mathbb{P}_X; E) \tag{3.3.12}$$

given by $\mathrm{eval}(t) : f \mapsto f(t)$ is computable.

Proof. The following is uniform from a name for $0 \le t \le T$.

By Theorem 3.3.1, $f \mapsto f(t)$ is computable on $\{f \in \mathbf{D}_E[0,T] : \Delta f(t) = 0\}$. Since $\Delta f(t) = 0$ a.s. $[\mathbb{P}_X]$, this means that $f \mapsto f(t)$ is $\mathbb{P}_X$-a.e. computable, hence $L^0(\mathbb{P}_X)$-computable by Fact 2.3.43. □

3.3.4 Counterexamples

Here we present a few basic negative results in effective Skorokhod space. For one thing, Theorem 3.3.13 (jump enumeration) cannot be improved:

Proposition 3.3.28. Given a sequence $(t_i)_{i\in\mathbb{N}}$ of distinct reals in $(0,T]$, we can compute an $f \in \mathbf{D}_V[0,T]$ such that f jumps exactly on $\{t_i : i \in \mathbb{N}\}$.

Corollary 3.3.29. A computably càdlàg function can jump on a dense set.

Proof of Proposition 3.3.28. The following is uniform from $(t_i)_{i\in\mathbb{N}}$.

We will compute a sequence $(f_n)_{n\in\mathbb{N}}$ of càdlàg functions converging in $\mathbf{D}_V[0,T]$ at a computable rate. First, fix any computable unit vector $\alpha \in V$. Recalling Example 3.1.18, let $f_0 = \mathbf{1}_{[t_0,T]}\alpha$.

Evidently, f_0 is computably càdlàg. We compute the rest of the sequence by induction. Suppose we have already computed piecewise-constant càdlàg functions $f_0, \dots, f_n$ in such a way that

(i) for all $k \le n$, f_k jumps exactly on $\{t_0, \dots, t_k\}$, and

(ii) for all $k < n$, $d_J(f_k, f_{k+1}) \le 2^{-(k+1)}$.

We define f_{n+1} as follows. Compute an $m \in \mathbb{N}$ such that $m \ge n+3$, $2^{-m} < T$, and $|t_i - t_j| > 2^{-m}$ for all $i < j \le n$. Set

$$f_{n+1} = f_n + 2^{-2m-1} \cdot \mathbf{1}_{[t_{n+1},T]}\alpha.$$

Evidently, f_{n+1} jumps exactly on $\{t_0, \dots, t_{n+1}\}$. Since t_{n+1} is computable, it is easy to see that f_{n+1} is computably càdlàg whenever f_n is.

By Fact 3.1.40, if $d_J^\circ(f_{n+1}, f_n) < \delta^2$ for some $\delta < \min\{T, 1/4\}$, then

$$d_J(f_{n+1}, f_n) \le 4\delta + \pi_{\mathbf{D}}(f_n, \delta),$$

where $\pi_{\mathbf{D}}$ is the càdlàg modulus. Set $\delta = 2^{-m}$. Then $\pi_{\mathbf{D}}(f_n, \delta) = 0$ by construction, as all the jump times are further than δ apart, and f_n is constant in between. The trivial time change $\phi = \mathrm{id}$ shows that

$$d_J^\circ(f_{n+1}, f_n) \le \|f_{n+1} - f_n\|_\infty = 2^{-2m-1} < \delta^2.$$

Therefore,

$$d_J(f_{n+1}, f_n) \le 4\delta + 0 = 2^{-m+2} \le 2^{-(n+1)}.$$

The Cauchy sequence $(f_n)_{n\in\mathbb{N}}$ of càdlàg functions converges in $\mathbf{D}_V[0,T]$ to some computably càdlàg function f. By construction, it is easy to see that f jumps exactly on $\{t_i : i \in \mathbb{N}\}$, as required. □

Theorem 3.3.6 was a useful alternative characterization of computably càdlàg functions. Of course, there may be other such characterizations. We know that every computably càdlàg function is a.e. computable (Corollary 3.3.4) and computably bounded (Theorem 3.3.6). Also, its discontinuities are computably enumerable (Theorem 3.3.13). But these properties do not characterize the computably càdlàg functions:

Example 3.3.30. The following real-valued function is a.e. computable and bounded by 1 with one discontinuity (at 1/2), but it is not even càdlàg:

$$f(t) = \begin{cases} \sin((1/2 - t)^{-1}) & \text{if } t < 1/2, \\ 0 & \text{if } t \ge 1/2. \end{cases} \tag{3.3.13}$$

Now suppose f is a càdlàg function taking values in a Banach space V. We might like to define—and ideally compute—a "continuous part" of f, which would be a continuous function $\widetilde{f} : [0,T] \to V$ such that for all t,

$$\widetilde{f}(t) = f(t) - \sum_{t_i \le t} \mathcal{J}_f(t_i) \quad \text{(hypothetically)}$$

where t_i ranges over the jump times of f. However, this is not possible:

Example 3.3.31. We define a real-valued computably càdlàg function f on $[0,1]$ that has no "continuous part" (computable or otherwise). For all $k \in \mathbb{N}$, let $t_k = 1/2 - 2^{-k-2}$ and let $a_k = 1/(k+1)$, i.e.,

$$(t_k)_{k\in\mathbb{N}} = (1/4, 3/8, 7/16, \ldots) \quad \text{and} \quad (a_k)_{k\in\mathbb{N}} = (1, 1/2, 1/3, \ldots).$$

Let $f = 0$ on $[0, t_0)$ and on $[1/2, 1]$. For all $k \in \mathbb{N}$, let f jump by a_k at t_k, then make f linear on $[t_k, t_{k+1})$ with slope $m_k = -a_k/(t_{k+1} - t_k)$, so that $f(t_{k+1}-)$ is zero. It is not difficult to see that f is computably càdlàg on $[0,1]$.

Let $\widetilde{f}$ be the continuous part of f on $[0, 1/2)$. $\widetilde{f}$ is well defined (and piecewise linear), but $\widetilde{f}(t) \downarrow -\infty$ as $t \uparrow 1/2$, since $\widetilde{f}(t_{k+1}) - \widetilde{f}(t_k) = -a_k$.

Fortunately, a càdlàg function like the one in Example 3.3.31 is rather unlikely to show up as the sample path of, say, a Lévy process—a class of stochastic processes to which we now turn.

Chapter 4

Effective theory of Lévy processes

A Lévy process is a type of vector-valued stochastic process in continuous time. In this chapter, we develop a basic effective theory of Lévy processes, first explaining what it means for such a process to be computable, then proving effective versions of two important results from the classical theory.

We assume the reader is familiar with the principles of classical measure-theoretic probability as in Pollard [61] or Grimmett and Stirzaker [42].

Without getting into all the details, a Lévy process $X = \{X_t : t \geq 0\}$ has independent, stationary increments, so that (i) successive changes in position, say $X_{t_1} - X_{t_0}$, $X_{t_2} - X_{t_1}$, and so on, are statistically independent, and (ii) the probability distribution of the increment $X_t - X_s$ depends, not on the initial time s, nor on the initial position X_s, but only on the elapsed time $t - s$. A Lévy process is a continuous-time analogue of a random walk.

Brownian motion is one example of a Lévy process; a compound Poisson process is another. These examples show that a Lévy process, in general, can move continuously (in space) or jump. In fact, this is about all it can do: according to an important classical theorem, which we will effectivize, every Lévy process admits a modification whose sample paths are càdlàg, so they live in Skorokhod space (which, of course, we just effectivized).

Moreover, the Lévy-Itô decomposition, which we will also effectivize, states that every Lévy process can be written as the sum of four processes: a deterministic linear drift, a Brownian motion, a compound Poisson process which consists of all the jumps above a certain size, and finally a purely discontinuous martingale which consists of all the other, small jumps.

In Section 4.1, we review this classical theory in a little more detail.

In Section 4.2, we show that a càdlàg modification of a Lévy process not only exists, but is computable from a suitable (effective) representation of the process. In Section 4.3, we show that the components of the Lévy-Itô decomposition are similarly computable from a representation.

4.1 Classical theory of Lévy processes

This section is review, based on the well-known Applebaum [3], which the reader may consult for further details, including proofs.

4.1.1 Stochastic processes

As a preliminary, we establish some notational conventions.

Notation 4.1.1. Let $\langle \Omega, \mathscr{F} \rangle$ be a measurable space. We may refer to the elements of $\mathscr{F}$, i.e., the measurable subsets of Ω, as *events*.

Let $\langle E, \mathscr{E} \rangle$ be a measurable space, too. We may refer to $\mathscr{F}/\mathscr{E}$-measurable functions $X : \Omega \to E$ as *random variables*, in which case Ω is said to be the *sample space* and E the *state space*.

Let $X : \Omega \to E$ be a random variable and let $A \subseteq E$ be an event. We may denote by $\{X \in A\}$ the event $\{\omega : X(\omega) \in A\} \subseteq \Omega$.

Let $\mathbb{P}$ be a probability measure on $\langle \Omega, \mathscr{F} \rangle$. Let $X : \Omega \to E$ be a random variable. As in Definition 2.3.44, we denote by $\mathbb{P}_X$ the *image measure* of $\mathbb{P}$ along X, which is the probability measure on $\langle E, \mathscr{E} \rangle$ given by

$$\mathbb{P}_X(A) = \mathbb{P}\{X \in A\}.$$

Finally, we adopt the *de Finetti notation* as in Pollard [61]:

(i) We will use the same symbol for a probability measure $\mathbb{P}$ on $\langle \Omega, \mathscr{F} \rangle$ and for the associated *expectation operator* on $\mathbb{R}$-valued random variables,

$$f \mapsto \mathbb{P}f = \int_\Omega f(\omega)\mathbb{P}(d\omega).$$

(ii) We may identify an event A with its *indicator function* $\mathbf{1}_A : E \to \{0, 1\}$, treating it as an $\mathbb{R}$-valued random variable.

We proceed to some basic definitions, beginning with two important equivalence relations on random variables. Throughout, $\langle \Omega, \mathscr{F}, \mathbb{P} \rangle$ denotes a probability space and $\langle E, \mathscr{E} \rangle$ a measurable space.

Definition 4.1.2. Random variables $X, Y : \Omega \to E$ are said to be *equal in distribution* if $\mathbb{P}_X = \mathbb{P}_Y$, in which case we write $X \sim Y$.

Definition 4.1.3. Random variables X, Y are said to be *equal almost surely* if $\mathbb{P}\{X = Y\} = 1$, in which case we write $X = Y$ a.s. $[\mathbb{P}]$.

Fact 4.1.4. If $X = Y$ a.s. $[\mathbb{P}]$, then $X \sim Y$.

To each $\mathbb{R}^n$-valued random variable we associate a $\mathbb{C}$-valued function:

Definition 4.1.5. Let $X : \Omega \to \mathbb{R}^n$ be a random variable. The *characteristic function* of X, $\phi_X : \mathbb{R}^n \to \mathbb{C}$, is given by

$$\phi_X(u) = \mathbb{P}\left[e^{\mathbf{i}(u,X)}\right] = \int_\Omega e^{\mathbf{i}(u,X(\omega))}\mathbb{P}(d\omega) = \int_{\mathbb{R}^n} e^{\mathbf{i}(u,y)}\mathbb{P}_X(dy). \qquad (4.1.1)$$

Here, $\mathbf{i}$ denotes $\sqrt{-1}$, and $(\cdot,\cdot)$ is the Euclidean inner product on $\mathbb{R}^n$.

Fact 4.1.6. ϕ_X is continuous at the origin, $\phi_X(0) = 1$, and $\|\phi_X\|_\infty \le 1$.

A stochastic process is a sequence of random variables representing a system whose state changes randomly over time:

Definition 4.1.7. An E-valued (continuous-time) *stochastic process* X on Ω is a family $\{X_t : t \in I\}$ of random variables $X_t : \Omega \to E$, where $I \subseteq [0, \infty)$. In that case, we may write $X(t)$ or $X(t, \cdot)$ instead of X_t. In general, I is an interval, but unless otherwise noted, our processes are indexed by $I = [0, \infty)$.

Definition 4.1.8. A function $I \to E$ of the form $t \mapsto X_t(\omega)$ for some $\omega \in \Omega$ is said to be a *sample path* of the stochastic process X.

There are two important equivalence relations on stochastic processes. Let $X = \{X_t : t \ge 0\}$ and $Y = \{Y_t : t \ge 0\}$ be E-valued processes on Ω.

Definition 4.1.9. X and Y are said to be *modifications* of one another if for all t, for $\mathbb{P}$-a.e. ω, $X_t(\omega) = Y_t(\omega)$.

Definition 4.1.10. X and Y are said to be *indistinguishable* if for $\mathbb{P}$-a.e. ω, for all t, $X_t(\omega) = Y_t(\omega)$.

Fact 4.1.11. Indistinguishable processes are modifications of one another.

Now assume that our state space E is a metric space (with the Borel σ-algebra). Our interest, again, is in stochastic processes with càdlàg sample paths—or at least with modifications with càdlàg sample paths.

Definition 4.1.12. A stochastic process $X = \{X_t : t \in I\}$ is said to be *càdlàg* if for $\mathbb{P}$-a.e. ω, the sample path $t \mapsto X_t(\omega)$ is càdlàg. X is said to be *continuous* if for $\mathbb{P}$-a.e. ω, the sample path $t \mapsto X_t(\omega)$ is continuous.

Fact 4.1.13. Let X be a stochastic process. Any two càdlàg modifications of X are indistinguishable.

4.1.2 Lévy processes

Lévy processes are characterized by three plus one properties: independent increments, stationary increments, stochastic continuity, and they have to start at the origin, which can be achieved by normalizing.

To define stochastic continuity, we need to assume that the state space is a metric space, say $\langle E, d\rangle$ (with the Borel σ-algebra), which is reasonable. So let $X = \{X_t : t \geq 0\}$ be an E-valued stochastic process.

Definition 4.1.14. Let $t \geq 0$. X is said to be *stochastically continuous at* t if for all $\epsilon > 0$, there exists a $\delta > 0$ such that for all $s \geq 0$,

$$|t - s| < \delta \implies \mathbb{P}\{d(X_t, X_s) \geq \epsilon\} < \epsilon. \tag{4.1.2}$$

X is said to be *stochastically continuous*[1] if it is stochastically continuous at t for all $t \geq 0$.

Fact 4.1.15. X is stochastically continuous at t if and only if the function $s \mapsto X_s$ of type $[0, \infty) \to L^0(\Omega, \mathbb{P}; E)$ is continuous at t.

For completeness, we include a proof, which is easy—and effective.

Proof. Suppose X is stochastically continuous at t and fix $\epsilon > 0$. For small enough δ, $|t - s| < \delta$ implies $\mathbb{P}(A) < \epsilon/2$, where $A = \{d(X_t, X_s)\| \geq \epsilon/2\}$. By a standard calculation,

$$d_{\mathbb{P}}(X_t, X_s) = \int_\Omega \min\{1, d(X_t, X_s)\}\, d\mathbb{P} \leq \int_A 1 d\mathbb{P} + \int_{\Omega \setminus A} \epsilon/2 d\mathbb{P} < \epsilon/2 + \epsilon/2 = \epsilon.$$

Hence $s \mapsto X_s$ is continuous at t.

Conversely, suppose $s \mapsto X_s$ is continuous at t and fix $\epsilon > 0$. Then for sufficiently small δ, $|t - s| < \delta$ implies

$$d_{\mathbb{P}}(X_t, X_s) = \int_\Omega \min\{1, d(X_t, X_s)\}\, d\mathbb{P} < \epsilon^2.$$

Hence $\mathbb{P}\{d(X_t, X_s) \geq \epsilon\}$ must be less than ϵ, which completes the proof. □

In order to define even an increment, we need to assume that the state space is a normed space, say $\langle V, \|\cdot\|\rangle$ (with the Borel σ-algebra again), which is still reasonable. So let X be a V-valued stochastic process.

Definition 4.1.16. X is said to have *independent increments* if $X_{t_1} - X_{t_0}$, $X_{t_2} - X_{t_1}$, …, $X_{t_n} - X_{t_{n-1}}$ are independent whenever $0 \leq t_0 < t_1 < \cdots < t_n$.

[1] The term "continuous in probability" is used in some sources.

Definition 4.1.17. X is said to have *stationary increments* if

$$(\forall t > s \geq 0) \quad X_t - X_s \sim X_{t-s} - X_0. \tag{4.1.3}$$

Fact 4.1.18. If X has independent and stationary increments, then X is stochastically continuous if and only if X is stochastically continuous at 0.

Now, Definitions 4.1.16 and 4.1.17 will not need to be effectivized at all when we turn to computability in Section 4.2. Definition 4.1.14 can easily be effectivized in either that or the equivalent form given by Fact 4.1.15.

In any case, now we can define a Lévy process. Fix $n \in \mathbb{N}$ and let the state space be $\mathbb{R}^n$ (with the Euclidean norm).

Definition 4.1.19. An $\mathbb{R}^n$-valued stochastic process $X = \{X_t : t \geq 0\}$ is said to be an (n-dimensional) *normalized Lévy process* if

(i) X has independent, stationary increments,

(ii) X is stochastically continuous, and

(iii) $X_0 = 0$ a.s. $[\mathbb{P}]$.

When we said that (iii) can be achieved by normalizing, we meant

Remark 4.1.20. Evidently, if an $\mathbb{R}^n$-valued process $Y = \{Y_t : t \geq 0\}$ satisfies (i) and (ii) in Definition 4.1.19, then $Y - Y_0$ is a normalized Lévy process. And if X is a normalized Lévy process and $v : \Omega \to \mathbb{R}^n$ is a random vector, then $Y = X + v$ satisfies (i), (ii), and $Y_0 = v$ a.s. $[\mathbb{P}]$.

In light of this remark, our next definition is not much of a generalization.

Definition 4.1.21. Let $v_0 : \Omega \to \mathbb{R}^n$ be a random vector. X is said to be a *Lévy process with initial position* v_0 if $X - v_0$ is a normalized Lévy process.

Let μ_0 a probability measure on $\mathbb{R}^n$. X is said to be a *Lévy process with initial distribution* μ_0 if $X - X_0$ is a normalized Lévy process and $\mathbb{P}_{X_0} = \mu_0$.

Typically, facts about normalized Lévy processes generalize to processes with a given initial position or distribution. Accordingly:

Notation 4.1.22. From now on, by a *Lévy process* we mean a normalized Lévy process: $X_0 = 0$ a.s. $[\mathbb{P}]$.

Next, we associate to each Lévy process a $\mathbb{C}$-valued continuous function, which is closely related to the characteristic function (see Definition 4.1.5).

Fact 4.1.23. Let X be a Lévy process. There exists a unique continuous function $\eta : \mathbb{R}^n \to \mathbb{C}$, called the *Lévy symbol* of X, such that

$$(\forall u \in \mathbb{R}^n) \quad (\forall t \geq 0) \quad \phi_{X_t}(u) = e^{t\eta(u)}. \tag{4.1.4}$$

Moreover, $\eta(u) = \ln \mathbb{P}[\exp(\mathbf{i}uX_1)]$ for all $u \in \mathbb{R}^n$.

We also associate to X a certain Borel measure, as follows.

Definition 4.1.24. A *Lévy measure* is a (Borel) measure ν on $\mathbb{R}^n$ such that

(i) $\nu\{0\} = 0$, and

(ii) $\int_{\mathbb{R}^n} \min\{1, |y|^2\}\nu(dy) < \infty$.

Remark 4.1.25. A Lévy measure is necessarily σ-finite.

The Lévy-Itô decomposition (Theorem 4.1.43) may be used to prove

Fact 4.1.26. Let X be a Lévy process. There exist a $b \in \mathbb{R}^n$, a positive-definite $n \times n$ symmetric matrix A, and a Lévy measure ν such that for all $u \in \mathbb{R}^n$, for all $t \geq 0$,

$$\begin{aligned} \phi_{X_t}(u) &= \exp\Big(t\big[\mathbf{i}(u,b) - \frac{1}{2}(u,Au) \\ &+ \int_{\mathbb{R}^n} (e^{\mathbf{i}(u,y)} - 1 - \mathbf{i}(u,y)\mathbf{1}_{|y|<1})\nu(dy)\big]\Big). \end{aligned} \tag{4.1.5}$$

This is called the *Lévy-Khintchine formula*.

Notation 4.1.27. $\langle b, A, \nu \rangle$ are called the *characteristics* of X. In particular, b is the *drift* and ν is the *intensity measure*.

Fact 4.1.28. Let Y be a modification of a Lévy process X. Then Y is a Lévy process with the same characteristics as X.

That brings us to the first major result from the classical theory of Lévy processes that we intend to effectivize in this chapter:

Fact 4.1.29. Every Lévy process has a càdlàg modification.

Finally, we need to be able to describe the jumps of a càdlàg process X.

Definition 4.1.30. The *jump process* ΔX is given by

$$\Delta X(t) = X(t) - X(t-) \quad \text{for all } t \geq 0. \tag{4.1.6}$$

Fact 4.1.31. For a Lévy process X, for all $t \geq 0$, for $\mathbb{P}$-a.e. ω, $\Delta X_t(\omega) = 0$.

But we can say much more about Lévy processes, particularly the jumps, if we use another important theorem: the Lévy-Itô decomposition.

4.1.3 Lévy-Itô decomposition

The second major result we plan to effectivize in this chapter is the Lévy-Itô decomposition. Explaining what it says will take up the remainder of this section. (For a proof of the result, see Section 2.4 in Applebaum [3].)

Let $X = \{X_t : t \geq 0\}$ be a Lévy process. We need to take a closer look at the jumps of X. For simplicity, and without any real loss of generality, we assume that X is one-dimensional, i.e., that the state space is $\mathbb{R}$.

Definition 4.1.32. The *Poisson random measure* associated with X is the family $\{N_t : t \geq 0\}$ of functions $N_t : \Omega \times \mathscr{B}(\mathbb{R}) \to [0, \infty]$ given by

$$N_t(\omega, A) = N_t(A)(\omega) = |\{s \leq t : 0 \neq \Delta X_s(\omega) \in A\}|. \tag{4.1.7}$$

In other words, the Poisson random measure counts jumps of certain sizes that occur by a certain time. We list some other basic properties in

Fact 4.1.33. For all $t \geq 0$,

(i) for all $\omega \in \Omega$, $A \mapsto N_t(\omega, A)$ is an $\mathbb{N} \cup \{\infty\}$-valued measure on $\mathbb{R}$,

(ii) for all $A \in \mathscr{B}(\mathbb{R})$, $N_t(A) : \omega \mapsto N_t(\omega, A)$ is a random variable, and

(iii) $A \mapsto \mathbb{P}(N_t(A))$ is a measure on $\mathbb{R}$.

Definition 4.1.34. A measurable subset A of $\mathbb{R}$ is said to be *bounded away from zero*[2] if $0 \notin \overline{A}$, i.e., if there exists an $\epsilon > 0$ such that $A \cap B(0, \epsilon) = \varnothing$.

We collect a few more basic properties of the Poisson random measure:

Fact 4.1.35. If A is bounded away from zero, then

(i) for all $t \geq 0$, $N_t(A) < \infty$ a.s. $[\mathbb{P}]$, and

(ii) $\{N_t(A) : t \geq 0\}$ is a Poisson process, called the *jump-counting process*, with intensity $\nu(A) < \infty$, where ν is the intensity measure of X.

Fact 4.1.36. If $A_1, \ldots, A_n$ are disjoint and bounded away from zero, and $t_1, \ldots, t_n \geq 0$, then the random variables $N_{t_i}(A_i)$ are independent.

Fact 4.1.37. The intensity measure ν is given by $\mathbb{P}(N_1)$: for all $A \in \mathscr{B}(\mathbb{R})$,

$$\nu(A) = \mathbb{P}(N_1(A)) = \mathbb{P}(|\{s \leq 1 : 0 \neq \Delta X_s \in A\}|). \tag{4.1.8}$$

Moreover, $\mathbb{P}(N_t(A)) = t\nu(A)$ for all $t \geq 0$.

[2]The term "bounded below" is used in some sources, including Applebaum [3].

The Poisson random measure is used to define other random variables, and even stochastic processes, based on the jumps of X. Let $A \subseteq \mathbb{R}$ be measurable. Let $f : \subseteq \mathbb{R} \to \mathbb{R}$ be measurable and defined on A.

