Math. Struct. in Comp. Science (2017), *vol.* 27, *pp.* 1287–1314. © Cambridge University Press 2016
doi:10.1017/S0960129516000098 First published online 28 July 2016

On computability and disintegration

NATHANAEL L. ACKERMAN[†], CAMERON E. FREER[‡] and DANIEL M. ROY[§]

[†]*Department of Mathematics, Harvard University, Cambridge, Massachusetts, U.S.A.*
Email: nate@math.harvard.edu
[‡]*Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, and Gamalon Labs, Cambridge, Massachusetts, U.S.A.*
Email: freer@mit.edu
[§]*Department of Statistical Sciences, University of Toronto, Toronto, Ontario, Canada*
Email: droy@utstat.toronto.edu

Received 3 April 2014; revised 3 March 2016

We show that the disintegration operator on a complete separable metric space along a projection map, restricted to measures for which there is a unique continuous disintegration, is strongly Weihrauch equivalent to the limit operator Lim. When a measure does not have a unique continuous disintegration, we may still obtain a disintegration when some basis of continuity sets has the Vitali covering property with respect to the measure; the disintegration, however, may depend on the choice of sets. We show that, when the basis is computable, the resulting disintegration is strongly Weihrauch reducible to Lim, and further exhibit a single distribution realizing this upper bound.

1. Introduction

Conditioning is a basic tool in probability theory and the core operation in Bayesian statistics. Given a pair of random variables θ and D, representing, e.g., an unobserved parameter of interest and some data collected in order to estimate the parameter, the conditional distribution of θ given $D = x$ can be understood as the distribution μ^x that assigns to every measurable set A the probability

$$\mu^x(A) = \frac{\mathbf{P}\{\theta \in A \text{ and } D = x\}}{\mathbf{P}\{D = x\}}, \tag{1}$$

provided that

$$\mathbf{P}\{D = x\} > 0. \tag{2}$$

When μ^x is defined for all x in the range of D, the map $x \mapsto \mu^x$ is called a *disintegration* of the distribution of θ with respect to D.

When property (2) does not hold, Equation (1) is meaningless. In modern probability theory, rather than defining the conditional probability $\mu^x(A)$ for a particular x, the map $x \mapsto \mu^x(A)$ is defined all at once as a function f satisfying

$$\mathbf{P}\{\theta \in A \text{ and } D \in B\} = \int_B f(x)\, \mathbf{P}_D(\mathrm{d}x), \tag{3}$$

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

for all measurable subsets B, where $\mathbf{P}_D := \mathbf{P}\{D \in \cdot\}$ is the distribution of D. The existence of such an f is guaranteed by the Radon–Nikodym theorem. The resulting disintegrations are, however, defined only up to a null set, and so their evaluation at points—namely, at the observed data values, as is statistical practice—has typically relied upon additional (often unstated) hypotheses.

One such hypothesis is the continuity of *some* disintegration, which ensures that any continuous disintegration is canonically defined everywhere in the support of the distribution of the conditioning variables. It is interesting to consider the computability of the conditioning operator in this context.

In Ackerman *et al.* (2010, 2011), we showed that there are computable random variables whose disintegrations are continuous but not computable on any measure one set (indeed, on any set of sufficiently large measure). Here we strengthen these results and make them uniform by providing precise bounds on how non-computable disintegration can be.

In the present paper, we make use of constructive definitions of disintegrations at a point x in terms of limits of the form

$$\lim_{n\to\infty} \frac{\mathbf{P}\{\theta \in A \text{ and } D \in B_n\}}{\mathbf{P}\{D \in B_n\}}. \tag{4}$$

for some sequence of sets B_n 'converging' to x in an appropriate sense. We make repeated use of notions developed by Tjur (1975, 1980). Namely, Tjur introduced a property that may hold of a distribution *at a point* which implies that all sensible choices of B_n lead to the same solution. Furthermore, under certain regularity conditions, if a distribution satisfies Tjur's property at every point, and so can be conditioned at every point, then the resulting disintegration is the unique continuous disintegration — the ideal case for statistical purposes. We show that the disintegration operator on the collection of distributions (on a fixed space) for which the Tjur property holds everywhere is strongly Weihrauch equivalent to the Lim operator.

Tjur's property is rather restrictive, and so we also explore some other attempts to give explicit formulas for disintegrations in terms of limits. Other work in this direction includes Kolmogorov's original axiomatization (Kolmogorov 1933) and the work of Pfanzagl (1979) and of Fraser and Naderi (1996). The latter two papers approach conditioning as the differentiation of a set function. Pfanzagl's definition extends Kolmogorov's own observation that limits in the case of the real line admit conditional probabilities. In general, the resulting disintegrations are only defined up to null sets and the resulting conditional probability will depend on the chosen sequence of sets B_n. Fraser and Naderi give conditions on a collection of sets that imply that there is no such dependence, and we use their notions in Section 5 to show that an everywhere-defined disintegration is strongly Weihrauch reducible to Lim, and moreover, that this upper bound is achieved by a particular disintegration.

Recently, there has been a great deal of progress towards the general program of determining the complexity of various operations in analysis by placing the relevant operators within the hierarchy of Weihrauch degrees (see, e.g., Brattka and Gherardi 2011a,b). The Lim operator plays a central role among the Weihrauch degrees, analogous to that of the halting set among Turing degrees. (For background and various equivalences,

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

see Brattka *et al.* 2012, Section 3.) Viewed within this program, the present paper extends certain consequences of results about Radon–Nikodym derivatives. Namely, a result by Hoyrup *et al.* (2012) (see also Hoyrup and Rojas 2011) can be shown to imply that the map taking a distribution and a real $\varepsilon > 0$ to a continuous disintegration on some $(1 - \varepsilon)$-measure set is Weihrauch reducible to the limit operator Lim, a result which is strengthened by our work.

1.1. *Outline of the paper*

In Section 2, we describe necessary and sufficient conditions for the existence of a continuous disintegration due to Tjur (1975), as well as weaker sufficient conditions for the existence of a disintegration due to Fraser and Naderi (1996). Then in Section 3, we recall the relevant definitions about Weihrauch degrees and represented spaces. We also define and describe some basic properties of the operators EC and Lim, which are key in our arguments (see Definitions 