Definition 4.1.38. Suppose A is bounded away from zero. For all $t \geq 0$, the *Poisson integral* of f over A at t is the random finite sum

$$\left(\int_A f(x)N_t(dx)\right)(\omega) = \sum_{x \in A} f(x)N_t(\omega, \{x\}). \tag{4.1.9}$$

Fact 4.1.39. $\left\{\int_A f(x)N_t(dx) : t \geq 0\right\}$ is a compound Poisson process.

Definition 4.1.40. Suppose A is bounded away from zero. For all $t \geq 0$, the *compensator* $\widetilde{N}_t(A)$ is the random variable

$$\widetilde{N}_t(A) = N_t(A) - t\nu(A). \tag{4.1.10}$$

Definition 4.1.41. Suppose f is ν-integrable over A. For all $t \geq 0$, the *compensated Poisson integral* of f over A at t is the random variable

$$\int_A f(x)\widetilde{N}_t(dx) = \int_A f(x)N_t(dx) - t\int_A f(x)\nu(dx). \tag{4.1.11}$$

(Note that we did not assume that A was bounded away from zero.)

Fact 4.1.42. $\left\{\int_A f(x)\widetilde{N}_t(dx) : t \geq 0\right\}$ is a purely discontinuous martingale.

That brings us to the Lévy-Itô decomposition, which is the second major result from the classical theory that we intend to effectivize:

Fact 4.1.43. Let X be a Lévy process. For all $\epsilon > 0$, there exist a $b \in \mathbb{R}$ and a Brownian motion $W = \{W_t : t \geq 0\}$ such that for all $t \geq 0$,

$$X_t = bt + W_t + \int_{|x|<\epsilon} x\widetilde{N}_t(dx) + \int_{|x|\geq\epsilon} xN_t(dx), \tag{4.1.12}$$

and all of these components are independent.

Notation 4.1.44. The third component, which is a purely discontinuous martingale, is called the *compensated sum of small jumps*. We can call the fourth component the *sum of large jumps*; it is a compound Poisson process.

But before we can effectivize any results, we need to answer more basic questions, like what it means for a Lévy process to be computable.

4.2 Computable Lévy processes

In this section, we present two candidates for what it might mean for a Lévy process to be computable, and show that they are essentially equivalent. Along the way, we will effectivize Fact 4.1.29 (and this represents most of the work), proving that a càdlàg modification of a Lévy process not only exists, but is computable from a suitable representation of the process.

4.2.1 Stochastic computability and effective measurability

There are at least two plausible candidates for what it might mean for a Lévy process to be computable: what we call stochastic computability and effective measurability. We state the definitions for a general stochastic process, although effective measurability only makes sense if the process has a càdlàg modification (which all Lévy processes do).

We need to assume that the sample space is a computable probability space, say $\langle \Omega, \mathbb{P} \rangle$ (with the Borel σ-algebra, of course). As for the state space, we already assumed—for simplicity, and without any real loss of generality—that our Lévy processes were one-dimensional, i.e., $\mathbb{R}$-valued. So let $X = \{X_t : t \geq 0\}$ be an $\mathbb{R}$-valued stochastic process on Ω.

Definition 4.2.1. X is said to be *stochastically computable* if the function $t \mapsto X_t$ is computable of type $[0, \infty) \to L^0(\Omega, \mathbb{P})$.

Remark 4.2.2. If a modification Y of X is stochastically computable, then X is stochastically computable, as $X_t = Y_t$ a.s. $[\mathbb{P}]$ means $d_{\mathbb{P}}(X_t, Y_t) = 0$.

In light of Fact 4.1.15, stochastic computability is the effective version of stochastic continuity. It is also a very reasonable assumption in the context of computable analysis. The proof of Fact 4.1.15 was effective, so we have

Proposition 4.2.3. Let X be stochastically computable. Let $t \geq 0$ and $\epsilon > 0$. Uniformly from names for t and ϵ, we can compute a $\delta = \delta(t, \epsilon) > 0$ such that for all $s \geq 0$,

$$|t - s| < \delta \implies \mathbb{P}\{|X_t - X_s| \geq \epsilon\} < \epsilon. \tag{4.2.1}$$

Moreover, if we restrict t to a closed and bounded interval $[0, T]$ (and we have a name for T), then $\delta = \delta_T(\epsilon)$ can be made independent of t.

Proposition 4.2.4. Assume that X is a Lévy process. If X_1 is $L^0(\mathbb{P})$-computable, then the Lévy symbol $\eta : \mathbb{R} \to \mathbb{C}$ of X (see Fact 4.1.23) is computable. In particular, this holds if X is stochastically computable.

Proof. By Fact 2.3.48, $\exp(iuX_1)$ is $L^0(\mathbb{P})$-computable uniformly from a name for $u \in \mathbb{R}$. Since $\exp(iuX_1)$ is bounded by 1 in absolute value, it is uniformly $L^1(\mathbb{P})$-computable by Fact 2.3.55.

Fact 2.3.53 generalizes to $\mathbb{C}$-valued functions in a straightforward way. It follows that $\mathbb{P}[\exp(iuX_1)] \in \mathbb{C}$ is computable. Setting

$$z = \mathbb{P}[\exp(iuX_1)] = r\exp(i\theta),$$

we can compute $\eta(u) = \ln z = \ln r + i\theta$. The branch cut in $\mathbb{C}$ is not an issue, because η is continuous (Fact 4.1.23). This completes the proof. □

Not every Lévy process is integrable, i.e., it is not necessarily the case that for all $t \geq 0$, $\mathbb{P}|X_t| < \infty$. We still get some use out of an effective analogue of integrability which is stronger than stochastic computability:

Definition 4.2.5. X is said to be L^1-*computable* if the function $t \mapsto X_t$ is computable of type $[0,\infty) \to L^1(\Omega,\mathbb{P})$.

Proposition 4.2.6. If X is L^1-computable, X is stochastically computable.

Proof. Just apply Fact 2.3.30 uniformly. □

Remark 4.2.7. Similarly, we may define L^2-*computability* for stochastic processes. Then L^2-computability implies L^1-computability.

To define effective measurability, our second notion of computability for stochastic processes, we need to assume that X is càdlàg. This is reasonable, since every Lévy process has a càdlàg modification (Fact 4.1.29).

Definition 4.2.8. Assume that X is càdlàg. Let $T > 0$. We denote by $X|T$ the function $\omega \mapsto X(\cdot,\omega)$ of type $\Omega \to \mathbf{D}[0,T]$, whose values are the sample paths of X restricted to the time interval $[0,T]$.

Of course, if X happens to be continuous, then $X|T : \Omega \to \mathbf{C}[0,T]$.

Remark 4.2.9. By Fact 3.3.24, $X|T : \Omega \to \mathbf{D}[0,T]$ is measurable.

Definition 4.2.10. Assume that X is càdlàg. X is said to be *effectively measurable* (in Skorokhod space, $\mathbf{D}$) if $X|T$ is $L^0(\mathbb{P})$-computable of type $\Omega \to \mathbf{D}[0,T]$, uniformly from T, for all dyadic rational $T > 0$.

Remark 4.2.11. We assumed that T was a dyadic rational for convenience. In fact, without any real loss of generality, we could assume that $T \in \mathbb{N}$.

Remark 4.2.12. There is no reason in general why time should be bounded for a Lévy process, and indeed in the classical theory time typically ranges over $[0, \infty)$. In Skorokhod space, however, time is bounded. The reader can see in Definition 4.2.10 how we have resolved this.

Effective measurability implies a relatively weak form of computability:

Definition 4.2.13. Assume that X is càdlàg. X is said to have a *computable distribution* (on Skorokhod space, $\mathbf{D}$) if $\mathbb{P}_{X|T}$ is computable in $\mathcal{M}_1(\mathbf{D}[0,T])$, uniformly from T, for all dyadic rational $T > 0$.

Proposition 4.2.14. Let X be càdlàg. If X is effectively measurable in $\mathbf{D}$, then X has a computable distribution on $\mathbf{D}$.

Proof. The image of $\mathbb{P}$ along $X|T$ is computable by Fact 2.3.46. □

Stochastic computability and effective measurability are closely related for Lévy processes. First of all, it is not difficult to prove

Proposition 4.2.15. Let X be a Lévy process with càdlàg sample paths. If X is effectively measurable, then X is stochastically computable.

Proof. The following is uniform from a name for $t \geq 0$. Fix a dyadic rational $T > t$. By assumption, $X|T$ is $L^0(\mathbb{P})$-computable of type $\Omega \to \mathbf{D}[0,T]$. By Fact 4.1.31 and Proposition 3.3.27, the evaluation mapping

$$\mathrm{eval} : [0,T] \to L^0(\mathbf{D}[0,T], \mathbb{P}_X) \tag{4.2.2}$$

is computable. Let $W_t(\omega) = \mathrm{eval}(t)(X(\cdot,\omega))$. By Fact 2.3.48, W_t is $L^0(\mathbb{P})$-computable. But $W_t = X_t$ a.s. $[\mathbb{P}]$, which completes the proof. □

By effectivizing Fact 4.1.29, we obtain a converse to Proposition 4.2.15:

Theorem 4.2.16. Every stochastically computable Lévy process X has an effectively measurable modification $\widetilde{X}$.

Actually, our proof will construct a càdlàg modification of X from the function $t \mapsto X_t$. In any case, combining this result with Propositions 4.2.15, 4.2.6, and 4.2.14 yields an important equivalence:

Corollary 4.2.17. Let X be a Lévy process. The following are equivalent:

(i) X is stochastically computable

(ii) X is effectively measurable up to a modification

Both (i) and (ii) hold if X is L^1-computable. Either (i) or (ii) implies that X (or a càdlàg modification thereof) has a computable distribution.

But before we prove Theorem 4.2.16, we should mention an application to continuous stochastic processes (for example, Brownian motion).

As usual, $X = \{X_t : t \geq 0\}$ denotes an $\mathbb{R}$-valued stochastic process on Ω.

Definition 4.2.18. Assume that X is continuous. X is said to be *effectively measurable* in $\mathbf{C}$ if $X|T$ is $L^0(\mathbb{P})$-computable of type $\Omega \to \mathbf{C}[0,T]$, uniformly from T, for all dyadic rational $T > 0$.

(Compare Definition 4.2.10.) Now, as we know, $\mathbf{C}[0,T]$ is a topological subspace of $\mathbf{D}[0,T]$ for all $T > 0$ (Fact 3.1.28). Better yet, we have

Proposition 4.2.19. Assume that X is continuous. Then X is effectively measurable in $\mathbf{C}$ if and only if X is effectively measurable in $\mathbf{D}$.

Proof. Just apply Corollary 3.3.3. □

Basically, if we know our process has an effectively measurable càdlàg modification, which we know classically is continuous, then Proposition 4.2.19 gives us an effectively measurable continuous modification for free.

The proof of Theorem 4.2.16 will take up the remainder of this section.

4.2.2 Computably càdlàg Lévy processes

Our proof of Theorem 4.2.16 uses several auxiliary results and is similar to the proof of Theorem 5.2.15, which in turn adapts a proof in Chan [19].

As a preliminary, the proof of Lemma 2.3.70 may be adapted to prove a technical result of no immediately obvious relevance:

Lemma 4.2.20. Let $a < b$ be computable reals. Let $(\mathbb{P}_n)_{n\in\mathbb{N}}$ be a sequence of uniformly computable probability measures on $\mathbb{R}$. We can compute an increasing, computably Cauchy sequence $(a_k)_{k\in\mathbb{N}}$ in (a,b) such that for $i \neq j$, the one-element sets $\{a_i\}$ and $\{a_i - a_j\}$ are $\mathbb{P}_n$-null for all $n \in \mathbb{N}$.

Proof. Similar to the proof of Lemma 2.3.70, for all $n, j \in \mathbb{N}$, let

$$U_{n,j} = \{a > 0 : \mu_n[-a,a] < \mu_n(-a,a) + 1/j\}.$$

$U_{n,j}$ is Σ^0_1 (uniformly from n and j). Obviously, for all $a > 0$,

$$\mu_n[-a,a] - \mu_n(-a,a) = \mu_n\{\pm a\}.$$

$\{\pm a\}$ and $\{\pm a_1\}$ are disjoint for $a \neq a_1$ and μ_n is finite, so $\mu_n\{\pm a\} \geq 1/j$ for only finitely many a. In particular, $U_{n,j}$ is dense in $[0,\infty)$. Hence

$$A_0 = \bigcap_{n,j\in\mathbb{N}} U_{n,j}$$

is a Π^0_2 set dense in $[0,\infty)$. Use the effective Baire category theorem to compute a point $a_0 \in \pm A_0 \cap (a,b)$. By construction, $\{a_0\}$ is μ_n-null for all $n \in \mathbb{N}$. Now suppose we have computed a_i for all $i < k$. For all $n, j \in \mathbb{N}$, let

$$V^k_{n,j} = \{a > 0 : (\forall i < k)\ \mu_n[a_i - a, a_i + a] < \mu_n(a_i - a, a_i + a) + 1/j\}.$$

Evidently, if $a \in V^k_{n,j}$, then $\mu_n\{a_i \pm a\} < 1/j$ for all $i < k$. Just as above, $V^k_{n,j}$ is a Σ^0_1 set dense in $[0,\infty)$ (uniformly from n, j, and k). Hence

$$A_k = \bigcap_{n,j\in\mathbb{N}} U_{n,j} \cap \bigcap_{n,j\in\mathbb{N}} V^k_{n,j}$$

is a Π^0_2 set dense in $[0,\infty)$. Use the effective Baire category theorem again to compute a point $a_k \in \pm A_k \cap (a_{k-1}, b) \cap (a_{k-1}, a_{k-1} + 2^{-k})$. By construction, both $\{a_k\}$ and $\{a_k - a_i\}$, for all $i < k$, are μ_n-null for all $n \in \mathbb{N}$.

Thus we construct $(a_k)_{k\in\mathbb{N}}$ by induction. By forcing $a_k - a_{k-1} < 2^{-k}$, we guarantee that $(a_k)_{k\in\mathbb{N}}$ is computably Cauchy. This completes the proof. □

Now let $X = \{X_t : t \geq 0\}$ be a stochastically computable Lévy process on a computable probability space $\langle \Omega, \mathbb{P}\rangle$. For the proofs to follow, we need a countable dense subset of times, for which we will use the non-negative dyadic rationals: let $D = \{q^\circ_i : i \in \mathbb{N}\}$ enumerate them (without repetition).

Our second technical result is

Lemma 4.2.21. Let $a < b$ be computable reals again. We can compute an increasing, computably Cauchy sequence $(a_k)_{k\in\mathbb{N}}$ in (a,b) such that for all $i \neq j$, for all (dyadic rational) $t \in D$,

$$\mathbb{P}\{|X_t - X_0| = a_i\} = \mathbb{P}\{|X_t - X_0| = a_i - a_j\} = 0. \tag{4.2.3}$$

Proof. Apply Lemma 4.2.20 to the sequence $(\mu_n)_{n\in\mathbb{N}}$, where μ_n is the image measure of $\mathbb{P}$ along $|X_t - X_0|$ and $t = q^\circ_n$. By (uniform) Fact 2.3.46, those measures are uniformly computable. □

Remark 4.2.22. It follows that $\{|X_t - X_0| > a_i\}$ and $\{|X_t - X_0| > a_i - a_j\}$ are $L^0(\mathbb{P})$-computable uniformly from $i, j \in \mathbb{N}$ and $t \in D$.

Remark 4.2.23. Lemma 4.2.21 holds for any stochastically computable process X, not just for a Lévy process.

Now we use Lemma 4.2.21 to prove two more lemmas. Both concern the regularity of the sample paths of X, i.e., how much they may vary.

Notation 4.2.24. For all $r, s \in D$, $\widehat{X}[r, r+s]$ denotes the random variable

$$\sup\{|X_t - X_r| : t \in [r, r+s] \cap D\}. \tag{4.2.4}$$

Our first regularity result is adapted from Lemma 5.3 in Chan [19].

Lemma 4.2.25. For all $r, s \in D$, for all $M > 0$, $\min\{\widehat{X}[r, r+s], M\}$ is $L^1(\mathbb{P})$-computable from r, s, and a name for M.

In the proof, we compute several objects used in later proofs.

Proof. We will prove the lemma for $r = 0$. It is easy to generalize to $r \in D$.

Fix $s \in D$. By Lemma 4.2.21, if $0 \le a < b$, we can compute an increasing, computably Cauchy sequence $(a_k)_{k\in\mathbb{N}}$ in (a, b) such that for all $t \in D$,

$$\mathbb{P}\{|X_t - X_0| = a_k\} = \mathbb{P}\{|X_t - X_0| = a_{k+1} - a_k\} = 0,$$

so that both $\{|X_t - X_0| > a_k\}$ and $\{|X_t - X_0| > a_{k+1} - a_k\}$ are $L^0(\mathbb{P})$-computable uniformly from k and t (see Remark 4.2.22).

For all $k \in \mathbb{N}$, let $\epsilon_k = (a_{k+1} - a_k)/2$. By Proposition 4.2.3, compute a δ_k such that $\mathbb{P}\{|X_t - X_0| \ge \epsilon_k\} < \epsilon_k$ whenever $t < \delta_k$. Then compute a strictly increasing sequence $(p_k)_{k\in\mathbb{N}} \subseteq \mathbb{N}$ such that $p_0 = 0$ and $2^{-p_k} s < \delta_k$ for all $k \ge 1$.

Divide $[0, s] \cap D$ into finite subsets as follows: for all $k \in \mathbb{N}$,

$$D_k = \{j 2^{-p_k} s : 0 \le j \le 2^{p_k}\}.$$

Then $\{0, s\} = D_0 \subseteq D_1 \subseteq \cdots$ and $\bigcup_{k\in\mathbb{N}} D_k = [0, s] \cap D$.

Compute, for each $t \in D$, the least $j \in \mathbb{N}$ such that $t \in D_j$, and call it $j(t)$. If $j(t) > 0$, compute the first time $r_t \in D_{j(t)-1}$ such that $r > t$.

Define a sequence $(A_k)_{k\in\mathbb{N}}$ of uniformly $L^0(\mathbb{P})$-computable sets by

$$A_k = \bigcap_{t\in D_k} \{|X_t - X_0| \le a_{j(t)}\}.$$

Then define a sequence $(T_k)_{k\in\mathbb{N}}$ of random variables (stopping times, in fact)

$$T_k = \sum_{t\in D_k} t \mathbf{1}_{C_k(t)} + s \mathbf{1}_{A_k}$$

where $C_k(t)$ is the uniformly $L^0(\mathbb{P})$-computable set

$$\{|X_t - X_0| > a_{j(t)} \text{ and } |X_u - X_0| \leq a_{j(u)} \text{ for all } u \in D_k \text{ with } u < t\}.$$

T_k represents the first time $t \in D_k$ when $|X_t - X_0| > a_{j(t)}$, or s if no such time exists. T_k is $L^1(\mathbb{P})$-computable uniformly from k.

For all $k \in \mathbb{N}$, $A_k \supseteq A_{k+1}$ by definition and

$$A_k \setminus A_{k+1} \subseteq \bigcup_{t \in D_{k+1}} \{T_{k+1} = t \text{ and } |X(r_t) - X(t)| > a_{k+1} - a_k\}$$

since $|X_t - X_0| > a_{j(t)} = a_{k+1}$ and $|X_{r_t} - X_0| \leq a_{j(r_t)} = a_k$. From this, we get

$$\begin{aligned} \mathbb{P}(A_k \setminus A_{k+1}) &\leq \sum_{t \in D_{k+1}} \mathbb{P}\{T_{k+1} = t \text{ and } |X(r_t) - X(t)| > a_{k+1} - a_k\} \\ &= \sum_{t \in D_{k+1}} \mathbb{P}\{T_{k+1} = t\} \cdot \mathbb{P}\{|X(r_t - t) - X(0)| > a_{k+1} - a_k\}, \end{aligned}$$

because X has independent, stationary increments. (This is the only place in the proof where we use the assumption that X is a Lévy process, and not merely stochastically computable.) And the last expression is bounded by

$$\sum_{t \in D_{k+1}} \mathbb{P}\{T_{k+1} = t\} \cdot (a_{k+1} - a_k) = a_{k+1} - a_k,$$

because $r_t - t < 2^{-p_k} s < \delta_k$ for all $t \in D_{k+1}$.

Since $\mathbb{P}(A_k \setminus A_{k+1}) \leq a_{k+1} - a_k$ and $\sum_{k \in \mathbb{N}} (a_{k+1} - a_k)$ converges (absolutely) at a computable rate, the event

$$B = \bigcap_{k \in \mathbb{N}} A_k = \bigcap_{t \in [0,s] \cap D} \{|X_t - X_0| \leq a_{j(t)}\}$$

is $L^0(\mathbb{P})$-computable, being a computable limit of $L^0(\mathbb{P})$-computable sets.

On B, $|X_t - X_0| < b$ for all $t \in [0, s] \cap D$, whereas on $\Omega \setminus B$, $|X_t - X_0| > a$ for some $t \in [0, s] \cap D$. (Recall that $a < a_k < b$ for all $k \in \mathbb{N}$.)

Now fix $M > 0$. Repeat the above construction, obtaining for all $n \in \mathbb{N}$ uniformly $L^0(\mathbb{P})$-computable sets $B(0), B(1), \ldots, B(2^n)$ such that

- on $B(k)$, $|X_t - X_0| < k2^{-n}M$ for all $t \in [0, s] \cap D$, and
- on $\Omega \setminus B(k)$, $|X_t - X_0| > (k-1)2^{-n}M$ for some $t \in [0, s] \cap D$.

Set $B(2^n + 1) = \mathbb{R}$ and define uniformly $L^1(\mathbb{P})$-computable random variables

$$Z_n = \sum_{k=0}^{2^n} (k+1)2^{-n} M \mathbf{1}_{B(k+1) \setminus B(k)}.$$

Note that $|Z_n - Z_{n+1}| \leq 2^{-n}M$ for all $n \in \mathbb{N}$. It follows that $Z = \lim Z_n$ is $L^1(\mathbb{P})$-computable.

Finally, it is easy to see that Z is the smaller of M and the supremum of $|X_t - X_0|$ over $t \in [0,s] \cap D$, which completes the proof. □

Our second regularity result is adapted from Lemma 5.4 in Chan [19].

Lemma 4.2.26. Let $r \in D$ and $\epsilon > 0$. Compute $\delta = \delta(r,\epsilon)$ as in Proposition 4.2.3 (uniformly from r and a name for ϵ). Then for all $s \in D$, for all $c > 2\epsilon$,

$$s < \delta \implies \mathbb{P}\{\widehat{X}[r, r+s] > c\} < 2\epsilon. \tag{4.2.5}$$

Proof. We will prove the lemma for $r = 0$. It is easy to generalize to $r \in D$.

First, of course, use Proposition 4.2.3 to compute δ. Next, fix $s \in D$, let $a = \epsilon$ and $b = 2\epsilon$, and compute the sequences $(a_k)_{k\in\mathbb{N}}$, $(p_k)_{k\in\mathbb{N}}$, $(D_k)_{k\in\mathbb{N}}$, and $(A_k)_{k\in\mathbb{N}}$ as in the proof of Lemma 4.2.25. For all $k \in \mathbb{N}$,

$$\mathbb{P}(A_k \setminus A_{k+1}) \leq a_{k+1} - a_k.$$

Moreover, if $s < \delta$, we have

$$\mathbb{P}(\Omega \setminus A_0) = \mathbb{P}\{|X_s - X_0| > a_0\} \leq \mathbb{P}\{|X_s - X_0| \geq \epsilon\} < \epsilon.$$

Therefore, if we let $B = \bigcap_k A_{k\in\mathbb{N}}$ (also as in the proof of Lemma 4.2.25),

$$\begin{aligned} \mathbb{P}(\Omega \setminus B) &= \mathbb{P}(\Omega \setminus A_0 \cup A_0 \setminus A_1 \cup A_1 \setminus A_2 \cup \cdots) \\ &< \epsilon + (a_1 - a_0) + (a_2 - a_1) + \cdots \\ &= \epsilon + \lim a_k - a_0 < 2\epsilon. \end{aligned}$$

Now, $B \subseteq \{\widehat{X}[0,s] \leq b\}$ by definition. Since $c > 2\epsilon = b$, $\{\widehat{X}[0,s] > c\} \subseteq \Omega \setminus B$, which completes the proof: $\mathbb{P}\{\widehat{X}[0,s] > c\} \leq \mathbb{P}(\Omega \setminus B) < 2\epsilon$. □

The next step is to use Lemmas 4.2.25 and 4.2.26 to prove the key result, adapted from Theorem 5.5 in Chan [19], about sample path regularity for stochastically computable Lévy processes on a given interval $[0,T]$.

Theorem 4.2.27. Given $0 \neq T \in D$, uniformly from a name for $0 < \epsilon < 1$, there exist $L^0(\mathbb{P})$-computable random variables

$$0 = U_0 = V_0 \leq U_1 \leq V_1 \leq \cdots \leq U_n \leq V_n, \tag{4.2.6}$$

and an $L^0(\mathbb{P})$-computable event G such that

(i) $\mathbb{P}(G \cup \{U_n \leq T\}) < 2\epsilon$, and

(ii) on $\Omega \setminus G$, $|X(t) - X(V_i)| < 2\epsilon$ whenever $U_i < t < U_{i+1}$ for some i.

Proof. First, compute $\delta = \delta_T(\epsilon/2)$ as in Proposition 4.2.3. Then compute a $K \in \mathbb{N}$ such that $K > 2$ and

$$K\delta(\epsilon/2)(1-\epsilon) > 2s,$$

and an $m \in \mathbb{N}$ such that $(K-2)^m/(K-1)^m < \epsilon$. Let $n = mK + 1$.

Let $a = \epsilon$ and $b = \epsilon + \epsilon/n$. Compute the sequences $(a_k)_{k\in\mathbb{N}}$, $(p_k)_{k\in\mathbb{N}}$, $(D_k)_{k\in\mathbb{N}}$, and $(j(t))_{t\in D}$ as in the proof of Lemma 4.2.25, but for the interval $[0,T]$ rather than $[0,s]$. For all $k \in \mathbb{N}$ and $s \in D$, let

$$Z_k(s) = \min\{2, \sup\{|X_t - X_s| : t \in [s, s + 2^{-p_k}T] \cap D\}\}.$$

Then $Z_k(s) \sim Z_k(0)$ is $L^1(\mathbb{P})$-computable (uniformly from k and s).

Next, for all $k \in \mathbb{N}$ and $s \in D$, let $\tau_k(s)$ be the least $t \in D_k$ such that $t \geq s$ and $|X(t) - X(s)| > a_{j(t)}$, or T if no such time exists. Then $\tau_k(s)$ is $L^1(\mathbb{P})$-computable as well (see the proof of Lemma 4.2.25).

Generalizing Lemma 4.2.21, compute a sequence $(a'_k)_{k\in\mathbb{N}}$ such that

- $a_k < a'_k < a_{k+1}$ for all k, and
- $\{Z_k(s) > a'_{k+1} - a_k\}$ is $L^0(\mathbb{P})$-computable for all k and s.

Let $B_k(s) = \{Z_k(\tau_{k+1}(s)) > a'_{k+1} - a_k\}$. $B_k(s)$ is $L^0(\mathbb{P})$-computable for all k and s, and

$$\mathbb{P}(B_k(s)) = \mathbb{P}\{Z_k(0) > a'_{k+1} - a_k\}.$$

Since $2^{-p_k}s < \delta((a_{k+1}-a_k)/2)$ and $a'_{k+1} - a_k > a_{k+1} - a_k$, Lemma 4.2.26 implies that $\mathbb{P}(B_k(s)) < a_{k+1} - a_k$. Hence $\sum_{k\in\mathbb{N}} \mathbb{P}(B_k(s))$ converges computably, so that $B(s) = \bigcup_{k\in\mathbb{N}} B_k(s)$ is $L^0(\mathbb{P})$-computable for all s. Furthermore,

$$\mathbb{P}(B(s)) < \sum_{k\in\mathbb{N}} (a_{k+1} - a_k) = \lim a_k - a_0 < b - a = \epsilon/n.$$

Note that $(\tau_k(s))_{k\in\mathbb{N}}$ is non-increasing and, for all $k \in \mathbb{N}$,

$$\{\tau_k(s) - \tau_{k+1}(s) > 2^{-p_k}T\} \subseteq B_k(s).$$

In particular, $\mathbb{P}\{\tau_k(s) - \tau_{k+1}(s) > 2^{-p_k}T\} \leq \mathbb{P}(B_k(s))$, and it follows that $\tau(s) = \lim_{k\to\infty} \tau_k(s)$ is $L^0(\mathbb{P})$-computable for all s.

Let $U_0 = V_0 = 0$. Then let $U_1 = \tau(0)$, $V_1 = \tau_1(0)$, and $G_1 = B(0)$. Now we repeat the entire construction starting at time $s = V_1$. By induction, after repeating the construction n times, we have, for all $i < n$,

- $V_{i+1} = V_i + \tau_1(V_i)$,
- $U_{i+1} = V_i + \tau(V_i)$, and
- $G_{i+1} = B(V_i)$.

Then let $G = \bigcup_{i \le n} G_i$. On $\Omega \setminus G = \bigcap_{i \le n} \Omega \setminus B(V_i)$, we have

$$|X(r) - X(\tau_{k+1}(V_i))| \le a'_{k+1} - a_k < a_{k+2} - a_k$$

whenever $k \ge 1$, $1 \le i \le n$, and r belongs to

$$[\tau_{k+1}(V_i), \tau_{k+1}(V_i) + 2^{-p_k}T] \supseteq [\tau_{k+1}(V_i), \tau_k(V_i)].$$

Summing over all $k \ge 1$, we have, for all $r \in (\tau(V_i), \tau_1(V_i)]$,

$$|X(r) - X(\tau(V_i))| < 4(\lim_{k\to\infty} a_k - a_1) < 2\epsilon.$$

Hence, for all $t \in (U_{i+1}, V_{i+1}]$, by the definitions of U_{i+1} and V_{i+1},

$$|X(V_{i+1}) - X(t)| < 2\epsilon.$$

On the other hand, if $t \in [V_{i+1}, U_{i+2})$, write $t = V_{i+1} + r$ where

$$r < U_{i+2} - V_{i+1} = \tau(V_{i+1}) \le \tau_{j(r)}(V_{i+1}).$$

Here, $\tau_{j(r)}(V_{i+1})$ is the least $u \ge V_{i+1}$ such that $u \in D_{j(r)}$ and $X(u)$ is outside the $a_{j(r)}$-neighborhood of $X(V_{i+1})$. Hence

$$|X(t) - X(V_{i+1})| = |X(V_{i+1} + r) - X(V_{i+1})| \le a_{j(r)} < \lim a_k \le b < 2\epsilon.$$

We have established claim (ii) in the theorem.

Now we need to show that $\mathbb{P}(G \cup \{U_n \le T\}) < 2\epsilon$. First, we have

$$\mathbb{P}(G_{i+1}) = \mathbb{P}(B(V_i)) < \epsilon/n,$$

so $\mathbb{P}(G) < \epsilon$. It remains only to show that $\mathbb{P}\{U_n \le T\} < \epsilon$.

For all $i < n$, for all $r \in D$,

$$\mathbb{P}(V_{i+1} - V_i) = \mathbb{P}[\tau_1(V_i)] \ge r\mathbb{P}\{\tau_1(V_i) \ge r\} \ge r\mathbb{P}\{\widehat{X}[V_i, V_i + r] \le a_1\}.$$

If moreover $r > \delta(\epsilon/2)$, then since $a = \epsilon > a_1$, Lemma 4.2.26 gives

$$\mathbb{P}\{\widehat{X}[V_i, V_i + r] > a_1\} \le \epsilon$$

(see Notation 4.2.24). Therefore, $\mathbb{P}(V_{i+1} - V_i) \geq r(1-\epsilon)$.

Summing over all $i < K$,

$$\mathbb{P}(V_K) \geq K\delta(\epsilon/2)(1-\epsilon) > 2T.$$

And since $0 \leq V_K \leq KT$, letting $q = \mathbb{P}\{V_K \leq T\}$, we have

$$2T < \mathbb{P}(V_K) \leq (1-q)KT + qT,$$

which, by a direct calculation, implies that $q < (K-2)/(K-1)$.

Finally, we put a bound on $\mathbb{P}\{U_n \leq T\}$:

$$\begin{aligned}
\mathbb{P}\{U_n \leq T\} &\leq \mathbb{P}\{V_{n-1} \leq T\} = \mathbb{P}\{V_{mK} \leq T\} \\
&\leq \mathbb{P}\{V_K \leq T,\ V_{2K} - V_K \leq T,\ \ldots,\ V_{mK} - V_{(m-1)K} \leq T\} \\
&= \mathbb{P}\{V_K \leq T,\ V_{2K} - V_K \leq T,\ \ldots\} \cdot \mathbb{P}\{V_K \leq T\} \\
&\leq \mathbb{P}\{V_K \leq T,\ V_{2K} - V_K \leq T,\ \ldots\} \cdot (K-2)/(K-1)
\end{aligned}$$

By induction, $\mathbb{P}\{U_n \leq T\} \leq (K-2)^m/(K-1)^m < \epsilon$, as required to establish claim (i) and complete the proof. □

We are now ready to prove Theorem 4.2.16. First, we define what we know classically to be a càdlàg modification of $X = \{X_t : t \geq 0\}$.

Definition 4.2.28. Let $\widetilde{X}$ extend $\{X_t : t \in D\}$ to $t \in [0, \infty)$ according to

$$\widetilde{X}_t(\omega) = \lim_{r \downarrow t} X_r(\omega) \tag{4.2.7}$$

where r is restricted to D, whenever the limit exists; and 0 otherwise.

As usual, see Applebaum [3] for the classical details, particularly

Fact 4.2.29. $\widetilde{X}$ is a càdlàg modification of X.

Remark 4.2.30. Since $\widetilde{X}$ is a modification of X and X is a Lévy process, Fact 4.1.28 implies (classically) that $\widetilde{X}$ itself is a Lévy process.

Now Theorem 4.2.16 is just the sum of Fact 4.2.29 and the following

Proposition 4.2.31. $\widetilde{X}$ is effectively measurable.

Proof. Let $0 \neq T \in D$ and $0 < \epsilon < 1$. Use Theorem 4.2.27 to define the $L^0(\mathbb{P})$-computable random variables

$$0 = U_0 = V_0 \leq U_1 \leq V_1 \leq \cdots \leq U_n \leq V_n.$$

Then define a random variable $Z : \Omega \to \mathbf{D}[0,T]$ by

$$Z(\cdot)(t) = X(V_i) \text{ if } t \in [U_i, U_{i+1}) \text{ for some } i < n,$$

and $Z(\cdot)(t) = 0$ by convention if $t > U_n$. By Theorem 3.3.26, Z is $L^0(\mathbb{P})$-computable (uniformly from a name for ϵ).

By Theorem 4.2.27 and the right-continuity of $\widetilde{X}$,

$$\mathbb{P}\{d_J(Z, \widetilde{X}) > 4\epsilon\} \leq \mathbb{P}\{d_\infty(Z, \widetilde{X}) > 4\epsilon\} < 2\epsilon,$$

which implies that $d_{\mathbb{P}}(Z, \widetilde{X}) < 2\epsilon + 4\epsilon$. Since ϵ was arbitrary, we have shown that $\widetilde{X}$ is $L^0(\mathbb{P})$-computable as a function from Ω to $\mathbf{D}[0,T]$. Since T was arbitrary, this completes the proof. □

4.3 Effective Lévy-Itô decomposition

In this section, we effectivize the Lévy-Itô decomposition (Fact 4.1.43).

4.3.1 Intensity measure and Poisson integral

Let $\langle \Omega, \mathbb{P} \rangle$ be a computable probability space and let $X = \{X_t : t \geq 0\}$ be an $\mathbb{R}$-valued Lévy process on Ω. Now, in this context (computable analysis), stochastic computability is a very reasonable assumption to make about a Lévy process, given that classically it is stochastically continuous. But in light of Theorem 4.2.16, we can even restrict our attention to effectively measurable (càdlàg) Lévy processes without any real loss of generality.

The reader should recall the following definitions and notation:

(i) the intensity measure ν (Notation 4.1.27)

(ii) the Poisson random measure N_t (Definition 4.1.32)

(iii) the Poisson integral $\int_A f(x) N_t(dx)$ (Definition 4.1.38)

(iv) the compensated Poisson integral $\int_A f(x) \widetilde{N}_t(dx)$ (Definition 4.1.41)

(v) the sum of large jumps $\int_{|x| \geq \epsilon} x N_t(dx)$ (Fact 4.1.43)

(vi) the compensated sum of small jumps $\int_{|x| < \epsilon} x \widetilde{N}_t(dx)$ (Fact 4.1.43 again)

All of these are relevant to the Lévy-Itô decomposition and will have to be effectivized somehow, in this section or the next. We will start with the intensity measure, the Poisson random measure, and the Poisson integral.

First of all, we can forget about computing the intensity measure:

Remark 4.3.1. The intensity measure may not be finite, so in general it cannot be computable even in the more general sense of Remark 2.3.5.

Still, the intensity measure has many computable properties, and we identify some of these in the following two lemmas.

Lemma 4.3.2. There exists a dense sequence $(\alpha_k)_{k\in\mathbb{N}} \subseteq \mathbb{R}$ of uniformly computable continuity points for ν, so that $\nu\{\alpha_k : k \in \mathbb{N}\} = 0$.

Proof. In the following, ϵ ranges over $\mathbb{Q}$ and t ranges over $[0,\infty)$.

By Fact 2.3.17 (computable Baire category theorem), it is enough to show that the set A of continuity points for ν is a dense Π^0_2 subset of $\mathbb{R}$.

If $x \in \mathbb{R} \setminus A$, i.e., if $x \neq 0$ is an atom of ν, then Fact 4.1.37 implies that X exhibits jump discontinuities $\Delta X(t) = x$ according to a Poisson process with intensity $r = \nu(\{x\}) > 0$. Clearly, therefore, $\mathbb{R} \setminus A$ cannot contain an open set, so A is dense (classically).

We just need to show that $\mathbb{R} \setminus A$ is Σ^0_2. By definition,

$$x \in \mathbb{R} \setminus A \iff (\exists \epsilon > 0)\ \mathbb{P}_X\{(\exists t \le 1)\ \mathcal{J}_f(t) = x\} \ge \epsilon.$$

By Proposition 3.3.18, $\{f : (\exists t \le 1)\ \mathcal{J}_f(t) = x\} \subseteq \mathbf{D}[0,1]$ is Π^0_1 uniformly from a name for $0 \neq x \in \mathbb{R}$. It follows that $\mathbb{P}_X\{f : (\exists t \le 1)\ \mathcal{J}_f(t) = x\}$ is upper semicomputable. Fixing ϵ, the following property of x is Π^0_1:

$$\mathbb{P}_X\{f : (\exists t \le 1)\ \mathcal{J}_f(t) = x\} \ge \epsilon.$$

Adding an existential quantifier over $\epsilon > 0$, we deduce that $\mathbb{R} \setminus A$ is Σ^0_2, which completes the proof. □

The reader may have noticed that, technically, in the proof of Lemma 4.3.2, $\mathbb{P}_X$ should be replaced by $\mathbb{P}_{X|1}$. For simplicity, though, we adopt

Notation 4.3.3. When $T > 0$ is fixed or clear from the context, we use X to denote $X|T : \Omega \to \mathbf{D}[0,T]$, the random variable whose values are the sample paths of X restricted to the time interval $[0,T]$ (see Definition 4.2.8).

Now fix, as in Lemma 4.3.2, a dense sequence $(\alpha_k)_{k\in\mathbb{N}} \subseteq \mathbb{R}$ of uniformly computable continuity points for ν.

Lemma 4.3.4. The intensity measure $\nu(U) \in [0,\infty]$ of an open subset U of $\mathbb{R}$ is lower semicomputable uniformly from a name for $U \in \tau(\mathbb{R})$.

Proof. It is enough to prove the lemma for an open rational interval $U = (p, q)$. Since we can compute continuity points $\alpha_i > p$ and $\alpha_j < q$ arbitrarily close p and q respectively, it is actually enough to show that $\nu((\alpha_i, \alpha_j))$ is lower semicomputable (uniformly from $i, j \in \mathbb{N}$).

Since $\nu(\{0\}) = 0$, $0 \notin [\alpha_i, \alpha_j]$ without loss of generality. Hence

$$\nu((\alpha_i, \alpha_j)) = \mathbb{P}|\{t \leq 1 : \alpha_i < \Delta X(t) < \alpha_j\}| = \mathbb{P}J_1(X) = \mathbb{P}_X J_1,$$

where $J_1 : \mathbf{D}[0,1] \to \mathbb{N}$ is given by $J_1(f) = |\{t \leq 1 : \alpha_i < \mathcal{J}_f(t) < \alpha_j\}|$. Because (α_i, α_j) is bounded away from zero, $J_1(f) < \infty$ for all $f \in \mathbf{D}[0,1]$.

The result follows from Proposition 3.3.25: uniformly from $n \in \mathbb{N}$, the event $\{J_1 = n\}$ is $L^0(\mathbb{P}_X)$-computable, so $\mathbb{P}_X\{J_1 = n\}$ is computable. Hence

$$\nu((\alpha_i, \alpha_j)) = \mathbb{P}_X(J_1) = \sum_{n \in \mathbb{N}} n\mathbb{P}_X\{J_1 = n\}$$

is (uniformly) lower semicomputable. This completes the proof. □

Now we turn from the intensity measure to the Poisson random measure, which we recall counts jumps of certain sizes that occur by a certain time.

Proposition 4.3.5. Let $A = (\alpha_i, \alpha_j)$, (α_i, ∞), or $(-\infty, \alpha_j)$, where in any case α_i, α_j are continuity points for ν and A is bounded away from zero. Then the intensity $\nu(A)$ is computable, and the jump-counting process $\{N_t(A) : t \geq 0\}$ is an L^1-computable Poisson process (uniformly from i, j).

Proof. We prove the proposition for $A = (\alpha_i, \alpha_j)$. The other cases are exactly analogous. Similar to the proof of Lemma 4.3.4, we write

$$N_t(\cdot, A) = |\{s \leq t : \alpha_i < \Delta X(s) < \alpha_j\}| = J_t(X)$$

where $J_t : \mathbf{D}[0,t] \to \mathbb{N}$ is given by $J_t(f) = |\{s \leq t : \alpha_i < \mathcal{J}_f(s) < \alpha_j\}|$. Because A is bounded away from zero, $J_t(f) < \infty$ for all $f \in \mathbf{D}[0,t]$ and $t \geq 0$. Also,

$$\nu(A) = \mathbb{P}(N_1(\cdot, A)) = \mathbb{P}(J_1(X)) = \mathbb{P}_X(J_1).$$

The following is uniform from a name for t. By Proposition 3.3.25, J_t is $L^0(\mathbb{P}_X)$-computable, so $J_t(X)$ is $L^0(\mathbb{P})$-computable (Fact 2.3.48).

The characteristic function of $J_t(X)$ (Definition 4.1.5) is given by

$$\phi(z) = \phi_{J_t(X)}(z) = \mathbb{P}[\exp(\mathbf{i}zJ_t(X))].$$

Now, $\exp(\mathbf{i}zJ_t(X))$ is (uniformly) $L^0(\mathbb{P})$-computable, but also bounded, so it is $L^1(\mathbb{P})$-computable (Fact 2.3.55). That lets us compute $\phi : \mathbb{R} \to \mathbb{C}$.

Since $\{J_t(X) : t \geq 0\}$ is a Poisson process with intensity $r = \nu(A)$,

$$(\forall u \in \mathbb{R}) \quad \phi(z) = \exp(rt(e^{\mathbf{i}z} - 1)).$$

Hence we can compute r from ϕ (see proof of Proposition 4.2.4). We have established the first claim.

For the second claim, we already know that $N_t(\cdot, A) = J_t(X)$ is $L^0(\mathbb{P})$-computable, and we just showed that its expected value $\mathbb{P}(N_t(A)) = t\nu(A)$ is computable. Therefore, $N_t(A)$ is $L^1(\mathbb{P})$-computable (Fact 2.3.55). It is a Poisson process by Fact 4.1.35, and that completes the proof. □

Remark 4.3.6. Assuming that $A = (\alpha_i, \alpha_j)$ is bounded away from zero, we can treat ν as a computable measure, in the more general sense of Remark 2.3.5, on the computable metric subspace $\overline{A} = [\alpha_i, \alpha_j]$.

Finally, we effectivize the Poisson integral.

Proposition 4.3.7. As in the premises of Proposition 4.3.5, let $A = (\alpha_i, \alpha_j)$, (α_i, ∞), or $(-\infty, \alpha_j)$, where in any case α_i, α_j are continuity points for ν and A is bounded away from zero. Let $f : \mathbb{R} \to \mathbb{R}$ be computable. Then the Poisson integral $\left\{\int_A f(x) N_t(dx) : t \geq 0\right\}$ is a stochastically computable compound Poisson process (uniformly from i, j).

Proof. As usual, the following is uniform from a name for t.

The Poisson integral is given by

$$\int_A f(x) N_t(dx) = \sum_{x \in A} f(x) N_t(\{x\}) = \sum_{x \in A} f(x) J_t(X, x),$$

where $J_t : \mathbf{D}[0, t] \times A \to \mathbb{N}$ is given by $J_t(g, x) = |\{s \leq t : \Delta g(s) = x\}|$.

By Proposition 3.3.25, we can $\mathbb{P}_X$-a.s. compute the list of all the jumps of g belonging to A. f is computable, hence computably continuous, so $\sum_{x \in A} f(x) J_t(X, x)$ is $L^0(\mathbb{P})$-computable. It is a compound Poisson process by Fact 4.1.39, which completes the proof. □

We have now effectivized (as much as we need to) the intensity measure, the Poisson random measure, and the Poisson integral. These are the basic ingredients of the Lévy-Itô decomposition. Now we turn to the jumps, both large and small, and the deterministic drift.

4.3.2 Jumps and drift

We begin by identifying a special sequence of continuity points for ν.

Proposition 4.3.8. There exists a decreasing sequence $(\epsilon_k)_{k\in\mathbb{N}}$ of uniformly computable positive reals, converging to 0, such that $\nu\{\pm\epsilon_k : k \in \mathbb{N}\} = 0$.

Proof. See the proof of Lemma 4.3.2. □

So fix $(\epsilon_k)_{k\in\mathbb{N}}$. For all $m \in \mathbb{N}$, let $B_m = \{x : \epsilon_{m+1} \leq |x| < \epsilon_m\} \subseteq \mathbb{R}$. For all $n \in \mathbb{N}$, let $A_n = \bigcup_{m=0}^{n} B_m = \{x : \epsilon_{n+1} \leq |x| < \epsilon_0\}$. Finally, let $C = \{x : |x| \geq \epsilon_0\}$.

Proposition 4.3.9. Each of B_m, A_n, and C is a ν-continuity set bounded away from 0. Its interior is Σ^0_1 and its closure is Π^0_1 (uniformly).

Proof. This is clear from the definitions. □

Recall the Lévy-Itô decomposition (Fact 4.1.43). If we set $\epsilon = \epsilon_0$, the sum of large jumps becomes $\int_C xN_t(dx)$. Then Proposition 4.3.7 implies

Corollary 4.3.10. The sum of large jumps is stochastically computable.

We can say more about the better-behaved of the two jump processes:

Proposition 4.3.11. The sum of large jumps is an effectively measurable (càdlàg) compound Poisson process with computable intensity.

Proof. Classically, $\int_C xN_t(dx)$ is càdlàg and a compound Poisson process (see Fact 4.1.39). By Corollary 4.3.10, it is stochastically continuous, so to prove it is effectively measurable, just apply Theorem 4.2.16.

To be precise, we can use Definition 4.2.28 to construct an effectively measurable modification of the sum of large jumps, and classically any two càdlàg modifications of a process are indistinguishable (Fact 4.1.13).

The intensity $r = \nu(C)$ is computable by Proposition 4.3.5. □

As for the sum of small jumps, it has to be compensated. Now, the compensator itself (Definition 4.1.40) is not actually used for this purpose, but it is worth mentioning that Proposition 4.3.5 implies

Corollary 4.3.12. Let $A = (\alpha_i, \alpha_j)$, (α_i, ∞), or $(-\infty, \alpha_j)$, where in any case α_i, α_j are continuity points for ν and A is bounded away from zero. Then the compensator $\widetilde{N}_t(A)$ is L^1-computable (uniformly from i, j).

More to the point, the small jumps are certainly bounded, and in general a Lévy process with bounded jumps is easier to control:

Proposition 4.3.13. Let $Y = \{Y_t : t \geq 0\}$ be a stochastically computable Lévy process. Suppose we are given a $K \in \mathbb{N}$ such that

$$\sup_{t \geq 0}\{|\Delta Y(t)|\} < K. \tag{4.3.1}$$

Then Y is L^2-computable (see Remark 4.2.7).

Proof. First, use Lemma 4.2.21 to compute a new bound $a > K$ such that $\mathbb{P}\{|Y_t| = a\} = 0$. Other than that, we are just effectivizing a classical proof from Applebaum [3] (specifically, Theorem 2.4.7).

We define, by induction, a sequence $(T_n)_{n \in \mathbb{N}}$ of stopping times. First, $T_0 = \inf\{t \geq 0 : |Y(t)| > a\}$. Then, for all $n \in \mathbb{N}$,

$$T_{n+1} = \inf\{t \geq T_n : |Y(t) - Y(T_n)| > a\}.$$

By Theorem 4.2.16, $T_0 = \inf\{t \in D : |Y(t)| > a\}$ a.s. $[\mathbb{P}]$, where D denotes the non-negative dyadic rationals. And classically we have

$$\mathbb{P}(e^{-T_n}) = [\mathbb{P}(e^{-T_0})]^n.$$

Denote $\inf\{t \in D : |Y(t)| > a\}$ by τ_0. The proof of Theorem 4.2.27 showed that $\tau_0 \wedge m$ is $L^0(\mathbb{P})$-computable uniformly from $m \in \mathbb{N}$. $e^{-(\tau_0 \wedge m)}$ is uniformly bounded, hence $L^1(\mathbb{P})$-computable (Fact 2.3.55). It follows that

$$e^{-\tau_0} = \lim_{m \to \infty} e^{-(\tau_0 \wedge m)}$$

is also $L^1(\mathbb{P})$-computable. Since $e^{-\tau_0}$ equals e^{-T_0} a.s. $[\mathbb{P}]$, the real number

$$\gamma = \mathbb{P}(e^{-\tau_0}) = \mathbb{P}(e^{-T_0})$$

is computable. Classically, we have $0 < \gamma < 1$, and for all $t \geq 0$ and $n \in \mathbb{N}$,

$$\mathbb{P}(\{|Y(t)| \geq 2na\}) \leq e^t \gamma^n,$$

a computable bound. Then, still classically, for all $p \in \mathbb{N}$,

$$\mathbb{P}(|Y(t)|^p \cdot \mathbf{1}_{|Y(t)| \geq 2na}) \leq (2a)^p e^t \sum_{j=n}^{\infty} (j+1)^p \gamma^j,$$

which converges absolutely at a computable rate.

Now, since $Y(t)$ is $L^0(\mathbb{P})$-computable (uniformly from a name for t) by hypothesis, this computable bound implies that $Y(t)$ is $L^2(\mathbb{P})$-computable, which completes the proof. □

We already took the trouble of separating out the large jumps of X, so we can apply Proposition 4.3.13 to whatever is left:

Corollary 4.3.14. For all $n \in \mathbb{N}$, $Y_n(t) = X(t) - \int_{|x|\geq\epsilon_n} xN_t(dx)$ defines an L^2-computable, effectively measurable Lévy process (uniformly from n).

Proof. Classically, Y_n is càdlàg and a Lévy process. By Proposition 4.3.7, $\left\{\int_{|x|\geq\epsilon_n} xN_t(dx) : t \geq 0\right\}$ is stochastically computable (uniformly). Hence the difference Y_n is also stochastically computable. By Proposition 4.3.13, Y_n is L^2-computable. Finally, for effective measurability, just apply Theorem 4.2.16 as in the proof of Proposition 4.3.11. □

Incidentally, this effectivizes another classical fact about Lévy processes:

Corollary 4.3.15. For all $n \in \mathbb{N}$, $Y_n(t) - \mathbb{P}(Y_n(t))$ defines an L^2-computable, effectively measurable martingale (uniformly from n).

But we are mainly interested in applying Corollary 4.3.14 to the drift:

Proposition 4.3.16. The drift b of X is computable.

Proof. We appropriate the relevant classical result from Applebaum [3]:

$$b = \mathbb{P}\left[X(1) - \int_C xN_1(dx)\right] = \mathbb{P}(Y_0(1)).$$

By Corollary 4.3.14, $\mathbb{P}(Y_0(1))$ is computable. □

With the sum of large jumps and the drift accounted for, we are left (classically, that is) with the sum of a Brownian motion and the small jumps. If we can pull out those jumps, we will have effectivized the decomposition.

We begin by effectivizing (as much as we need to) the compensated Poisson integral. (We effectivized the Poisson integral in Proposition 4.3.7.) As with Corollary 4.3.14, bounding the jumps is a key part of the proof.

Proposition 4.3.17. Let $A = (\alpha_i, \alpha_j)$ where α_i, α_j are continuity points for ν and A is both bounded away from zero and bounded. Let $f : \mathbb{R} \to [0, M]$ be computable and bounded. Then the compensated Poisson integral

$$\int_A f(x)\widetilde{N}_t(dx) = \int_A f(x)N_t(dx) - t\int_A f(x)\nu(dx)$$

defines an L^2-computable martingale (uniformly from i, j, M).

Proof. By Proposition 4.3.7, the Poisson integral $\left\{\int_A f(x)N_t(dx) : t \geq 0\right\}$ is stochastically computable. Since f and A are bounded, this process has bounded jumps, and Proposition 4.3.13 implies that it is $L^2(\mathbb{P})$-computable.

Classically, $\int_A f(x)\nu(dx)$ is just the expected value of $\int_A f(x)N_1(dx)$. Hence it is computable, which is enough to prove that the compensated Poisson integral is L^2-computable. Classically, the process is a martingale (Fact 4.1.42), and that completes the proof. □

That brings us to the compensated sum of small jumps.

Proposition 4.3.18. The compensated sum of small jumps

$$Z(t) = \int_{|x|<\epsilon_0} x\widetilde{N}_t(dx)$$

defines an L^2-computable, effectively measurable martingale.

Proof. Classically, for all $n \in \mathbb{N}$,

$$M_n(t) = \int_{A_n} x\widetilde{N}_t(dx) = \sum_{m=0}^{n} \int_{B_m} x\widetilde{N}_t(dx) = \int_{\epsilon_{n+1}\leq|x|<\epsilon_0} x\widetilde{N}_t(dx)$$

defines a martingale. By Proposition 4.3.17, it is L^2-computable.

Classically, for all $t \geq 0$, $\|M_n(t)\|_{L^2}$ is increasing in n and bounded above, and $M_n(t)$ converges to the martingale $Z(t)$ in L^2-norm; see Theorem 2.4.11 in Applebaum [3]. We can then use a classical bound to ensure that the rate of convergence is computable; see, for example, Dia [27].

Classically, Z is a purely discontinuous martingale (Fact 4.1.42). Finally, for effective measurability, just apply Theorem 4.2.16 as in the proof of Proposition 4.3.11. □

We can now state the *effective Lévy-Itô decomposition*.

Theorem 4.3.19. Let X be an effectively measurable Lévy process. We can compute an $\epsilon > 0$ such that in the Lévy-Itô decomposition (Fact 4.1.43),

(i) the drift vector b is computable,

(ii) the sum of large jumps $\int_{|x|\geq\epsilon} xN_t(dx)$ defines an effectively measurable compound Poisson process with computable intensity,

(iii) the compensated sum of small jumps $\int_{|x|<\epsilon} x\widetilde{N}_t(dx)$ defines an L^2-computable, effectively measurable martingale, and

(iv) the Brownian motion W is effectively measurable in $\mathbf{C}$.

Proof. $\epsilon = \epsilon_0$ was given by Proposition 4.3.8. Claims (i), (ii), and (iii) are established in Propositions 4.3.16, 4.3.11, and 4.3.18, respectively.

Finally, if we subtract from X all the other components, which are stochastically computable, we are left with a Brownian motion W, which is therefore stochastically computable itself. Apply Theorem 4.2.16 as in the proof of Proposition 4.3.11 to show that W is effectively measurable in $\mathbf{D}$. Classically, the sample paths of W are continuous, so Proposition 4.2.19 establishes (iv) and completes the proof. □

The value we chose for ϵ is not in any way special. It is really just a positive real number such that $\nu(\{\pm\epsilon\}) = 0$. So it may be worth mentioning

Remark 4.3.20. Throughout the proofs in this section, we can replace ϵ_0 by ϵ_k for any $k \in \mathbb{N}$. Hence we have actually shown that we can compute a sequence $(\epsilon_k)_{k\in\mathbb{N}}$ of uniformly computable positive reals, converging to 0, such that (i)–(iv) hold for $\epsilon = \epsilon_k$, uniformly from k, for all $k \in \mathbb{N}$.

Indeed, the proof of Lemma 4.3.2 shows that we can compute a sequence $(\alpha_k)_{k\in\mathbb{N}}$ of uniformly computable reals that is dense in $(0, \infty)$ and such that (i)–(iv) hold for $\epsilon = \alpha_k$, uniformly from k, for all $k \in \mathbb{N}$.

Chapter 5

Effective theory of Feller processes

Every Lévy process is also a Feller process, and every Feller process also admits a càdlàg modification. In this chapter, we extend our effective theory of Lévy processes to the larger class of Feller processes in a way that is broadly consistent with the constructive theory of Chan [19, 20].[1]

A Feller process is a type of Markov process, which is to say a process with no memory: the distribution of its future state depends, not on its entire past, but only on its present state. A Feller process must also satisfy a continuity condition, which is enough to ensure that a càdlàg modification exists. All of this is also true of a Lévy process, of course.

One difference is that, in many cases, we would rather treat a Feller (or a Markov) process, not as a given collection of sample paths, but as a so-called transition function describing how the process, whatever it may be, is supposed to change from one time to another. We would then try to construct sample paths that are consistent with the transition function. For that reason alone, our effective theory of Feller processes looks quite different from the theory developed in Chapter 4.

In Section 5.1, we review the relevant classical theory of Feller processes. In Section 5.2, we effectivize its basic objects, including transition functions. Finally, in Section 5.3, we generalize Theorem 4.2.16, proving that a càdlàg modification of a Feller process not only exists, but is computable from a suitable representation of the process—which, being based on a transition function, is quite different from how we represented a Lévy process.

[1] In the constructive theory, Feller processes are described as Markov processes with strongly continuous Markov semigroups (which they are); the term "Feller" is not used.

5.1 Classical theory of Feller processes

In this section, we review basic concepts from the classical theory of Markov processes—to be precise, time-homogeneous Markov processes in continuous time—with an emphasis on Feller processes, of course. For details, the reader may consult Rogers and Williams [65] or Revuz and Yor [63].

5.1.1 Probability kernels

Probability kernels are fundamental to Markov process theory. They extend the concept of a conditional probability distribution.

Suppose the state x of a stochastic process is observed at time s. Where will the process be at a later time t? The answer is a probability distribution on the state space. If the process has no memory, that distribution depends only on x, s, and t. If s and t are fixed, we have a mapping from states to probability distributions: this is the concept of a probability kernel.

Definition 5.1.1. A *kernel* from a measurable space $\langle E, \mathscr{E}\rangle$ to another measurable space $\langle F, \mathscr{F}\rangle$ is a function $\kappa : E \times \mathscr{F} \to [0, \infty]$ such that

(i) for all $x \in E$, $B \mapsto \kappa(x, B)$ is a measure on $\langle F, \mathscr{F}\rangle$, and

(ii) for all $B \in \mathscr{F}$, $x \mapsto \kappa(x, B)$ is $\mathscr{E}$-measurable.

This chapter will introduce a great deal of notation, starting with

Notation 5.1.2. Let κ be a kernel. The measure $B \mapsto \kappa(x, B)$ is denoted by κ^x. The integral with respect to κ^x, assuming it exists, of a measurable function $f : F \to [-\infty, \infty]$ may be denoted by any of these:

$$\int \kappa(x, dy) f(y) = \int f(y) \kappa^x(dy) = \kappa f(x). \tag{5.1.1}$$

With this notation, for all $x \in E$, for all $B \in \mathscr{F}$, $\kappa^x(B) = \kappa \mathbf{1}_B(x)$.

Definition 5.1.3. A kernel κ from $\langle E, \mathscr{E}\rangle$ to $\langle F, \mathscr{F}\rangle$ is said to be a *sub-probability kernel* if $\kappa^x(F) \le 1$ for all $x \in E$. κ is said to be a *probability kernel* if $\kappa^x(F) = 1$ for all $x \in E$. In both cases, $0 \le \kappa \le 1$.

Remark 5.1.4. A sub-probability kernel describes a stochastic process that can disappear or die. Later we will formalize this with a "coffin state."

Definition 5.1.5. For a measurable space $\langle E, \mathscr{E}\rangle$, $\mathscr{E}^+$ denotes the cone of $\mathscr{E}$-measurable functions $E \to [0, \infty]$. $b\mathscr{E}$ denotes the Banach space of bounded $\mathscr{E}$-measurable functions $E \to \mathbb{R}$ under the sup-norm.

Let κ be a kernel from $\langle E, \mathscr{E}\rangle$ to $\langle F, \mathscr{F}\rangle$.

Fact 5.1.6. If $f \in \mathscr{F}^+$, then $(x \mapsto \kappa f(x)) \in \mathscr{E}^+$. Thus κ induces a function from $\mathscr{F}^+$ to $\mathscr{E}^+$, which is also denoted by κ.

Fact 5.1.7. Assume that κ is a sub-probability kernel. If $f \in b\mathscr{F}$, then $(x \mapsto \kappa f(x)) \in b\mathscr{E}$. Thus κ induces a linear operator from $b\mathscr{F}$ to $b\mathscr{E}$, which is still denoted by κ. This operator satisfies

$$0 \le f \le 1 \implies 0 \le \kappa f \le 1. \tag{5.1.2}$$

Moreover, for any sequence $(f_n)_{n\in\mathbb{N}} \subseteq b\mathscr{F}$,

$$f_n \downarrow 0 \text{ (pointwise)} \implies \kappa f_n \downarrow 0 \text{ (pointwise)}. \tag{5.1.3}$$

Finally, κ is a probability kernel if and only if $\kappa(1) = 1$.

Definition 5.1.8. A linear operator $\varphi : b\mathscr{F} \to b\mathscr{E}$ is said to be *Markov* if

(i) $0 \le f \le 1$ implies $0 \le \varphi f \le 1$, and

(ii) $f_n \downarrow 0$ implies $\varphi f_n \downarrow 0$ (pointwise in both cases).

A Markov operator is said to be *strictly Markov* if $\varphi(1) = 1$.

With this terminology, Fact 5.1.7 states that each sub-probability kernel κ induces a Markov operator, and this operator is strictly Markov if and only if κ is a probability kernel. We have a converse in

Fact 5.1.9. To each Markov operator $\varphi : b\mathscr{F} \to b\mathscr{E}$ there corresponds a unique sub-probability kernel κ inducing φ, namely the one given by

$$(\forall x \in E)\ (\forall B \in \mathscr{F}) \quad \kappa(x, B) = \varphi \mathbf{1}_B(x). \tag{5.1.4}$$

Again, κ is a probability kernel if and only if φ is strictly Markov.

We collect a few more basic facts about Markov operators in

Fact 5.1.10. Every Markov operator $\varphi : b\mathscr{F} \to b\mathscr{E}$ is

(i) *positive*, i.e., $f \ge 0$ implies $\varphi f \ge 0$,

(ii) *monotone*, i.e., $f \le g$ implies $\varphi f \le \varphi g$, and

(iii) a *contraction*, i.e., $|f| \le 1$ implies $|\varphi f| \le 1$.

Of course, every contraction is bounded: $\|\varphi\| \leq 1$, with equality if and only if φ is strictly Markov. Moreover, φ satisfies *Jensen's inequality*:

(iv) if $H : \mathbb{R} \to \mathbb{R}$ is convex and $H(0) = 0$, then $H(\varphi f) \leq \varphi(H(f))$.

The following example of a probability kernel is simple but important.

Definition 5.1.11. If $x \in E$, ε_x denotes the *unit mass* at x: the probability measure on $\langle E, \mathscr{E} \rangle$ defined by $\varepsilon_x(A) = 1$ if $x \in A$, and 0 otherwise.

Example 5.1.12. We define a probability kernel κ_0 from $\langle E, \mathscr{E} \rangle$ to $\langle E, \mathscr{E} \rangle$ by setting $\kappa_0(x, \cdot) = \varepsilon_x$ for all $x \in E$. κ_0 induces the identity operator on $b\mathscr{E}$.

We conclude by defining some important kernel-related operations.

Notation 5.1.13. To simplify our notation, we may identify a measurable space $\langle A, \mathscr{A} \rangle$ with the underlying set A. This should not be confusing.

First, we define a product operation on kernels.

Definition 5.1.14. If κ_1 is a kernel from E to F and κ_2 is a kernel from F to G, then we define the *product kernel* $\kappa_1\kappa_2$ from E to G by

$$(\forall x \in E)\ (\forall C \in \mathscr{G}) \quad (\kappa_1\kappa_2)(x, C) = \int \kappa_1(x, dy)\kappa_2(y, C). \tag{5.1.5}$$

Fact 5.1.15. The product of sub-probability kernels is a sub-probability kernel. The product of probability kernels is a probability kernel.

Fact 5.1.16. For sub-probability kernels κ_1 and κ_2, the Markov operator $b\mathscr{G} \to b\mathscr{E}$ induced by the product $\kappa_1\kappa_2$ is given by

$$(\kappa_1\kappa_2)f(x) = (\kappa_1(\kappa_2 f))(x) = \int \kappa_1(x, dy) \int \kappa_2(y, dz) f(z). \tag{5.1.6}$$

We can also integrate kernels, i.e., apply measures to them.

Definition 5.1.17. If κ is a sub-probability kernel from E to F and μ is a probability measure on E, then $\mu\kappa : b\mathscr{F} \to \mathbb{R}$ denotes the linear operator

$$(\mu\kappa)f = \int (\kappa f) d\mu = \int \mu(dx) \int \kappa(x, dy) f(y). \tag{5.1.7}$$

Remark 5.1.18. The operator $\mu\kappa$ is a special case of the product operator in Fact 5.1.16, as each measure μ on E induces a kernel from E to E which is constant in its first argument: $\langle x, A \rangle \mapsto \mu(A)$.

Notation 5.1.19. Naturally, a kernel from E to E is called a kernel *on* E.

5.1.2 Markov processes

With a kernel, we can describe how the state of a memoryless process changes from one time to another, provided that both times are fixed. We introduce time dependence (in continuous time) via transition functions.

Definition 5.1.20. A *transition function* on a measurable space E is a family $\{P_{s,t} : 0 \le s \le t\}$ of sub-probability kernels on E such that

$$(\forall s \le t \le u) \quad P_{s,u} = P_{s,t}P_{t,u}. \tag{5.1.8}$$

In other words, for all $x \in E$, for all $A \in \mathscr{E}$,

$$P_{s,u}(x, A) = \int P_{s,t}(x, dy)P_{t,u}(y, A). \tag{5.1.9}$$

This is known as the *Chapman-Kolmogorov equation.*

For our purposes, Definition 5.1.20 is unnecessarily general, because Feller processes are necessarily time-homogeneous in the following sense.

Definition 5.1.21. A transition function $\{P_{s,t} : 0 \le s \le t\}$ is said to be *time-homogeneous* or simply *homogeneous* if $P_{s,t} = P_{0,t-s}$ for all $0 \le s \le t$.

In other words, $P_{s,t}$ depends on s and t only through the difference $t - s$.

Notation 5.1.22. If $\{P_{s,t} : 0 \le s \le t\}$ is homogeneous, we write P_t instead of $P_{0,t}$ for all $t \ge 0$, and denote the transition function by $\{P_t : t \ge 0\}$.

Fact 5.1.23. If $\{P_t : t \ge 0\}$ is homogeneous, the Chapman-Kolmogorov equation can be rewritten as

$$(\forall s, t \ge 0) \quad P_{s+t} = P_sP_t. \tag{5.1.10}$$

Hence $\{P_t : t \ge 0\}$ is a semigroup under the kernel product. In other words, for all $x \in E$, for all $A \in \mathscr{E}$,

$$P_{s+t}(x, A) = \int P_s(x, dy)P_t(y, A). \tag{5.1.11}$$

Notation 5.1.24. From this point forward, we consider only homogeneous transition functions, so we can omit the word "homogeneous."

Fact 5.1.25. According to Fact 5.1.7, a transition function $\{P_t : t \ge 0\}$ on E induces a family of Markov operators $P_t : b\mathscr{E} \to b\mathscr{E}$ via

$$P_tf(x) = \int P_t(x, dy)f(y). \tag{5.1.12}$$

By (5.1.10), these operators form a semigroup under composition.

Now fix a sample space $\langle \Omega, \mathscr{F} \rangle$ and a state space $\langle E, \mathscr{E} \rangle$. To define an E-valued Markov process on Ω, we also need a filtration.

Definition 5.1.26. A *filtration* on Ω is a family $\{\mathscr{F}_t : t \geq 0\}$ of sub-σ-algebras of $\mathscr{F}$ such that $\mathscr{F}_s \subseteq \mathscr{F}_t$ whenever $s \leq t$.

Fix a filtration $\{\mathscr{F}_t : t \geq 0\}$ on Ω as well.

Definition 5.1.27. An E-valued stochastic process $X = \{X_t : t \geq 0\}$ on Ω is said to be $\mathscr{F}_t$-*adapted* if X_t is $\mathscr{F}_t$-measurable for each $t \geq 0$.

Now we can define a Markov process. (A time-homogeneous Markov process, to be precise, but these are the only Markov processes we will consider. For a more general definition, see Revuz and Yor [63].)

Definition 5.1.28. A *Markov process* with transition function $\{P_t : t \geq 0\}$ and initial distribution $\mu \in \mathcal{M}_1(E)$ is a pair $\langle X, \nu \rangle$ where

(i) $X = \{X_t : t \geq 0\}$ is an E-valued, $\mathscr{F}_t$-adapted stochastic process,

(ii) ν is a probability measure on Ω such that $\nu_{X_0} = \mu$, and

(iii) the *Markov property* holds: for all $s, t \geq 0$, for all $f \in b\mathscr{E}$,

$$\nu[f(X_{s+t})|\mathscr{F}_s] = (P_t f)(X_s) \quad \text{a.s. } [\nu]. \tag{5.1.13}$$

We would rather handle every initial distribution in one process:

Definition 5.1.29. A *Markov process* with transition function $\{P_t : t \geq 0\}$ (and no specified initial distribution) is a pair $\langle X, \mathbf{P} \rangle$ where

(i) $X = \{X_t : t \geq 0\}$ is an E-valued, $\mathscr{F}_t$-adapted stochastic process,

(ii) $\mathbf{P} = \{\mathbf{P}^\mu : \mu \in \mathcal{M}_1(E)\}$ is a family of probability measures on Ω such that $\mathbf{P}^\mu_{X_0} = \mu$ for all μ, and

(iii) for all μ, the Markov property holds: for all $s, t \geq 0$, for all $f \in b\mathscr{E}$,

$$\mathbf{P}^\mu[f(X_{s+t})|\mathscr{F}_s] = (P_t f)(X_s) \quad \text{a.s. } [\mathbf{P}^\mu]. \tag{5.1.14}$$

Let us assume that a Markov process $\langle X, \mathbf{P} \rangle$ actually exists. An initial distribution $\mu \in \mathcal{M}_1(E)$, together with the transition function $\{P_t : t \geq 0\}$, determines the finite-dimensional distributions of the process:

Fact 5.1.30. Let $0 = t_0 < t_1 < \cdots < t_n$. For all $f \in b\mathscr{E}^{n+1}$, the expected value $\mathbf{P}^\mu[f(X_{t_0}, \ldots, X_{t_n})]$ is given by the formula

$$\int \mu(dx_0) \int P_{t_1}(x_0, dx_1) \cdots \int P_{t_n - t_{n-1}}(x_{n-1}, dx_n) f(x_0, \ldots, x_n). \quad (5.1.15)$$

(In particular, this determines the finite-dimensional distributions of X.)

Remark 5.1.31. Fact 5.1.30 says that the initial position x_0 of the process is distributed according to μ, its position x_1 at time t_1 according to $P_{t_1}(x_0, \cdot)$, its position x_2 at time t_2, according to $P_{t_2 - t_1}(x_1, \cdot)$, and so on.

Notation 5.1.32. Let $s, t \geq 0$. We may write $\mathbf{P}^{X_s}[f(X_t)]$ for $P_t f(X_s)$, in which case we can rewrite the Markov property (5.1.14) as

$$\mathbf{P}^\mu[f(X_{s+t}) | \mathscr{F}_s] = \mathbf{P}^{X_s}[f(X_t)] \quad \text{a.s. } [\mathbf{P}^\mu]. \quad (5.1.16)$$

A Markov process is particularly well behaved if its transition function satisfies two extra assumptions:

Definition 5.1.33. $\{P_t : t \geq 0\}$ is said to be *normal* if $P_0(x, \cdot) = \varepsilon_x$ for all $x \in E$. In other words, P_0 is the identity operator on $b\mathscr{E}$.

Definition 5.1.34. $\{P_t : t \geq 0\}$ is said to be *conservative* if P_t is a probability kernel for each $t \geq 0$. In other words, $P_t(1) = 1$.

Remark 5.1.35. $1 - P_t(x, E)$ represents the probability that the process disappears or dies, so to speak, before or at time t. See Remark 5.1.4.

The first condition will not be a problem for us, since all Feller transition functions are normal. Neither will the second condition, in fact, as we can make any transition function conservative by extending the state space:

Definition 5.1.36. The *extended state space* for a given transition function $\{P_t : t \geq 0\}$ is $\langle E^*, \mathscr{E}^* \rangle$, where

(i) $*$ is a new symbol, called the *coffin state*,

(ii) $E^* = E \cup \{*\}$, and

(iii) $\mathscr{E}^* = \sigma(\mathscr{E}, *)$ is the smallest σ-algebra containing $\mathscr{E}$ and $\{*\}$.

Fact 5.1.37. Every transition function $\{P_t : t \geq 0\}$ induces a conservative transition function $\{P_t^* : t \geq 0\}$ on E^* as follows: for all $t \geq 0$,

(i) for all $x \in E$, $P_t^*(x, \{*\}) = 1 - P_t(x, E)$,

(ii) $P_t^*(*, \cdot) = \varepsilon_*$ is the unit mass at $*$, and

(iii) $P_t^* = P_t$ on $E \times \mathscr{E}$.

Remark 5.1.38. Fact 5.1.37 (ii) says that the coffin state is *absorbing*.

5.1.3 Feller processes

The class of Markov processes is a bit too large to effectivize: in general, the sample paths of a Markov process do not live in a well-behaved space of functions. Feller processes, on the other hand, have càdlàg modifications.

We have to assume that the state space E is *lccb*, i.e., a locally compact Hausdorff space with a countable base (equipped with the Borel σ-algebra). Such a space has a one-point compactification. The definition of a Feller process will be stated in terms of real-valued functions on the state space that vanish at infinity. Recall Definitions 2.2.88 and 2.2.90.

Definition 5.1.39. A function $f : E \to \mathbb{R}$ is said to *vanish at infinity* if for all $\epsilon > 0$ there exists a compact $K \subseteq E$ such that $|f(x)| < \epsilon$ for all $x \notin K$.

Definition 5.1.40. $\mathbf{C}_0(E)$ denotes the vector space of continuous functions $E \to \mathbb{R}$ that vanish at infinity. $\mathbf{C}_0(E)$ is a complete, separable subspace of $b\mathscr{E}$ under the sup-norm.

Fact 5.1.41. Let $\varphi : \mathbf{C}_0(E) \to b\mathscr{E}$ be a linear operator such that

$$0 \le f \le 1 \implies 0 \le \varphi f \le 1. \tag{5.1.17}$$

Then there exists a unique sub-probability kernel κ on E such that

$$(\forall f \in \mathbf{C}_0(E))\ (\forall x \in E) \quad \varphi f(x) = \int \kappa(x, dy) f(y). \tag{5.1.18}$$

Thus, by Fact 5.1.7, φ extends to a Markov operator $b\mathscr{E} \to b\mathscr{E}$.

Definition 5.1.42. A *Feller semigroup* on $\langle E, \mathscr{E} \rangle$ is a family $\{T_t : t \ge 0\}$ of linear operators $\mathbf{C}_0(E) \to \mathbf{C}_0(E)$ with the following properties:

(i) $T_0 = \mathrm{id}$, and for all $s, t \ge 0$, $T_{t+s} = T_t T_s$

(ii) for all $t \ge 0$, for all $f \in \mathbf{C}_0(E)$, $0 \le f \le 1$ implies $0 \le T_t f \le 1$

(iii) for all $f \in \mathbf{C}_0(E)$, $\lim_{t \downarrow 0} \|T_t f - f\|_\infty = 0$

Property (iii) is called *strong continuity*.

Remark 5.1.43. By Fact 5.1.41, every Feller semigroup $\{T_t : t \ge 0\}$ extends to a unique transition function on E, which we denote by $\{P_t : t \ge 0\}$.

Definition 5.1.44. A transition function $\{P_t : t \ge 0\}$ associated with a Feller semigroup is called a *Feller transition function*. A Markov process $\langle X, \mathbf{P} \rangle$ with a Feller transition function is called a *Feller process*.

Feller semigroups have another characterization:

Fact 5.1.45. Let $\{T_t : t \geq 0\}$ be a family of linear operators $\mathbf{C}_0(E) \to \mathbf{C}_0(E)$ satisfying properties (i) and (ii) in Definition 5.1.42. Then property (iii), strong continuity, is implied by the apparently weaker condition

(iii)$'$ for all $f \in \mathbf{C}_0(E)$, for all $x \in E$, $\lim_{t \downarrow 0} |T_t f(x) - f(x)| = 0$.

This, in turn, leads to a characterization of Feller transition functions:

Fact 5.1.46. A transition function $\{P_t : t \geq 0\}$ is Feller if and only if

(i) for all $t \geq 0$, for all $f \in \mathbf{C}_0(E)$, $P_t f \in \mathbf{C}_0(E)$, and

(ii) for all $f \in \mathbf{C}_0(E)$, for all $x \in E$, $\lim_{t \downarrow 0} |P_t f(x) - f(x)| = 0$.

Notation 5.1.47. In light of Remark 5.1.43, we can identify—classically, at least—a Feller semigroup with the associated transition function.

Let us assume that a Feller process $\langle X, \mathbf{P} \rangle$ actually exists. Of course, for all $\mu \in \mathcal{M}_1(E)$, the Markov property holds: for all $s, t \geq 0$, for all $f \in b\mathscr{E}$,

$$\mathbf{P}^\mu[f(X_{s+t})|\mathscr{F}_s] = (P_t f)(X_s) \quad \text{a.s. } [\mathbf{P}^\mu].$$

It is often possible to reduce the theory of Feller processes to the case where the state space E is compact by taking a one-point compactification of E. Recall the extended state space E^* (Definition 5.1.36), which makes the transition function conservative (according to Fact 5.1.37).

Definition 5.1.48. Whenever E is lccb, let the extended state space be a one-point compactification of E. That is, let $*$ be a point at infinity.

Fact 5.1.49. Every Feller transition function induces a conservative Feller transition function on its extended state space.

Feller processes do, in fact, exist, and like Lévy processes, they have càdlàg modifications. The main result in this chapter will effectivize

Fact 5.1.50. For every Feller semigroup $\{P_t : t \geq 0\}$, there exists an E^*-valued càdlàg Feller process $\langle X, \mathbf{P} \rangle$ with transition function $\{P_t : t \geq 0\}$.

Technically, the transition function for the E^*-valued Feller process in Fact 5.1.50 is the conservative transition function $\{P_t^* : t \geq 0\}$ induced by $\{P_t : t \geq 0\}$ on E^*. However, for simplicity, we adopt

Notation 5.1.51. We will continue to refer to $\{P_t : t \geq 0\}$ as the transition function, to E as the state space, and to E^* as the extended state space.

5.2 Computable Feller processes

In this section, we explain what it means to compute a probability kernel, a transition function, a Feller semigroup, and a Feller process.

5.2.1 Computable probability kernels

From now on, of course, our measurable spaces must be computable metric spaces, equipped with the Borel σ-algebra.

Definition 5.2.1. Let E and F be computable metric spaces. A probability kernel κ from E to F is said to be *computable* if the function $x \mapsto \kappa^x$ is computable of type $E \to \mathcal{M}_1(F)$. A *name* for κ is a name for this function.

Gács [39] uses a different definition of a computable probability kernel, based on the sequence $(g_i)_{i\in\mathbb{N}}$ from Definition 2.3.49. But our definition is equivalent to his, and the proof is routine computable analysis:

Proposition 5.2.2. A probability kernel κ from E to F is computable if and only if $\kappa g_i : E \to \mathbb{R}$ is computable, uniformly from i, for all $i \in \mathbb{N}$.

In the classical theory of Feller processes, we had to assume that the state space was locally compact. Let us now assume that E is effectively locally compact. Then the space $\mathbf{C}_0(E)$ of continuous functions $E \to \mathbb{R}$ that vanish at infinity is a computable Banach space (Corollary 2.2.101).

We can also characterize computable probability kernels in terms of the associated Markov operator:

Proposition 5.2.3. Let κ be a probability kernel on E. The following are equivalent, and this can be made uniform:

(a) κ is computable in the sense of Definition 5.2.1

(b) $f, x \mapsto \kappa f(x)$ is computable of type $\mathbf{C}_0(E) \times E \to \mathbb{R}$

(c) $f, x \mapsto \kappa f(x)$ is computable of type $\mathbf{C}(E, [0,1]) \times E \to [0,1]$

(d) $f \mapsto \kappa f$ is computable of type $\mathbf{C}(E, [0,1]) \to \mathbf{C}(E, [0,1])$

Proof. The equivalence of (a), (b), and (c) follows from Proposition 2.3.9. The equivalence of (c) and (d) follows from Corollary 2.2.85. □

We can also see that the kernel product is a computable operation:

Proposition 5.2.4. Let κ_1 and κ_2 be probability kernels on E. Then $\kappa_1\kappa_2$ is computable uniformly from names for κ_1 and κ_2.

Proof. Uniformly from names for κ_1 and κ_2, the composition of $f \mapsto \kappa_1 f$ and $f \mapsto \kappa_2 f$ is computable from $\mathbf{C}(E, [0,1])$ to $\mathbf{C}(E, [0,1])$ by Proposition 5.2.3. One more application of that result completes the proof. □

These are more than enough facts about computable probability kernels in general. We proceed to computable transition functions.

5.2.2 Computable Feller semigroups

We continue to assume that E is effectively locally compact.

Definition 5.2.5. A transition function $\{P_t : t \geq 0\}$ on E is said to be *computable* if $t, x \mapsto P_t^x$ is computable of type $[0,\infty) \times E \to \mathcal{M}_1(E)$.

Proposition 5.2.6. $\{P_t : t \geq 0\}$ is computable if and only if $t, f, x \mapsto P_t f(x)$ is computable of type $[0,\infty) \times \mathbf{C}_0(E) \times E \to \mathbb{R}$.

Proof. This is immediate from Proposition 5.2.3. □

Definition 5.2.7. A Feller semigroup $\{T_t : t \geq 0\}$ is said to be *computable* if $t, f \mapsto T_t f$ is computable of type $[0,\infty) \times \mathbf{C}_0(E) \to \mathbf{C}_0(E)$.

Proposition 5.2.8. The transition function associated with a computable Feller semigroup is itself computable. This can be made uniform.

Proof. This is immediate from Proposition 5.2.6. □

Notation 5.2.9. In light of Proposition 5.2.8, we can identify a Feller semigroup with the associated transition function. (Recall Notation 5.1.47.)

Let $\{P_t : t \geq 0\}$ be a computable Feller transition function. Recall the property of strong continuity, from Definition 5.1.42 (iii):

$$\lim_{t \downarrow 0} \|P_t f - f\|_\infty = 0 \quad \text{for all } f \in \mathbf{C}_0(E).$$

The next result effectivizes this key property.

Proposition 5.2.10. Let $K \subseteq E$ be compact. Given $\epsilon > 0$, uniformly from a name for $K \in \mathcal{H}(E)$, we can compute a $\delta > 0$ such that

$$1 - P_t(x, B(x, \epsilon)) < \epsilon \tag{5.2.1}$$

whenever $0 \leq t < \delta$ and $x \in K$.

Proof. We effectivize a standard classical argument [12]. Define a family $\{f_x : x \in E\}$ of continuous functions on E by:

$$f_x(y) = \begin{cases} 1 - d(x,y)/\epsilon & \text{if } d(x,y) < \epsilon, \\ 0 & \text{otherwise.} \end{cases}$$

Evidently, $f_x \in \mathbf{C}_0(E)$ is computable uniformly from a name for x, and it is not difficult to show that $\|f_x - f_z\|_\infty \leq d(x,z)/\epsilon$ for all $x, z \in E$.

Uniformly a name for $K \in \mathcal{H}(E)$, we can compute a finite set of points $x_0, \ldots, x_m \in E$ such that

$$K \subseteq \bigcup_{i \leq m} B_{\epsilon^2/4}(x_i).$$

In other words, for all $x \in K$, $\min_{i \leq m} d(x, x_i) < \epsilon^2/4$. It follows that the family of functions $\{f_x : x \in K\}$ is similarly bounded: for all $x \in K$,

$$\min_{i \leq m} \|f_x - f_{x_i}\|_\infty < (\epsilon^2/4)/\epsilon = \epsilon/4.$$

Next, we have

$$P_t(x, B(x,\epsilon)) = \int P_t(x, dy) \mathbf{1}_{B(x,\epsilon)}(y) \geq \int P_t(x, dy) f_x(y) = P_t f_x(x),$$

which implies that

$$1 - P_t(x, B(x,\epsilon)) \leq f_x(x) - P_t f_x(x).$$

We will bound the right-hand side uniformly. Taking a supremum over E, we have, for all $i \leq m$,

$$\|f_x - P_t f_x\|_\infty \leq \|f_x - f_{x_i}\|_\infty + \|f_{x_i} - P_t f_{x_i}\|_\infty + \|P_t f_{x_i} - P_t f_x\|_\infty.$$

Since $x \in K$, we have $\|f_x - f_{x_i}\|_\infty < \epsilon/4$ for some $i \leq m$. Furthermore, $\|P_t\| \leq 1$ (linear operator norm), so $\|P_t f_{x_i} - P_t f_x\|_\infty \leq \|f_{x_i} - f_x\|_\infty < \epsilon/4$ as well.

By hypothesis, $t, f \mapsto P_t f$ is computable from $[0,\infty) \times \mathbf{C}_0(E)$ to $\mathbf{C}_0(E)$. Since we can compute a name for $f_{x_i} \in \mathbf{C}_0(E)$ from $i \in \mathbb{N}$, we can compute a δ such that whenever $0 \leq t < \delta$, $\|f_{x_i} - P_t f_{x_i}\|_\infty < \epsilon/2$ for all $i \leq m$. Hence

$$\|f_x - P_t f_x\|_\infty < \epsilon/4 + \epsilon/2 + \epsilon/4 = \epsilon.$$

Therefore, $0 \leq t < \delta$ implies $1 - P_t(x, B_\epsilon(x)) < \epsilon$, as required. □

Proposition 5.2.10 says that for t near 0, P_t^x is nearly concentrated around x. Hence if a bounded, continuous function f vanishes near x, then $P_t f(x) = \int f(y) dP_t^x(y)$ nearly vanishes too. The proof is routine:

Corollary 5.2.11. Let $K \subseteq E$ be compact. Given $\epsilon > 0$, uniformly from a name for K in $\mathcal{H}(E)$, we can compute a $\delta > 0$ such that $|P_t f(x)| < \epsilon$ whenever $0 \leq t < \delta$, $x \in K$, and $f \in \mathbf{C}(E, [0,1])$ vanishes on $B(x, \epsilon)$.

Let $\langle \Omega, \mathbb{P} \rangle$ be a computable probability space equipped with a filtration $\{\mathscr{F}_t : t \geq 0\}$, representing the sample space of a stochastic process.

Definition 5.2.12. An E-valued adapted process X on Ω is said to be *computably Markov* if it is a Markov process with a computable transition function. X is said to be *computably Feller* if it is a Feller process with a computable Feller semigroup.

Proposition 5.2.13. A computable Feller process is computably Markov.

Proof. This is immediate from Proposition 5.2.8. □

Remark 5.2.14. Without any real loss of generality, replacing E by E^*, we can assume that a computable Feller transition function is conservative and that the state space E is effectively compact: see Fact 5.1.49, Theorem 2.2.73, and Proposition 2.2.104. Obviously, in that case, $\mathbf{C}_0(E) = \mathbf{C}(E)$.

The main goal of this chapter is to effectivize Fact 5.1.50 as follows.

Theorem 5.2.15. Let $\{P_t : t \geq 0\}$ be a computable, conservative Feller semigroup on an effectively compact space E. There exists a càdlàg Feller process $\langle X, \mathbf{P} \rangle$ with transition function $\{P_t : t \geq 0\}$ such that

(i) $t, \omega \mapsto X(t, \omega)$ is $L^0(\lambda \times \mathbf{P}^\mu)$-computable of type $[0, \infty) \times \Omega \to E$, uniformly from a name for $\mu \in \mathcal{M}_1(E)$, where λ is Lebesgue measure;

(ii) for all $t \geq 0$, $X(t, \cdot)$ is $L^0(\mathbf{P}^\mu)$-computable of type $\Omega \to E$, uniformly from names for t and $\mu \in \mathcal{M}_1(E)$; and

(iii) $\omega \mapsto X(\cdot, \omega)$ is $L^0(\mathbf{P}^\mu)$-computable of type $\Omega \to \mathbf{D}_E[0, T]$, uniformly from a dyadic rational $T > 0$ and a name for $\mu \in \mathcal{M}_1(E)$.

Remark 5.2.16. In the terminology of Chapter 4, Theorem 5.2.15 (ii) says that X is stochastically computable, and (iii) says that X is effectively measurable. Incidentally, (i) says that X is effectively *jointly measurable*.

The proof of Theorem 5.2.15 is carried out in the next section, but before that, we specialize effective Feller process theory to Lévy processes.

Essentially, a one-dimensional Lévy process (see Section 4.1) is an $\mathbb{R}$-valued Feller process whose transition function is induced by a so-called *convolution semigroup*, which means it takes a particularly simple form [3]:

Fact 5.2.17. The transition function for a Lévy process X is given by

$$P_t f(x) = \mathbb{P}[f(X_t - X_0 + x)], \tag{5.2.2}$$

or equivalently by $P_t(x, A) = \mathbb{P}\{X_t - X_0 + x \in A\}$.

In particular, stochastic continuity (Definition 4.1.14) at $t = 0$ implies

Fact 5.2.18. For all $\epsilon > 0$, there exists a $\delta > 0$ such that for all $s < \delta$,

$$1 - P_s(x, B(x, \epsilon)) = 1 - \mathbb{P}\{\|X_s - X_0\| < \epsilon\} < \epsilon. \tag{5.2.3}$$

Thanks to these classical results, we can easily prove

Proposition 5.2.19. Every stochastically computable Lévy process is also a computable Feller process.

Proof. Let $X = \{X_t : t \geq 0\}$ be the Lévy process. Fact 5.2.17 gives us the Feller semigroup $\{P_t : t \geq 0\}$ of X, so most of our work is done:

$$P_t f(x) = \mathbb{P}[f(X_t - X_0 + x)].$$

Now, uniformly from a name for t, X_t and X_0 are $L^0(\mathbb{P})$-computable. It follows that, uniformly from names for $x \in \mathbb{R}$ and $f \in \mathbf{C}_0(\mathbb{R})$, $f(X_t - X_0 + x)$ is $L^1(\mathbb{P})$-computable, which lets us compute $P_t f(x)$, as required. □

Stochastic computability implies an effective version of Fact 5.2.18. This is essentially Proposition 5.2.10 specialized to Lévy processes:

Proposition 5.2.20. Let X be a stochastically computable Lévy process. Given $\epsilon > 0$, we can compute a $\delta > 0$ such that for all $0 \leq t < \delta$,

$$1 - P_t(x, B(x, \epsilon)) = 1 - P\{\|X_t - X_0\| < \epsilon\} < \epsilon. \tag{5.2.4}$$

Proof. Set $t = 0$ in Proposition 4.2.3. □

5.3 Computably càdlàg Feller processes

In this section, we develop the effective theory of Feller processes in a way that is broadly consistent with the constructive theory of Chan [19, 20]. Our goal is to prove Theorem 5.2.15 by adapting the constructive argument.

5.3.1 Pseudo-canonical process

Classically, the existence of a Feller process in a state space E with a given transition function is established using the Daniell-Kolmogorov extension theorem, which defines a so-called canonical process on the sample space of all functions from, say, $[0,1]$ to $\mathbb{R}$. After establishing the existence of some such process, we can show that its sample paths are actually càdlàg.

In computable analysis, this approach must be modified. For one thing, the extension theorem is not effective; for another, the space of all functions (with the product topology) is not separable. Instead, we define what we call a "pseudo-canonical" process X indexed by dyadic rational times. Later, we will prove that X satisfies an appropriate version of the Markov property.

The set-up, here and for the rest of this section, is as follows.

Notation 5.3.1. $\langle E, d, S\rangle$ denotes an effectively compact metric space, which we identify with E, and $\mathscr{E}$ denotes its Borel σ-algebra. $\{P_t : t \geq 0\}$ denotes a computable, conservative Feller transition function on E.

Remark 5.3.2. We are assuming that E is effectively compact and that $\{P_t\}$ is conservative without loss of generality: see Remark 5.2.14.

Notation 5.3.3. Let $D = \{q_i^\circ : i \in \mathbb{N}\}$ enumerate the non-negative dyadic rationals: D will be the index set for our pseudo-canonical process. Let Ω be the product space E^D (Definition 2.2.46), which is effectively compact. $\mathscr{F} = \mathscr{B}(\Omega)$ denotes the Borel σ-algebra of Ω.

Definition 5.3.4. The *coordinate process* on Ω is the computable function $X : D \times \Omega \to E$ given by $X(t,\omega) = \omega(t)$. Of course, X is the evaluation operator; here, it represents a D-indexed, E-valued stochastic process on Ω.

Remark 5.3.5. $\mathscr{F} = \mathscr{B}(\Omega) = \sigma\{X_t : t \in D\}$.

With the set-up in place, we need to define probability measures $\mathbf{P}^\mu$ on Ω, which amounts to integrating bounded (Borel-) measurable functions from Ω to $\mathbb{R}$. It is enough to integrate continuous functions, and the Stone-Weierstrass theorem allows us to approximate them conveniently.

Definition 5.3.6. Let $\mathbf{A} \subseteq \mathbf{C}(\Omega)$ be the set of functions of the form

$$f_1(X_{t_1})f_2(X_{t_2})\cdots f_n(X_{t_n}), \tag{5.3.1}$$

where $f_1, f_2, \ldots, f_n \in \mathbf{C}(E)$ and $t_1 < t_2 < \cdots < t_n$ belong to D. Let $\mathbf{A}_0 \subseteq \mathbf{A}$ be the subset of functions of the form (5.3.1) where $f_1, f_2, \ldots, f_n$ are ideal points from $\mathbf{C}(E)$. Let $\mathrm{Lin}(\mathbf{A}) \subseteq \mathbf{C}(\Omega)$ be the set of linear combinations of elements of $\mathbf{A}$ with coefficients from $\mathbb{R}$, and let $\mathrm{Lin}_0(\mathbf{A}_0) \subseteq \mathrm{Lin}(\mathbf{A})$ be the set of linear combinations of elements of $\mathbf{A}_0$ with coefficients from $\mathbb{Q}$.

Recall that Ω is compact. The Stone-Weierstrass theorem implies

Fact 5.3.7. $\mathrm{Lin}(\mathbf{A})$ is dense in $\mathbf{C}(\Omega)$ (with respect to the sup-norm).

In fact, we can say more (the proof is routine computable analysis):

Proposition 5.3.8. $\mathrm{Lin}_0(\mathbf{A}_0)$ is a computably equivalent set of ideal points for the computable Banach space $\mathbf{C}(\Omega)$ (under the sup-norm).

We know how the transition function $\{P_t : t \geq 0\}$ acts on the state space: it maps each point $x \in E$ to a distribution P_t^x on E. We need to define, for each $\mu \in \mathcal{M}_1(E)$, a measure $\mathbf{P}^\mu$ on the sample space Ω that represents the law of the coordinate process X with initial distribution μ. For that, we need to define a probability kernel κ from E to Ω, i.e., a Markov operator $\kappa : b\mathscr{F} \to b\mathscr{E}$. We begin by defining κ on functions of the form (5.3.1).

Definition 5.3.9. For $f = f_1(X_{t_1})f_2(X_{t_2})\cdots f_n(X_{t_n}) \in \mathbf{A}$,

$$\kappa f = P_{t_1}(f_1 \cdot P_{t_2-t_1}(f_2 \cdots P_{t_n-t_{n-1}}(f_n)\cdots)), \tag{5.3.2}$$

where $f_1, f_2, \ldots, f_n \in \mathbf{C}(E)$ and $t_1 < t_2 < \cdots < t_n$ belong to D.

Proposition 5.3.10. The function κ maps $\mathbf{A}$ to $\mathbf{C}(E)$, and extends to a computable linear operator $\kappa : \mathbf{C}(\Omega) \to \mathbf{C}(E)$ such that (i) $\kappa(1) = 1$, and (ii) for all $f \in \mathbf{C}(\Omega)$, $0 \leq f \leq 1$ implies $0 \leq \kappa f \leq 1$.

Proof. Since $\{P_t : t \geq 0\}$ is a Feller semigroup, we can see from Definition 5.3.9 that $\kappa f \in \mathbf{C}(E)$ for all $f \in \mathbf{A}$. Since $\{P_t : t \geq 0\}$ is computable, κf is computable (as a point in $\mathbf{C}(E)$) uniformly from a name for f (in $\mathbf{C}(\Omega)$).

We extend κ, defined by (5.3.2), from $\mathbf{A}$ to $\mathrm{Lin}(\mathbf{A})$ by linearity:

$$\kappa\left(\sum_{i\leq k} a_i f_1^i(X_{t_1})\cdots f_n^i(X_{t_n})\right) = \sum_{i\leq k} a_i P_{t_1}(f_1^i \cdot P_{t_2-t_1}(f_2^i \cdots P_{t_n-t_{n-1}}(f_n^i)\cdots)).$$

It is not difficult to show that κ is well-defined on $\mathrm{Lin}(\mathbf{A})$; that is, if $f = 0$, then $\kappa f = 0$ as well. And it is clear that κ is a linear operator from $\mathrm{Lin}(\mathbf{A})$ to $\mathbf{C}(E)$, and that κf is computable uniformly from a name for f.

By a standard continuity argument, the extension of κ from $\mathrm{Lin}(\mathbf{A})$ to $\mathbf{C}(\Omega)$ is linear. In fact, the standard continuity argument is effective, and since κ is computable on $\mathrm{Lin}_0(\mathbf{A}_0)$, its extension is computable on $\mathbf{C}(\Omega)$.

It is easy to see from Definition 5.3.9 that (i) $\kappa(1) = 1$, and (ii) for all $f \in \mathbf{A}$, $0 \leq f \leq 1$ implies $0 \leq \kappa f \leq 1$. The latter property extends to $\mathrm{Lin}(\mathbf{A})$ by linearity, then to $\mathbf{C}(\Omega)$ by continuity, which completes the proof. □

Proposition 5.3.11. For all $\mu \in \mathcal{M}_1(E)$, κ induces a probability measure $\mu\kappa$ on Ω. A name for $\mu\kappa$ is computable uniformly from a name for μ.

Proof. The integral operator associated with $\mu\kappa$ is given by $(\mu\kappa)(f) = \mu(\kappa f)$ for $f \in \mathbf{C}(\Omega)$. By Proposition 5.3.10, $\kappa : \mathbf{C}(\Omega) \to \mathbf{C}(E)$ is computable, so $\mu(\kappa f)$ is computable uniformly from names for μ and f. Finally, $\kappa(1) = 1$, so $\mu\kappa$ is a probability measure. This completes the proof. □

Remark 5.3.12. κ defines a computable probability kernel from E to Ω.

The computable probability kernel κ is essential to the theory.

Remark 5.3.13. We review some important facts about probability kernels. We know that κ induces a computable linear operator $\kappa : b\mathscr{F} \to b\mathscr{E}$. This operator has the following properties:

(i) κ is positive: $f \geq 0$ implies $\kappa f \geq 0$ (for all $f \in b\mathscr{F}$)

(ii) κ is monotone: $f \leq g$ implies $\kappa f \leq \kappa g$

(iii) κ is a contraction: $\|f\|_\infty \leq 1$ implies $\|\kappa f\|_\infty \leq 1$ (in fact, $\|\kappa\| = 1$)

(iv) $f_n \downarrow 0$ (pointwise) implies $\kappa f_n \downarrow 0$ (pointwise)

Notation 5.3.14. For all $\mu \in \mathcal{M}_1(E)$, $\mathbf{P}^\mu$ denotes $\mu\kappa \in \mathcal{M}_1(\Omega)$.

Going forward, μ represents an initial distribution (on the state space) for a Feller process, and $\mathbf{P}^\mu$ the corresponding measure on the sample space.

Proposition 5.3.15. For all $\mu \in \mathcal{M}_1(E)$ and $t \in D$, $\mathbf{P}^\mu_{X_t} = \mu P_t$.

Proof. Because of all the new notation, we will work through the details. Let $f \in \mathbf{C}(E)$. By the change-of-variables formula,

$$\mathbf{P}^\mu_{X_t}(f) = \mathbf{P}^\mu(f \circ X_t) = \mu\kappa(f \circ X_t).$$

By the definition (5.3.9) of κ on $\mathbf{A}$, $\kappa(f(X_t)) = P_t f$, so $\mu\kappa(f \circ X_t) = \mu(P_t f)$. Since $f \in \mathbf{C}(E)$ was arbitrary, this completes the proof. □

Remark 5.3.16. As a special case of Proposition 5.3.15, $\mathbf{P}^\mu_{X_0} = \mu$.

5.3.2 Pseudo-Markov property

In the last part, given a conservative Feller semigroup $\{P_t : t \geq 0\}$, we defined a probability kernel κ which took a distribution μ on E to a distribution $\mathbf{P}^\mu = \mu\kappa$ on $\Omega = E^D$. Here, we verify a version of the Markov property (5.1.14) for the coordinate process $X : D \times \Omega \to E$.

To state this "pseudo-Markov" property, we need a number of tools.

Definition 5.3.17. For all $s \in D$, $\theta_s : \Omega \to \Omega$ denotes the *left-shift operator*

$$\theta_s(\omega) : t \mapsto \omega(t + s). \tag{5.3.3}$$

Notation 5.3.18. If $Z : \Omega \to \mathbb{R}$ is a random variable, we may denote the random variable $(\kappa Z)(X_s) : \Omega \to \mathbb{R}$ by $\kappa^{X_s}(Z)$. This will be used later.

Evidently, the Markov property—the entire concept of a memoryless process—is based on conditional expectations. Classically, these are defined in terms of a filtration: conditioning on σ-algebras. However, σ-algebras are not easily effectivized, so we will condition on sets of measurable functions.

Definition 5.3.19. Let $\langle M, \rho\rangle$ be a complete, separable metric space with the Borel σ-algebra $\mathscr{B}(M)$. Let ν be a probability measure on M. A set $\mathscr{G}$ of measurable functions from M to $\mathbb{R}$ is said to be a *Borel family* if

1. $\mathscr{G}$ is closed under the Ky-Fan metric d_ν, and
2. if $g_1, \ldots, g_n \in \mathscr{G}$ and $f \in \mathbf{C}(\mathbb{R}^n)$, then $f(g_1, \ldots, g_n) \in \mathscr{G}$.

Definition 5.3.20. Let A be a subset of $L^0(M, \nu)$. $\mathscr{G}(A)$ denotes the set of functions of the form $f(g_1, \ldots, g_n)$ where $g_1, \ldots, g_n \in A$ and $f \in \mathbf{C}_0(\mathbb{R}^n)$.

Fact 5.3.21. The closure $\overline{\mathscr{G}}(A)$ of $\mathscr{G}(A)$ under d_ν is a Borel family, and every Borel family on M arises in this way for some countable A.

Notation 5.3.22. $\overline{\mathscr{G}}(A)$ is called the Borel family *generated by* A.

Next, we effectivize Borel families. Assume that M is an effectively compact space and ν is a computable probability measure on M.

Fact 5.3.23. Let A enumerate uniformly $L^0(\nu)$-computable functions from M to $\mathbb{R}$. For all $m \in \mathbb{N}$, let $(f_k^m)_{k\in\mathbb{N}}$ enumerate the ideal points of $\mathbf{C}_0(\mathbb{R}^m)$, so that by combining these sequences with A we have an enumeration of

$$\widehat{A} = \{f_k^m(g_1, \ldots, g_m) : m, k \in \mathbb{N} \text{ and } g_1, \ldots, g_m \in A\}.$$

Then the Borel family $\overline{\mathscr{G}}(A)$ generated by A is a computable metric space under d_ν with $\widehat{A}$ for ideal points.

Proof. $\widehat{A}$ is dense in $\overline{\mathscr{G}}(A)$, and the elements of $\widehat{A}$ are uniformly $L^0(\nu)$-computable by Fact 2.3.48. Apply Fact 2.2.18. □

By Fact 2.2.18, the computable points in $\overline{\mathscr{G}}(A)$ are exactly the L^0-computable functions that belong to $\overline{\mathscr{G}}(A)$, which we state as

Fact 5.3.24. A name for f in $\overline{\mathscr{G}}(A)$ is computable uniformly from a name for f in $L^0(M,\nu)$, and vice versa; in particular, a point $f \in \overline{\mathscr{G}}(A)$ is computable in $\overline{\mathscr{G}}(A)$ if and only if f is computable in $L^0(M,\nu)$.

Notation 5.3.25. The computable metric space $\langle \overline{\mathscr{G}}(A), d_\nu, \widehat{A}\rangle$ is called a *computable Borel family* in $L^0(M,\nu)$.

We are now able to condition on Borel families:

Definition 5.3.26. Let $f \in L^1(M,\nu)$. A *conditional expectation* of f given the Borel family $\overline{\mathscr{G}}(A)$ is a random variable $\widetilde{f} \in \overline{\mathscr{G}}(A)$ such that

$$(\forall g \in \overline{\mathscr{G}}(A)) \quad \nu(g \cdot \widetilde{f}) = \nu(g \cdot f). \tag{5.3.4}$$

Notation 5.3.27. Conditional expectations, so defined, are unique up to ν-a.e. equality. We denote the equivalence class by $\nu[f|\overline{\mathscr{G}}(A)]$.

Remark 5.3.28. It is enough to check that (5.3.4) holds for all $g \in \widehat{A}$.

Definition 5.3.29. The (equivalence class of a) conditional expectation $\widetilde{f}$ of f given $\overline{\mathscr{G}}(A)$ is said to be *computable* if $\widetilde{f}$ is computable in $\overline{\mathscr{G}}(A)$.

The Borel families we will be using are intended to simulate the natural filtration of a stochastic process, as follows. Let Z be a computable function from $D \times \Omega$ to E (where E is compact). We denote $Z(t,\cdot)$ by Z_t or $Z(t)$.

Definition 5.3.30. Let $s \in D$. For all $m \in \mathbb{N}$, let $(h_k^m)_{k\in\mathbb{N}}$ enumerate the ideal points of $\mathbf{C}(E^n)$. Combine these sequences to enumerate

$$A_s^Z = \{h_k^m(Z(t_1),\dots,Z(t_m)) : m,k \in \mathbb{N} \text{ and } t_1 < \cdots < t_m \in [0,s] \cap D\}$$

as a sequence of uniformly computable functions from Ω to $\mathbb{R}$.

Remark 5.3.31. Evidently, a function in $A_s = A_s^Z$ depends only on the state of Z at a finite collection of times bounded by s.

Notation 5.3.32. Fix a probability measure ν on Ω. We denote by $\mathscr{F}_s = \mathscr{F}_s^Z(\nu)$ the Borel family in $L^0(\Omega,\nu)$ generated by A_s. It represents complete information about the state of the process Z up to (and including) time s.

Remark 5.3.33. For all ν, $\mathscr{F}_s(\nu) \subseteq \mathscr{F}_t(\nu)$ for all $s \leq t$.

Remark 5.3.34. Recall Remark 5.3.28. For any $f \in L^1(\Omega, \nu)$, to verify that $\widetilde{f} = \nu[f|\mathscr{F}_s(\nu)]$, it is enough to check that (5.3.4) holds for $g \in A_s$.

Notation 5.3.35. The choice of measure ν for $\mathscr{F}_s$ is implicit in the notation for conditional expectation: $\nu[f|\mathscr{F}_s] = \nu[f|\mathscr{F}_s(\nu)]$.

Example 5.3.36. The most important cases are $Z = X$ and $Z = X \circ \theta_s$ for some $s \in D$. Since $X_t \circ \theta_s = X_{t+s}$ for all $s, t \in D$, we have, for all ν,

$$\mathscr{F}_t^{X\circ\theta_s}(\nu) \subseteq \mathscr{F}_{s+t}^X(\nu) \tag{5.3.5}$$

where the right-hand side represents complete information about X up to time $s+t$, while the left-hand side lacks information about X on $[0, s]$.

We can now state the pseudo-Markov property. $\mu \in \mathcal{M}_1(E)$ continues to represent an initial distribution for the Feller process we are constructing, while $\mathbf{P}^\mu = \mu\kappa$ represents the corresponding distribution on $\Omega = E^D$.

Notation 5.3.37. Unless otherwise specified, $\mathscr{F}_s$ denotes $\mathscr{F}_s^X$.

Proposition 5.3.38. Let $s \in D$, $\mu \in \mathcal{M}_1(E)$, and $Z \in L^1(\Omega, \mu P_s \kappa)$. Then

(i) the random variable $Z \circ \theta_s$ is L^1-computable on $\langle \Omega, \mathbf{P}^\mu \rangle$ uniformly from s and names for μ and Z, and

(ii) the *pseudo-Markov property* holds:

$$\mathbf{P}^\mu[Z \circ \theta_s|\mathscr{F}_s] = \kappa^{X_s}(Z) \quad (\mathbf{P}^\mu\text{-a.e.}) \tag{5.3.6}$$

Proof. Let $s \in D$. The following is uniform in a name for μ.

First, consider the special case where Z is of the form

$$Z = g_1(X_{s_1})g_2(X_{s_2})\cdots g_n(X_{s_m}), \tag{5.3.7}$$

where $g_1, \ldots, g_m \in \mathbf{C}(E)$ and $s_1 \leq \cdots \leq s_m$ belong to D (see Definition 5.3.6). In that case, it is clear that

$$Z \circ \theta_s = g_1(X_{s_1+s})g_2(X_{s_2+s})\cdots g_n(X_{s_m+s})$$

is $L^1(\mathbf{P}^\mu)$-computable. We still need to verify (ii), i.e., that

$$\mathbf{P}^\mu[Z \circ \theta_s|\mathscr{F}_s] = \kappa^{X_s}(Z) \quad (\mathbf{P}^\mu\text{-a.e.}).$$

In other words (Definition 5.3.26), for all $Y \in \mathscr{F}_s = \mathscr{F}_s(\mathbf{P}^\mu)$,

$$\mathbf{P}^\mu[Y \cdot (Z \circ \theta_s)] = \mathbf{P}^\mu[Y \cdot (\kappa Z)(X_s)]. \tag{5.3.8}$$

It is enough to check (5.3.8) for Y of the special form

$$Y = f_1(X_{t_1}) f_2(X_{t_2}) \cdots f_n(X_{t_n}),$$

where $f_1, \ldots, f_n \in \mathbf{C}(E)$ and $t_1 \leq \cdots \leq t_n = s$ belong to D. But in that case, (5.3.8) follows from Definition 5.3.9 by direct substitution.

Now consider the general case $Z \in L^1(\Omega, \mu P_s \kappa)$. Uniformly from a name for Z, compute a series $\sum Z_n$ converging to Z in $L^1(\Omega, \mu P_s \kappa)$ such that $\forall n$, Z_n has the form (5.3.7), and $\forall m$, $\sum_{n>m} \mu P_s \kappa |Z_n| < 2^{-m}$. Then

$$\sum_{n>m} \mathbf{P}^\mu |Z_n \circ \theta_s| = \sum_{n>m} \mathbf{P}^\mu (|Z_n| \circ \theta_s) = \sum_{n>m} \mathbf{P}^\mu [(\kappa |Z_n|)(X_s)]$$

by the special case (5.3.7), and by Proposition 5.3.15 this equals

$$\sum_{n>m} \mathbf{P}^\mu_{X_s}(\kappa |Z_n|) = \sum_{n>m} \mu P_s \kappa |Z_n| < 2^{-m}.$$

Hence the sequence of partial sums $\sum_{n>m} Z_n \circ \theta_s$ is fast Cauchy in $L^1(\Omega, \mathbf{P}^\mu)$, and it evidently converges to $Z \circ \theta_s$. We have established claim (i).

Finally, we establish claim (ii) in the general case by linearity:

$$\mathbf{P}^\mu[Y \cdot (Z \circ \theta_s)] = \sum \mathbf{P}^\mu[Y \cdot (Z_n \circ \theta_s)] = \sum \mathbf{P}^\mu[Y \cdot (\kappa Z_n)(X_s)],$$

for all $Y \in \mathscr{F}_s$ by the special case (5.3.7), while on the other hand

$$\sum \mathbf{P}^\mu[Y \cdot (\kappa Z_n)(X_s)] = \mathbf{P}^\mu[Y \cdot (\kappa Z)(X_s)]$$

(by linearity again). This completes the proof. □

5.3.3 Strong pseudo-Markov property

In the last part, we showed that the pseudo-canonical process X indexed by dyadic rationals (defined in Part 5.3.1) satisfies a version of the Markov property. The goal now is to extend X computably to a Markov process on the index set $[0, \infty)$. This will be accomplished in Part 5.3.4 using sample path regularity properties. In order to state and prove those regularity properties, as it turns out, we need a few more tools.

Classically, the canonical process satisfies the *strong Markov property*, which in essence states that the Markov property (5.1.14) holds at random times: for any non-negative stopping time τ, for all $x \in E$, $s \geq 0$, and $f \in b\mathscr{E}$,

$$\mathbf{P}^x[f(X_{s+\tau})|\mathscr{F}_s] = (P_\tau f)(X_s) \quad (\mathbf{P}^x\text{-a.s.}). \tag{5.3.9}$$

Here, we define effective stopping times (a few of them, at least) and use them to state and prove a "strong pseudo-Markov property."

By this point, we have accumulated quite a bit of notation:

Notation 5.3.39. Recall E and P_t (Notation 5.3.1), D and Ω (Notation 5.3.3), X (Definition 5.3.4), κ (Definition 5.3.9), $\mathbf{P}^\mu$ (Notation 5.3.14), θ_s (Definition 5.3.17), κ^{X_s} (Notation 5.3.18), and $\mathscr{F}_s^Z(\nu)$ (Notation 5.3.32).

$\mu \in \mathcal{M}_1(E)$ is still an initial distribution for X, while $\mathbf{P}^\mu = \mu\kappa$ is the corresponding distribution on $\Omega = E^D$.

Definition 5.3.40. A random variable τ on Ω taking values in a finite subset D° of D is said to be a *discrete stopping time* for the process X with initial distribution μ if $\forall s \in D^\circ$, the event $\{\tau = s\}$ belongs to $\mathscr{F}_s(\mathbf{P}^\mu)$.

Definition 5.3.41. A discrete stopping time τ is said to be *computable* if a name for $\{\tau = s\}$ in $\mathscr{F}_s(\mathbf{P}^\mu)$ is computable from s and a name for μ.

Remark 5.3.42. $Z = \sum_{s \in D^\circ} Z \cdot \mathbf{1}_{\{\tau = s\}}$ for any random variable Z on Ω.

Discrete stopping times define Borel families:

Definition 5.3.43. For a discrete stopping time τ for X starting from μ, $\mathscr{F}_\tau(\mathbf{P}^\mu) = \mathscr{F}_\tau^X(\mathbf{P}^\mu)$ denotes the set of random variables $Z : \Omega \to \mathbb{R}$ such that for all $s \in D^\circ$, for all $f \in \mathbf{C}_0(\mathbb{R})$, $f(Z)\mathbf{1}_{\{\tau = s\}}$ belongs to $\mathscr{F}_s(\mathbf{P}^\mu)$.

Fact 5.3.44. $\mathscr{F}_\tau(\mathbf{P}^\mu)$ is a Borel family, so we can condition $\mathbf{P}^\mu$ on $\mathscr{F}_\tau(\mathbf{P}^\mu)$. If τ is computable, then so is $\mathscr{F}_\tau(\mathbf{P}^\mu)$.

Remark 5.3.45. τ itself belongs to $\mathscr{F}_\tau(\mathbf{P}^\mu)$. Also, if τ' is another stopping time such that $\tau' \leq \tau$ (surely), then $\mathscr{F}_{\tau'}(\mathbf{P}^\mu) \subseteq \mathscr{F}_\tau(\mathbf{P}^\mu)$.

Remark 5.3.46. $\mathscr{F}_\tau(\mathbf{P}^\mu)$ represents complete information about X with initial distribution μ up to the random time τ.

We define a *random left-shift operator*:

Definition 5.3.47. For a discrete stopping time τ for X and μ, let

$$\theta_\tau(\omega) = \theta_{\tau(\omega)}(\omega).$$

Similarly, let $X_\tau(\omega) = X_{\tau(\omega)}(\omega)$. We may write $X(\tau)$ instead of X_τ.

Proposition 5.3.48. If τ and τ' are computable discrete stopping times (for X and μ), then $\tau + \tau' \circ \theta_\tau$ is also a computable discrete stopping time.

Proof. Let τ and τ' be D°-valued. Fix $s \in D^\circ$. $\tau + \tau' \circ \theta_\tau = s$ if and only if for some $q \in [0, s] \cap D^\circ$, $\tau = q$ and $\tau' \circ \theta_q = s - q$. $\{\tau = q\}$ is computable in $\mathscr{F}_q(\mathbf{P}^\mu) \subseteq \mathscr{F}_s(\mathbf{P}^\mu)$ by hypothesis. Similarly, $\{\tau' \circ \theta_q = s - q\}$ is computable in $\mathscr{F}_{s-q}^{X \circ \theta_q}(\mu) = \mathscr{F}_s(\mathbf{P}^\mu)$. So $\{\tau + \tau' \circ \theta_\tau = s\}$ is a finite union of sets

$$\{\tau = q\} \cap \{\tau' \circ \theta_q = s - q\}$$

that are computable in $\mathscr{F}_s(\mathbf{P}^\mu)$. This completes the proof. □

That brings us to the strong pseudo-Markov property:

Proposition 5.3.49. Uniformly from t, for all $t \in D$, let τ be a computable discrete stopping time for X and $\mu P_t \kappa$ and let Z be $L^1(\mu P_t \kappa)$-computable.

(i) $Z \circ \theta_\tau$ is $L^1(\mu P_t \kappa)$-computable, uniformly from t, for all $t \in D$.

(ii) The *strong pseudo-Markov property* holds:

$$\mathbf{P}^\mu(Z \circ \theta_\tau | \mathscr{F}_\tau) = \kappa^{X_\tau} Z \quad (\mathbf{P}^\mu\text{-a.e.}) \tag{5.3.10}$$

(iii) $\kappa(Z \circ \theta_\tau) = \kappa(\kappa^{X_\tau} Z)$ holds μP_t-a.s. for all $t \in D$.

Proof. All three claims follow from the definitions. Consider claim (i).

Let τ be D°-valued. Since Z is $L^1(\mu P_t \kappa)$-computable for all $t \in D$ and $\{P_u : u \in D\}$ is a semigroup, Z is $L^1(\mu P_t P_s \kappa)$-computable for all $s, t \in D$. By Proposition 5.3.38 (i), $Z \circ \theta_s$ is $L^1(\mu P_t \kappa)$-computable for all $s, t \in D$. Now claim (i) follows from $Z \circ \theta_\tau = \sum_{s \in D^\circ} (Z \circ \theta_s) \mathbf{1}_{\{\tau = s\}}$.

Claims (ii) and (iii) follow from the definitions in similar fashion. □

5.3.4 Effective sample path regularity

Here, we adapt two lemmas from Chan [19] (Lemmas 5.3 and 5.4 in that source). Those lemmas are used to prove a key result (Theorem 5.3.52) on the regularity of sample paths, which we need to prove the main result.

The first lemma from Chan [19], translated into computable analysis:

Lemma 5.3.50. The function

$$Y_s = \sup\{d(X_0, X_t) : t \in [0, s] \cap D\}$$

is $L^1(\mathbf{P}^\mu)$-computable uniformly from $s \in D$ and a name for $\mu \in \mathcal{M}_1(E)$.

Proof. Fix s and μ. Let $0 \le a < b$. By Lemma 4.2.20 (and see Lemma 4.2.21), compute a strictly increasing sequence $(a_k)_{k\in\mathbb{N}}$ in (a,b) such that for all $t \in D$ and $k \in \mathbb{N}$, $\{d(X_0,X_t) = a_k\}$ and $\{d(X_0,X_t) = a_{k+1} - a_k\}$ are $\mu P_r \kappa$-null for all $r \in D$. It follows that $\{d(X_0,X_t) > a_k\}$ and $\{d(X_0,X_t) > a_{k+1} - a_k\}$ are computable in $\mathscr{F}_t(\mu P_r\kappa) = \mathscr{F}_t^X(\mu P_r\kappa)$. Let $a_\infty = \lim a_k$.

Compute a strictly increasing sequence $(p_k)_{k\in\mathbb{N}} \subseteq \mathbb{N}$ with $p_0 = 0$ and such that for all $k \ge 1$, $2^{-p_k} s < \delta_k$, where δ_k is given by Corollary 5.2.11, setting $\epsilon = (a_{k+1} - a_k)/2$; that is, $|P_t f(x)| < (a_{k+1} - a_k)/2$ whenever $0 \le t < \delta_k$, $x \in E$, and $f \in \mathbf{C}(E,[0,1])$ vanishes on $B(x,(a_{k+1} - a_k)/2)$.

Approximate the dyadic rational times $[0,s] \cap D$ by finite sets as follows:

$$D_k = \{j 2^{-p_k} s : 0 \le j \le 2^{p_k}\}.$$

Then $\{0,s\} = D_0 \subseteq D_1 \subseteq \cdots$ and $\bigcup_k D_k = [0,s] \cap D$.

Compute, for each $t \in D$, the least $j = j(t) \in \mathbb{N}$ such that $t \in D_j$. If $j > 0$, also compute the least dyadic rational $r = r(t) \in D_{j-1}$ such that $r > t$.

Compute a sequence $(A_k)_{k\in\mathbb{N}}$ of measurable sets given by

$$A_k = \bigcap_{t\in D_k} \{d(X_0,X_t) \le a_{j(t)}\}.$$

Define, for each $k \in \mathbb{N}$, a random variable $T_k = s\mathbf{1}_{A_k} + \sum_{t\in D_k} t\mathbf{1}_{C_k(t)}$ where

$$C_k(t) = \{d(X_0,X_t) > a_{j(t)}\} \cap \{d(X_0,X_u) \le a_{j(u)} \text{ for all } u \in D_k \text{ with } u < t\}.$$

T_k represents the first time $t \in D_k$ when $d(X_0,X_t) > a_{j(t)}$, or s if no such time exists. T_k is a D_k-valued computable discrete stopping time for $\mathscr{F}_t(\mu P_r\kappa)$, uniformly from k. Moreover, for all $k \in \mathbb{N}$, $A_k \supseteq A_{k+1}$ and

$$A_k \smallsetminus A_{k+1} \subseteq \bigcup_{t\in D_{k+1}} \{T_{k+1} = t \text{ and } d(X_t, X_{r(t)}) > a_{k+1} - a_k\}.$$

From this, we can calculate

$$\begin{aligned}
\mathbf{P}^\mu(A_k \smallsetminus A_{k+1}) &\le \sum_{t\in D_{k+1}} \mathbf{P}^\mu\{T_{k+1} = t \text{ and } d(X_t, X_{r(t)}) > a_{k+1} - a_k\} \\
&\le \sum_{t\in D_{k+1}} \mathbf{P}^\mu[\mathbf{1}_{\{T_{k+1}=t\}} \cdot \kappa^{X_t}\{d(X_0, X_{r(t)-t}) > a_{k+1} - a_k\}]
\end{aligned}$$

by the pseudo-Markov property, Proposition 5.3.38 (ii), using the fact that $\{T_{k+1} = t\} \in \mathscr{F}_t(\mathbf{P}^\mu)$. And the last expression above is bounded by

$$\sum_{t\in D_{k+1}} \mathbf{P}^\mu[\mathbf{1}_{\{T_{k+1}=t\}}] \cdot (a_{k+1} - a_k)/2 = (a_{k+1} - a_k)/2,$$

because for $t \in D_{k+1}$, $v = r(t) - t < 2^{-p_k} s < \delta_k$. Therefore,

$$B = \bigcap_{k\in\mathbb{N}} A_k = \bigcap_{t\in[0,s]\cap D} \{d(X_0, X_t) \le a_{j(t)}\}$$

is $L^0(\mathbf{P}^\mu)$-computable, being an effective limit of $L^0(\mathbf{P}^\mu)$-computable sets.

Evidently, for all $\omega \in B$, $d(X_0(\omega), X_t(\omega)) < b$ for all $t \in [0,s] \cap D$, and for all $\omega \notin B$, $d(X_0(\omega), X_t(\omega)) > a$ for some $t \in [0,s] \cap D$.

Compute $M = \operatorname{diam}(E)$. Repeat the above construction to obtain, for each $n \in \mathbb{N}$, uniformly $L^0(\mathbf{P}^\mu)$-computable sets $B_0, B_1, \ldots, B_{2^n}$ such that for all $i \le 2^n$, for all $\omega \in B_i$, $d(X_0(\omega), X_t(\omega)) < i2^{-n}M$ for all $t \in [0,s] \cap D$, and for all $\omega \notin B_i$, $d(X_0(\omega), X_t(\omega)) > (i-1)2^{-n}M$ for some $t \in [0,s] \cap D$.

Finally, let

$$Z_n = \sum_{i=1}^{2^n} k2^{-n} M \mathbf{1}_{B_i \setminus B_{i-1}}.$$

Z_n is $L^1(\mathbf{P}^\mu)$-computable, uniformly from n, and $|Z_n - Z_{n+1}| \le 2^{-n}M$ for all n, so $Z = \lim Z_n$ is $L^1(\mathbf{P}^\mu)$-computable as well. It is easy to see that Z, when it is defined, is the supremum of $\{d(X_0, X_t) : t \in [0,s] \cap D\}$. Hence $Y_s = Z$ is $L^1(\mathbf{P}^\mu)$-computable, and the proof was uniform in s and μ. □

The second lemma from Chan [19], translated into computable analysis:

Lemma 5.3.51. Let $\epsilon > 0$. Compute $\delta = \delta(\epsilon)$ as in Corollary 5.2.11. Fix $s < \delta$ and $c > 2\epsilon$. Then $\kappa\{Y_s > c\} < 2\epsilon$ (where Y_s is given by Lemma 5.3.50).

Proof. For $x \in E$, let $\mu = \epsilon_x$, so that $\mathbf{P}^\mu = \mu\kappa = \kappa^x$, where we recall that

$$\kappa\{Y_s > c\}(x) = \kappa^x\{Y_s > c\}.$$

Let $a = \epsilon$ and $b = 2\epsilon$, and compute $(a_k)_{k\in\mathbb{N}}$, $(p_k)_{k\in\mathbb{N}}$, $(D_k)_{k\in\mathbb{N}}$, and $(A_k)_{k\in\mathbb{N}}$ as in the proof of Lemma 5.3.50. Then, for all $k \in \mathbb{N}$,

$$\kappa^x(A_k \setminus A_{k+1}) = \mathbf{P}^x(A_k \setminus A_{k+1}) < a_{k+1} - a_k.$$

Fix any $f \in \mathbf{C}(E^2)$ such that $\mathbf{1}_{\{d(x,y)>a_0\}} \le f(x,y) \le \mathbf{1}_{\{d(x,y)>\epsilon\}}$. Then

$$\kappa^x(\Omega \setminus A_0) = \kappa^x\{d(X_0, X_s) > a_0\} \le \kappa^x(f(X_0, X_s)) = P_s(f(x,\cdot))(x) \le \epsilon$$

where the last inequality requires that $s < \delta$. Therefore, letting $B = \bigcap_{k\in\mathbb{N}} A_k$,

$$\begin{aligned} \kappa^x(\Omega \setminus B) &= \mathbf{P}^x(\Omega \setminus A_0 \cup A_0 \setminus A_1 \cup A_1 \setminus A_2 \cup \cdots) \\ &< \epsilon + (a_1 - a_0) + (a_2 - a_1) + \cdots \\ &= \epsilon + a_\infty - a_0 < 2\epsilon \end{aligned}$$

Finally, $\{Y_s > c\} \subseteq \Omega \setminus B$ provided that $c > 2\epsilon$, which completes the proof. □

We use Lemmas 5.3.50 and 5.3.51 to prove the key result, adapted from Theorem 5.5 in Chan [19] (and similar to Theorem 4.2.27), on sample path regularity for the pseudo-canonical process X.

Theorem 5.3.52. Given $0 \neq T \in D$, uniformly from names for $0 < \epsilon < 1$ and $\mu \in \mathcal{M}_1(E)$, there exist $L^0(\mathbf{P}^\mu)$-computable random variables

$$0 = U_0 = V_0 \leq U_1 \leq V_1 \leq \cdots \leq U_n \leq V_n, \tag{5.3.11}$$

and an $L^0(\mathbf{P}^\mu)$-computable event G such that

(i) $\mathbf{P}^\mu(G \cup \{U_n \leq T\}) < 2\epsilon$, and

(ii) on $\Omega \setminus G$, $d(X(t), X(V_i)) < 2\epsilon$ whenever $U_i < t < U_{i+1}$ for some i.

Proof. Compute $\delta = \delta(\epsilon)$ as in Corollary 5.2.11. Compute a $q \in \mathbb{N}$ such that

$$q > \max\{2, 2T/[(1-\epsilon)\delta(\epsilon/2)]\}$$

and an $m \in \mathbb{N}$ such that $(q-2)^m/(q-1)^m < \epsilon$. Let $n = qm + 1$.

The following is uniform from a name for μ. As usual, $\mathbf{P}^\mu = \mu\kappa$. Let $a = \epsilon$ and $b = \epsilon + \epsilon/n$, and compute $(a_k)_{k\in\mathbb{N}}$, $(p_k)_{k\in\mathbb{N}}$, $(D_k)_{k\in\mathbb{N}}$, $\{j(t) : t \in D\}$, $(A_k)_{k\in\mathbb{N}}$, and $(T_k)_{k\in\mathbb{N}}$ as in the proof of Lemma 5.3.50. The random variable

$$Z_k = Y_{2^{-p_k}T} = \sup\{d(X(0), X(t)) : t \in D \cap [0, 2^{-p_k}T]\}.$$

is $L^1(\mathbf{P}^\mu)$-computable uniformly from k. By an easy generalization of Lemma 4.2.20 (and see Lemma 4.2.21), compute a sequence (a'_k) such that $a_k < a'_k < a_{k+1}$ for all k and such that $\{Z_k > a'_{k+1} - a_k\}$ is $L^0(\mu P_u \kappa)$-computable uniformly from $u \in D$. Set $B_k = \{Z_k \circ \theta_{T_{k+1}} > a'_{k+1} - a_k\}$. Then

$$\mathbf{1}_{B_k} = \mathbf{1}_{\{Z_k > a'_{k+1} - a_k\}} \circ \theta_{T_{k+1}}.$$

By Lemma 5.3.49, B_k is $L^0(\mu P_u \kappa)$-computable uniformly from u, and

$$\kappa(B_k) = \kappa\left(\kappa^{X(T_{k+1})}\{Z_k > a'_{k+1} - a_k\}\right).$$

Since $2^{-p_k}T < \delta((a_{k+1} - a_k)/2)$ and $a'_{k+1} - a_k > a_{k+1} - a_k$, Lemma 5.3.51 implies that $\kappa(B_k) < a_{k+1} - a_k$. Hence $\sum \mu P_u \kappa(B_k)$ converges (absolutely), so $B = \bigcup_k B_k$ is $L^0(\mu P_u \kappa)$-computable uniformly from u. Furthermore,

$$\kappa(B) < \sum_k (a_{k+1} - a_k) = a_\infty - a_0 < b - a = \epsilon/n.$$

Similarly, because T_{k+1} and T_k are $L^0(\mu P_u \kappa)$-computable uniformly from u, $T_{k+1} \leq T_k$, and $\{T_k - T_{k+1} > 2^{-p_k} T\} \subseteq B_k$, $T_\infty = \lim T_k$ is $L^0(\mu P_u \kappa)$-computable uniformly from u as well.

Set $U_1 = T_\infty$, $V_1 = T_1$, and $G_1 = B$. Now we restart the process from time T_1, and repeat the construction n times: for $0 < i \leq n$, $V_i = V_{i-1} + V_1 \circ \theta_{V_{i-1}}$, $U_i = V_{i-1} + U_1 \circ \theta_{V_{i-1}}$, and $G_i = \theta_{V_{i-1}}^{-1}(G_1)$. Set $G = G_1 \cup \cdots \cup G_n$.

Suppose $\omega \notin G$. For $0 < i \leq n$, set $\omega' = \theta_{V_{i-1}}(\omega)$. Then ω' belongs to $G_1^c = B^c = \bigcap_k B_k^c$, which implies that

$$d(X(T_{k+1}), X(r))(\omega') \leq a'_{k+1} - a_k < a_{k+2} - a_k$$

for $k \geq 1$ and for $r \in [T_{k+1}(\omega'), T_{k+1}(\omega') + 2^{-p_k} T] \subseteq [T_{k+1}(\omega'), T_k(\omega')]$.

Summing over all $k \geq 1$, we have, for all $r \in (T_\infty(\omega'), T_1(\omega')]$,

$$d(X(T_1), X(r))(\omega') < 4(a_\infty - a_1) < 2\epsilon.$$

Substituting $\omega' = \theta_{V_{i-1}}(\omega)$ and simplifying, we get $d(X(V_i), X(t))(\omega) < 2\epsilon$ for all $t \in (U_i(\omega), V_i(\omega)]$.

Now, if $t \in [V_i(\omega), U_{i+1}(\omega))$, write $t = r + V_i(\omega)$ where

$$r < U_{i+1}(\omega) - V_i(\omega) = T_\infty(\theta_{V_i}(\omega)) \leq T_{j(r)}(\theta_{V_i}(\omega)).$$

$T_{j(r)}$ is the first time $u \in D_{j(r)}$ when X leaves the $a_{j(r)}$-neighborhood of $X(0)$. Hence

$$d(X(t), X(V_i))(\omega) = d(X(r), X(0))(\theta_{V_i}(\omega)) \leq a_{j(r)} < a_\infty < 2\epsilon$$

Claim (ii) follows. It remains only to show that $G \cup \{U_n \leq T\}$ has small measure. By Lemma 5.3.49 (iii),

$$\mathbf{P}^\mu(G_i) = \mathbf{P}^\mu(B \circ \theta_{V_{i-1}}) = \mathbf{P}^\mu(\kappa^{X(V_{i-1})}(B)) < \epsilon/n,$$

so $\mathbf{P}^\mu(G) < \epsilon$. We just need to show that $\mathbf{P}^\mu\{U_n \leq T\}$.

For all $i \in \mathbb{N}$, for all $r \in D$ with $r > \delta(\epsilon/2)$,

$$\kappa(V_{i+1} - V_i) = \kappa(T_1 \circ \theta_{V_i}) = \kappa(\kappa^{X(V_i)}(T_1)) \geq \kappa(r\kappa^{X(V_i)}\{T_1 \geq r\})$$

which is $\geq \kappa(r\kappa^{X(V_i)}(Y_r \leq a_1)) \geq r(1-\epsilon)$, where Y_r is given in Lemma 5.3.50. Summing over i, $\kappa(V_q) \geq q\delta(\epsilon/2)(1-\epsilon) > 2T$. Since $0 \leq V_q \leq qT$, we have

$$\kappa(V_q \leq T) \leq (q-2)/(q-1).$$

By induction, $V_{2q} - V_q = V \circ \theta_V, \ldots, V_{mq} - V_{(m-1)q} = V \circ V_{(m-1)q}$. Applying Lemma 5.3.49 (ii) repeatedly,

$$\begin{aligned}
\mathbf{P}^\mu\{U_n \le T\} &\le \mathbf{P}^\mu\{V_{n-1} \le T\} = \mathbf{P}^\mu\{V_{mq} \le T\} \\
&\le \mathbf{P}^\mu\{V_q \le T, V_{2q} - V_q \le T, \ldots, V_{mq} - V_{(m-1)q} \le T\} \\
&= \mathbf{P}^\mu\{V_q \le T, V_{2q} - V_q \le T, \ldots, V \circ \theta_{(m-1)q} \le T\} \\
&= \mathbf{P}^\mu\{V_q \le T, V_{2q} - V_q \le T, \ldots, \kappa^{X(V_{(m-1)q})}(V \le T)\} \\
&\le \mathbf{P}^\mu\{V_q \le T, V_{2q} - V_q \le T, \ldots\}(q-2)/(q-1) \le \cdots \\
&\le (q-2)^m/(q-1)^m < \epsilon.
\end{aligned}$$

This proves (ii) and completes the proof. □

The three claims in Theorem 5.2.15 now follow from Theorem 5.3.52.

Definition 5.3.53. The *canonical process* X is the extension of the pseudo-canonical process from the index set D to $[0, \infty)$ according to the formula

$$X(t, \omega) = \lim_{r \in D,\ r \downarrow t} X(r, \omega).$$

Proposition 5.3.54. X is $L^0(\lambda \times \mathbf{P}^\mu)$-computable of type $[0, \infty) \times \Omega \to E$, uniformly from a name for $\mu \in \mathcal{M}_1(E)$.

Proof. It suffices to show that X is $L^0(\lambda \times \mathbf{P}^\mu)$-computable on $[0, T] \times \Omega$ for an arbitrary fixed T. Fix an initial distribution $\mu \in \mathcal{M}_1(E)$. Let $k \in \mathbb{N}$.

By Theorem 5.3.52, there exist $L^0(\mathbf{P}^\mu)$-computable random variables

$$0 = U_0^k = V_0^k \le U_1^k \le V_1^k \le \cdots \le U_{n(k)}^k \le V_{n(k)}^k$$

and an $L^0(\mathbf{P}^\mu)$-computable event G_k such that

(i) $\mathbf{P}^\mu(G_k \cup \{U_{n(k)}^k \le T\}) < 2^{-k}$, and

(ii) on $\Omega \setminus G$, $d(X(t), X(V_i^k)) < 2^{-k}$ whenever $U_i^k < t < U_{i+1}^k$ for some i.

Define a random variable $X^k : [0, \infty) \times \Omega \to E$ by

$$X^k(t, \omega) = X(V_i^k(\omega), \omega) \quad \text{if } U_i^k(\omega) < t < U_{i+1}^k(\omega) \text{ for some } i < n(k).$$

Evidently, X^k is $L^0(\lambda \times \mathbf{P}^\mu)$-computable (uniformly from k).

Whenever $j > k$, $d(X^j, X^k) < 2^{-j} + 2^{-k} < 2^{-k+1}$ except possibly on the set $B = [0, T] \times (G_j \cup G_k)$. But $(\lambda \times \mathbf{P}^\mu)(B) < 2^{-k+1}T$, so $(X^k)_{k \in \mathbb{N}}$ converges effectively almost surely with respect to $\lambda \times \mathbf{P}^\mu$. Evidently, the limit is X, which proves that X is $L^0(\lambda \times \mathbf{P}^\mu)$-computable. □

Proposition 5.3.55. For all $t \geq 0$, $X(t) = X(t, \cdot)$ is $L^0(\mathbf{P}^\mu)$-computable from Ω to E, uniformly from names for t and $\mu \in \mathcal{M}_1(E)$.

Proof. Fix t. As usual, $\mathbf{P}^\mu = \mu\kappa$. Suppose $(\epsilon_k)_{k\in\mathbb{N}}$ is a convergent sequence of positive reals. By Lemma 5.3.51, let $(r_k)_{k\in\mathbb{N}} \uparrow t$ and $(s_k)_{k\in\mathbb{N}} \downarrow t$ be sequences in D such that for all $k \in \mathbb{N}$, $\mathbf{P}^\mu(A_k) < \epsilon_k$, where

$$A_k = \{\epsilon_k < \sup\{d(X(r_k), X(s)) : s \in [r_k, s_k] \cap D\}\}.$$

Let $A = \liminf A_k = \bigcup_{n\in\mathbb{N}} \bigcap_{k\geq n} A_k$. Then $\mathbf{P}^\mu(A) = 0$ and X_t is defined except possibly on A. Furthermore, $d(X(t), X(r_k)) \leq \epsilon_k$ except possibly on A_k, so $(X(r_k))_{k\in\mathbb{N}}$ converges effectively almost surely to $X(t)$ with respect to $\mathbf{P}^\mu$; in particular, $X(t)$ is $L^0(\mathbf{P}^\mu)$-computable. □

Proposition 5.3.56. $\omega \mapsto X(\cdot, \omega)$ is $L^0(\mathbf{P}^\mu)$-computable of type $\Omega \to \mathbf{D}_E[0, T]$, uniformly from a dyadic rational $T > 0$ and a name for $\mu \in \mathcal{M}_1(E)$.

Proof. Fix T. Let $\epsilon \in (0, 1)$. Compute, as in Theorem 5.3.52,

$$0 = U_0 = V_0 \leq U_1 \leq V_1 \leq \cdots \leq U_n \leq V_n.$$

Then define a random variable $Z : \Omega \to \mathbf{D}_E[0, T]$ by the formula

$$Z(\cdot)(t) = X(V_i) \quad \text{if } t \in [U_i, U_{i+1}) \text{ for some } i \leq n,$$

and $Z(\cdot)(t) = 0$ by convention if $t > U_n$. By Theorem 3.3.26, Z is $L^0(\mathbf{P}^\mu)$-computable. Moreover, $d_{\mathbf{P}^\mu}(Z, X) < 2\epsilon + 4\epsilon$ by Theorem 5.3.52. Since ϵ was arbitrary, $\omega \mapsto X(\cdot, \omega)$ is $L^0(\mathbf{P}^\mu)$-computable. □

Actually, we still need to verify that the extension of X to $[0, \infty)$ is a Markov process, but this can be done classically with a continuity argument; see Chan [19]. As an illustration, we re-check Proposition 5.3.15:

Fact 5.3.57. For all $\mu \in \mathcal{M}_1(E)$ and $t \geq 0$, $\mathbf{P}^\mu_{X_t} = \mu P_t$.

Proof. Note that this result does not need to be effective in any way. Let $f \in \mathbf{C}(E)$. (Recall that E is compact.) By a change of variables,

$$\mathbf{P}^\mu_{X_t} f = \mathbf{P}^\mu(f \circ X_t) = \mu\kappa(f \circ X_t).$$

By Proposition 5.3.55, X is stochastically continuous. Classically, for $r > t$ close to t, $\mu\kappa(f \circ X_t)$ is close to $\mu\kappa(f \circ X_r)$. By Proposition 5.3.15, $\mu\kappa(f \circ X_r) = \mu(P_r f)$. It is also true classically that $\mu(P_r f)$ is close to $\mu(P_t f)$ for r close to t. Since $f \in \mathbf{C}(E)$ was arbitrary, this completes the proof. □

We also need to verify the Markov property by re-checking Proposition 5.3.38, but this is covered by the next two facts. As with Fact 5.3.57, they can be proved classically with a continuity argument [19].

Fact 5.3.58. For all $s \geq 0$, the (generalized) left-shift operator

$$\theta_s(\omega) : t \mapsto X(t+s, \omega) \tag{5.3.12}$$

is measurable from Ω to Ω.

Fact 5.3.59. Let $s \geq 0$, $\mu \in \mathcal{M}_1(E)$, and $Z \in L^1(\Omega, \mu P_s \kappa)$. The Markov property holds:

$$\mathbf{P}^\mu[Z \circ \theta_s | \mathscr{F}_s] = \kappa^{X_s}(Z) \quad (\mathbf{P}^\mu\text{-a.e.}), \tag{5.3.13}$$

where $\mathscr{F}_s$ is the Borel family

$$\{g(X(t_1), \ldots, X(t_n)) : g \in \mathbf{C}(E^n) \text{ and } t_1 < \cdots < t_m \leq s\}.$$

Once Facts 5.3.58 and 5.3.59 have been verified, Theorem 5.2.15 is just the sum of Propositions 5.3.54, 5.3.55, and 5.3.56.

Chapter 6

Future directions

In the introduction, we stated that the purpose of this dissertation was first to develop a computational framework for the study of stochastic processes with càdlàg sample paths, then within that framework to effectivize three important classical results. This we have accomplished:

(i) Theorem 4.2.16 states that a càdlàg modification of a Lévy process is computable from a suitable representation of the process

(ii) According to Theorem 4.3.19, the Lévy-Itô decomposition of a Lévy process is also computable from such a representation

(iii) By Theorem 5.2.15, a càdlàg modification of a Feller process, too, is computable from a suitable representation of that type of process

And we picked up a few other results along the way.

At this point, with our framework in place, and with some idea of how to effectivize (some parts of) stochastic process theory, there are a number of promising directions for future research. In this chapter, we outline two: an effective theory of càdlàg martingales, and algorithmic randomness for Lévy processes.

6.1 Càdlàg martingales

Markov processes, which we already discussed in detail, and martingales, which we did not, are two of the largest and most important classes of stochastic processes (in continuous time). A Markov process is memoryless: the distribution of its future state depends, not on its entire past, but only on its present state. A martingale, on the other hand, tends to stay put:

its expected future state, given its present state, is that same state. Both properties—Markov and martingale—are defined in terms of a filtration.

To be more precise: let $X = \{X_t : t \geq 0\}$ be an $\mathbb{R}$-valued integrable stochastic process on $\langle \Omega, \mathbb{P} \rangle$, adapted to a filtration $\{\mathscr{F}_t : t \geq 0\}$. X is said to be a *martingale* if for all $s \leq t$,

$$\mathbb{P}[X_t | \mathscr{F}_s] = X_s. \tag{6.1.1}$$

Much of martingale theory generalizes to sub- and supermartingales, so it is worth defining those as well: X is said to be a *submartingale* if for all $s \leq t$,

$$\mathbb{P}[X_t | \mathscr{F}_s] \geq X_s, \tag{6.1.2}$$

and a *supermartingale* if for all $s \leq t$,

$$\mathbb{P}[X_t | \mathscr{F}_s] \leq X_s. \tag{6.1.3}$$

For a proper introduction to martingales, see Revuz and Yor [63].

Now, recall that a Feller process is just a Markov process that satisfies a very reasonable assumption (strong continuity), which is enough to ensure that any Feller process has a càdlàg modification. As it turns out, any martingale—or sub- or supermartingale—that satisfies a very reasonable assumption, *right-continuity in probability*, also has a càdlàg modification.

A natural question is whether such a modification is computable from a representation of the martingale. As usual, we need to assume that the sample space is a computable probability space. As for the martingale, stochastic computability is still a very reasonable assumption, and it implies right-continuity in probability. L^1-computability is stronger, but also seems reasonable, given that martingales are integrable. This leads to

Conjecture. Every L^1-computable (sub- or super-) martingale X has an effectively measurable modification $\widetilde{X}$.

It may be possible to effectivize the proof given in Revuz and Yor [63] for the corresponding classical result (Theorem II.2.5 in that source).

In any case, there is an interesting connection to Lévy processes. Recall that Fact 4.1.29 states that every Lévy process has a càdlàg modification, and Theorem 4.2.16 effectivizes that fact. Now, our proof of Theorem 4.2.16 had nothing to do with martingales; essentially, it treated a Lévy process as a very special type of Markov process. On the other hand, there is a proof of Fact 4.1.29 (that is, a classical proof) that gets at Lévy processes through martingales; see, for example, Theorem 2.1.8 in Applebaum [3].

Of course, a Lévy process X is not, in general, a martingale. It is, however, closely related to a certain martingale Z (or rather a family of martingales Z_u where $u \in \mathbb{R}$). Classically, if we can find a càdlàg modification of Z, we can turn that into a càdlàg modification of X.

Fact 6.1.1. If X is a (one-dimensional) Lévy process, then for all $u \in \mathbb{R}$,

$$Z(t) = Z_u(t) = \exp(\mathbf{i}uX(t) - t\eta(u)) \tag{6.1.4}$$

defines a $\mathbb{C}$-valued martingale $Z = Z_u = \{Z(t) : t \geq 0\}$, where η is the Lévy symbol (Fact 4.1.23)

We can say something about the computability of Z:

Proposition 6.1.2. If X is a stochastically computable Lévy process, then $Z = Z_u$ is L^1-computable, uniformly from a name for u, for all $u \in \mathbb{R}$.

Proof. The following is uniform from names for $u \in \mathbb{R}$ and $t \geq 0$.

By hypothesis, $X(t)$ is $L^0(\mathbb{P})$-computable. By Lemma 4.2.4, $\eta : \mathbb{R} \to \mathbb{C}$ is computable. It follows immediately that $Z(t)$ is $L^0(\mathbb{P})$-computable. Since $Z(t)$ is also bounded in absolute value, Fact 2.3.55 implies that $Z(t)$ is actually $L^1(\mathbb{P})$-computable, as required. □

In light of Proposition 6.1.2, it seems likely that our conjecture would provide another, martingale-based proof of Theorem 4.2.16. In any case, an effective theory of martingales is a promising area for future research.

6.2 Algorithmic randomness

In the introduction, we mentioned that computable aspects of Brownian motion have been studied in the context of algorithmic randomness, and that our own project was conceived as an extension of that theory to the entire class of Lévy processes. We will now make this as precise as possible.

Let $\langle X, \mu \rangle$ be a computable probability space.

Definition 6.2.1. A *Martin-Löf test* (with respect to μ) is a sequence $(U_n)_{n\in\mathbb{N}}$ of uniformly Σ^0_1 sets such that for all $n \in \mathbb{N}$, $\mu(U_n) < 2^{-n}$.

Note that if $(U_n)_{n\in\mathbb{N}}$ is a Martin-Löf test, then $\bigcap_{n\in\mathbb{N}} U_n$ is a Π^0_2 null set. Essentially, then, a Martin-Löf test is an effectively null Π^0_2 set.

Definition 6.2.2. Let $(U_n)_{n\in\mathbb{N}}$ be a Martin-Löf test. A point $x \in X$ is said to *pass* the test if $x \notin \bigcap_{n\in\mathbb{N}} U_n$.

Tests, and null sets in general, represent statistically unusual properties, whereas random points ought to be statistically typical:

Definition 6.2.3. x is said to be *Martin-Löf random* or *MLR* (with respect to μ) if it passes every Martin-Löf test.

For a proper introduction to algorithmic randomness, the reader should really consult Nies [60] or Downey and Hirschfeldt [28].

Now suppose X is a stochastically computable Lévy process on $\langle \Omega, \mathbb{P} \rangle$. Without loss of generality, by Theorem 4.2.16, X is effectively measurable.

For all dyadic rational $T > 0$, $\omega \mapsto X(\cdot, \omega)$ defines an $L^0(\mathbb{P})$-computable function of type $\Omega \to \mathbf{D}[0,T]$, whose values are the sample paths of X restricted to $[0,T]$. For simplicity, and without any real loss of generality, let $T = 1$. Then X induces a probability measure $\mathbb{P}_X$ (technically, $\mathbb{P}_{X|1}$) on $\mathbf{D}[0,1]$. This measure is computable (Fact 2.3.46).

Definition 6.2.4. A Martin-Löf random point in $\langle \mathbf{D}[0,1], \mathbb{P}_X \rangle$ is said to be a *Martin-Löf random sample path* of the Lévy process X.

The properties of Martin-Löf random sample paths of Brownian motion have already been studied extensively by Fouché [32], Kjøs-Hanssen and Nerode [49], and Allen, Bienvenu, and Slaman [2], among other authors. An extension of their work to Lévy processes—cataloging all the properties of these statistically typical sample paths—is now possible.

A key tool will be the *conservation of randomness* [45]:

Fact 6.2.5. Let $\langle X, \mu \rangle$ be a computable probability space and let Y be a computable metric space. Let $f : X \to Y$ be μ-a.e. computable. If $x \in X$ is MLR with respect to μ, then $f(x)$ is MLR with respect to μ_f.

This fact, which pushes randomness forward onto another space, already implies a number of basic results about Martin-Löf random sample paths of Lévy processes. For example, we can use it to sketch a proof of

Proposition 6.2.6. An MLR sample path of a Lévy process can only jump at MLR times (with respect to Lebesgue measure on $[0,1]$).

Proof. By Proposition 4.3.8, we can compute a decreasing sequence (ϵ_k) of positive reals, converging to 0, such that $\nu\{\pm\epsilon_k : k \in \mathbb{N}\} = 0$.

Obviously, the size of each jump lies between ϵ_j and ϵ_j for some j and k. For a fixed j and k, let $\Psi(x) = \langle t_1, \ldots, t_k \rangle$ where $0 < t_1 < \cdots < t_k \leq T$ are exactly the times when $0 < \epsilon_j < \|\mathcal{J}_x(t)\| < \epsilon_k$. By Proposition 3.3.25, $\Psi : \mathbf{D}_V[0,T] \to [0,T]^*$ is $\mathbb{P}_X$-a.s. computable.

One application of the conservation of randomness (Fact 6.2.5) tells us that each time t such that $\epsilon_j < \|\mathcal{J}_x(t)\| < \epsilon_k$ will be Martin-Löf random—but with respect to $\mathbb{P}_\Psi(X)$.

Now, this image measure is not the usual Lebesgue measure on the time interval $[0, 1]$. However, because so-called large jumps (in this case, bounded below by ϵ_j) always occur as a compound Poisson process, i.e., at a constant rate, according to a Poisson distribution, it is not difficult to show that the $\mathbb{P}_\Psi(X)$-MLR reals are exactly the MLR reals for Lebesgue measure. □

What else can we say about Martin-Löf random sample paths of Lévy processes? Any non-trivial property that holds, classically, with probability one is probably worth investigating. And what about the sample paths of Feller processes? And martingales? And what about Schnorr randomness?

Unfortunately, we cannot explore any of these exciting possibilities in this dissertation. All questions are open.

References

[1] Kelty Ann Allen. "Martin-Löf Randomness and Brownian Motion". PhD thesis. University of California, Berkeley, 2014.

[2] Kelty Allen, Laurent Bienvenu, and Theodore A. Slaman. "On zeros of Martin-Löf random Brownian motion". In: *Journal of Logic and Analysis* 6.9 (2014), pp. 1–34.

[3] David Applebaum. *Levy Processes and Stochastic Calculus*. 2nd ed. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2009.

[4] E. A. Asarin and A. V. Pokrovskii. "Use of the Kolmogorov complexity in analyzing control system dynamics". In: *Automation and Remote Control* 47.21–28 (1986).

[5] Jeremy Avigad and Vasco Brattka. "Computability and analysis: The legacy of Alan Turing". In: *Turing's Legacy: Developments from Turing's Ideas in Logic*. Ed. by Robert Downey. Lecture Notes in Logic. Cambridge University Press, 2014.

[6] Laurent Bienvenu et al. "Algorithmic tests and randomness with respect to a class of measures". In: *Proceedings of the Steklov Institute of Mathematics* 274.1 (2011), pp. 34–89.

[7] Patrick Billingsley. *Convergence of Probability Measures*. 2nd ed. Wiley Series in Probability and Statistics. Wiley-Interscience, 1999.

[8] Paul Bilokon and Abbas Edalat. "A domain-theoretic approach to Brownian motion and general continuous stochastic processes". In: *Logic in Computer Science*. 15. CSL-LICS 2014. ACM, 2014.

[9] Errett Bishop. *Foundations of Constructive Analysis*. McGraw-Hill, 1967.

[10] Errett Bishop and Henry Cheng. *Constructive Measure Theory*. AMS Memoirs 116. American Mathematical Society, 1972.

[11] Jens Blanck. "Effective domain representations of $\mathscr{H}(X)$, the space of compact subsets". In: *Theoretical Computer Science* 219 (1999), pp. 19–48.

[12] Stefano Bonaccorsi and Enrico Priola. "From Brownian motion to stochastic differential equations". In: *Brownian motion and stochastic differential equations: 10th International Internet Seminar*. International School on Evolution Equations. 2006.

[13] Volker Bosserhoff. "Notions of probabilistic computability on represented spaces". In: *Electronic Notes in Theoretical Computer Science* 202 (2008), pp. 137–170.

[14] Vasco Brattka. "Recursive and Computable Operations over Topological Structures". PhD thesis. The University of Hagen, 1999.

[15] Vasco Brattka. "Computable versions of Baire's category theorem". In: *Mathematical Foundations of Computer Science*. Lecture Notes in Computer Science 2136. MFCS 2001. Springer, 2001, pp. 224–235.

[16] Vasco Brattka and Gero Presser. "Computability on subsets of metric spaces". In: *Theoretical Computer Science* 305.1-3 (2003), pp. 43–76.

[17] Vasco Brattka and Klaus Weihrauch. "Computability on subsets of Euclidean space. I. Closed and compact subsets". In: *Theoretical Computer Science* 219.1–2 (1999), pp. 65–93.

[18] Yuen-Kwok Chan. "A constructive approach to the theory of stochastic processes". In: *Transactions of the American Mathematical Society* 165 (1972), pp. 37–44.

[19] Yuen-Kwok Chan. "Notes on constructive probability theory". In: *Annals of Probability* 2.1 (1974), pp. 51–75.

[20] Yuen-Kwok Chan. "On some open problems in constructive probability theory". In: *Constructive Mathematics*. Ed. by Fred Richman. Lecture Notes in Mathematics 873. Springer, 1981, pp. 44–53.

[21] Donald L. Cohn. *Measure Theory*. Birkhäuser, 1980.

[22] Pieter Collins. "Computable stochastic processes". In: *arXiv.org* (2014). Submitted. arXiv: `1409.4667v1` (math.PR).

[23] John B. Conway. *A Course in Functional Analysis*. 2nd ed. Graduate Texts in Mathematics 96. Springer, 1990.

[24] Nigel Cutland. *Computability: An Introduction to Recursive Function Theory*. Cambridge University Press, 1980.

[25] George Davie. "The Borel-Cantelli lemmas, probability laws and Kolmogorov complexity". In: *Annals of Probability* 29.4 (2001), pp. 1426–1434.

[26] George Davie and Willem L. Fouché. "On the computability of a construction of Brownian motion". In: *Mathematical Structures in Computer Science* 23.6 (2013), pp. 1257–1265.

[27] El Hadj Aly Dia. "Error bounds for small jumps of Levy processes". In: *Advances in Applied Probability* 45.1 (2013), pp. 86–105.

[28] Rodney G. Downey and Denis R. Hirschfeldt. *Algorithmic Randomness and Complexity*. Theory and Applications of Computability. Springer, 2010.

[29] Abbas Edalat. "Domain theory and integration". In: *Theoretical Computer Science* 151.1 (1995), pp. 163–193.

[30] Abbas Edalat. "Domain theory in stochastic processes". In: *Logic in Computer Science*. LICS 1995. IEEE, 1995, pp. 244–254.

[31] Abbas Edalat. "A computable approach to measure and integration theory". In: *Information and Computation* 207.5 (2009), pp. 642–659.

[32] Willem L. Fouché. "Arithmetical representations of Brownian motion I". In: *Journal of Symbolic Logic* 65.1 (2000), pp. 421–442.

[33] Willem L. Fouché. "The descriptive complexity of Brownian motion". In: *Advances in Mathematics* 155.2 (2000), pp. 317–343.

[34] Willem L. Fouché. "Dynamics of a generic Brownian motion: Recursive aspects". In: *Theoretical Computer Science* 394 (2008), pp. 175–186.

[35] Willem L. Fouché. "Fractals generated by algorithmically random Brownian motion". In: *Mathematical Theory and Computational Practice*. Lecture Notes in Computer Science 5635. CiE 2009. Springer, 2009, pp. 208–217.

[36] Willem L. Fouché. "Diophantine properties of Brownian motion: recursive aspects". In: *Logic, Computation, Hierarchies*. Ed. by Vasco Brattka, Hannes Diener, and Dieter Spreen. Ontos Mathematical Logic 4. De Gruyter, 2014, pp. 139–156.

[37] Willem L. Fouché, Safari Mukeru, and George Davie. "Fourier spectra of measures associated with algorithmically random Brownian motion". In: *Logical Methods in Computer Science* 10.3 (2014), pp. 1–24.

[38] Cameron E. Freer and Daniel M. Roy. "Computable de Finetti measures". In: *Annals of Pure and Applied Logic* 163.5 (2012), pp. 530–546.

[39] Peter Gács. "Uniform test of algorithmic randomness over a general space". In: *Theoretical Computer Science* 341.1–3 (2005), pp. 91–137.

[40] Peter Gács, Mathieu Hoyrup, and Cristóbal Rojas. "Randomness on computable probability spaces—a dynamical point of view". In: *Symposium on Theoretical Aspects of Computer Science*. Dagstuhl Seminar Proceedings 9001. STACS 2009. Schloss Dagstuhl, 2009, pp. 469–380.

[41] Peter Gács, Mathieu Hoyrup, and Cristóbal Rojas. "Randomness on computable probability spaces, a dynamical point of view". In: *Theory of Computing Systems* 48.3 (2011), pp. 465–485.

[42] Geoffrey Grimmett and David Stirzaker. *Probability and Random Processes*. 3rd ed. Oxford University Press, 2001.

[43] Peter Hertling and Klaus Weihrauch. "Randomness spaces". In: *Automata, Languages and Programming*. Lecture Notes in Computer Science 1443. ICALP 1998. Springer, 1998, pp. 796–807.

[44] Peter Hertling and Klaus Weihrauch. "Random elements in effective topological spaces with measure". In: *Information and Computation* 181.1 (2003), pp. 32–56.

[45] Mathieu Hoyrup and Cristóbal Rojas. "An application of Martin-Löf randomness to effective probability theory". In: *Mathematical Theory and Computational Practice*. Lecture Notes in Computer Science 5635. CiE 2009. Springer, 2009, pp. 260–269.

[46] Mathieu Hoyrup and Cristóbal Rojas. "Applications of effective probability theory to Martin-Löf randomness". In: *Automata, Languages and Programming*. Lecture Notes in Computer Science 5555. ICALP 2009. Springer, 2009, pp. 549–561.

[47] Mathieu Hoyrup and Cristóbal Rojas. "Computability of probability measures and Martin-Löf randomness over metric spaces". In: *Information and Computation* 207.7 (2009), pp. 830–847.

[48] Alexander Kechris. *Classical Descriptive Set Theory*. Springer, 1995.

[49] Bjorn Kjøs-Hanssen and Anil Nerode. "The law of the iterated logarithm for algorithmically random Brownian motion". In: *Logical Foundations of Computer Science*. Lecture Notes in Computer Science 4514. LFCS 2007. Springer, 2007, pp. 310–317.

[50] Bjorn Kjøs-Hanssen and Anil Nerode. "Effective dimension of points visited by Brownian motion". In: *Theoretical Computer Science* 410 (2009), pp. 347–354.

[51] Bjorn Kjøs-Hanssen, Paul Kim Long V. Nguyen, and Jason M. Rute. "Algorithmic randomness for Doob's martingale convergence theorem in continuous time". In: *Logical Methods in Computer Science* 10.4 (2014). Paper 12, pp. 1–35.

[52] Bjorn Kjøs-Hanssen and Tamás Szabados. "Kolmogorov complexity and strong approximation of Brownian motion". In: *Proceedings of the American Mathematical Society* 139.9 (2011), pp. 3307–3316.

[53] Christoph Kreitz and Klaus Weihrauch. "Theory of representations". In: *Theoretical Computer Science* 38 (1985), pp. 35–53.

[54] Mark Mandelkern. "Metrization of the one-point compactification". In: *Proceedings of the American Mathematical Society* 107.4 (1989), pp. 1111–1115.

[55] Kenshi Miyabe. "L^1-computability, layerwise computability and Solovay reducibility". In: *Computability* 2 (2013), pp. 15–29.

[56] Takakazu Mori, Yoshiki Tsujii, and Mariko Yasugi. "Computability structures on metric spaces". In: *Combinatorics, Complexity, and Logic*. Springer Series in Discrete Mathematics and Theoretical Computer Science. DMTCS 1996. Springer, 1997, pp. 351–362.

[57] Safari Mukeru. "The descriptive complexity of stochastic integration". In: *Journal of Complexity* 31.1 (2015), pp. 57–74.

[58] Norbert Th. Müller. "Computability on random variables". In: *Theoretical Computer Science* 219 (1999), pp. 287–299.

[59] James R. Munkres. *Topology*. 2nd ed. Pearson, 1999.

[60] André Nies. *Computability and Randomness*. Oxford University Press, 2009.

[61] David Pollard. *A User's Guide to Measure Theoretic Probability*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2002.

[62] Marian B. Pour-El and J. Ian Richards. *Computability in Analysis and Physics*. Perspectives in Mathematical Logic. Springer, 1989.

[63] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*. 3rd ed. A Series of Comprehensive Studies in Mathematics. Springer, 1999.

[64] Hartley Rogers Jr. *Theory of Recursive Functions and Effective Computability*. MIT Press, 1987.

[65] L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes, and Martingales*. 2nd ed. Vol. 1. Cambridge Mathematical Library. Cambridge University Press, 1994.

[66] Walter Rudin. *Principles of Mathematical Analysis*. 3rd ed. McGraw-Hill, 1976.

[67] Jason Rute. "Topics in Algorithmic Randomness and Computable Analysis". PhD thesis. Carnegie Mellon University, 2013.

[68] Matthias Schröder. "Admissible representations of probability measures". In: *Electronic Notes in Theoretical Computer Science* 167 (2007).

[69] Matthias Schröder and Alex Simpson. "Representing probability measures using probabilistic processes". In: *Journal of Complexity* 22 (2006), pp. 768–788.

[70] A. V. Skorokhod. "Limit theorems for stochastic processes". In: *Theory of Probability and Its Applications* 1.3 (1956), pp. 261–290.

[71] Shinzo Watanabe and Nobuyuki Ikeda. *Stochastic differential equations and diffusion processes*. 2nd ed. North Holland Mathematical Library 24. North-Holland/Kodansha, 1989.

[72] Klaus Weihrauch. "Computability on the probability measures on the Borel sets of the unit interval". In: *Theoretical Computer Science* 219 (1999), pp. 421–437.

[73] Klaus Weihrauch. *Computable Analysis: An Introduction*. Texts in Theoretical Computer Science (EATCS). Springer, 2000.

[74] Klaus Weihrauch and Nazanin Tavana-Roshandel. "Representations of measurable sets in computable measure theory". In: *Logical Methods in Computer Science* 10.3 (2014).

[75] Yongcheng Wu. "Computable Measure Theory". PhD thesis. Nanjing University, 2005.

[76] Yongcheng Wu. "Computability on random events and variables in a computable probability space". In: *Theoretical Computer Science* 460 (2012), pp. 54–69.

[77] Yongcheng Wu and Decheng Ding. "Computability of measurable sets via effective metrics". In: *Mathematical Logic Quarterly* 51.6 (2005), pp. 543–559.

[78] Yongcheng Wu and Decheng Ding. "Computability of measurable sets via effective topologies". In: *Archive for Mathematical Logic* 45.3 (2006), pp. 365–379.

[79] Yongcheng Wu and Klaus Weihrauch. "A computable version of the Daniell-Stone theorem on integration and linear functionals". In: *Theoretical Computer Science* 359 (2006), pp. 28–42.

[80] Mariko Yasugi, Takakazu Mori, and Yoshiki Tsujii. "Effective properties of sets and functions in metric spaces with computability structure". In: *Theoretical Computer Science* 219 (1999), pp. 467–486.

Vita

Adrian Maler

Born: September 2, 1985 in Ottawa, Ontario, Canada

Education

PhD	December 2015	The Pennsylvania State University Dissertation Adviser: Stephen G. Simpson
BS	May 2008	University of Ottawa

Publications

1. *Effective Theory of Levy and Feller Processes.*
   Ph.D. dissertation, December 2015.

2. Cell cycle times and the tumour control probability.
   Adrian Maler, Frithjof Lutscher.
   Mathematics in Medicine and Biology 27(4): 313–342, 2010.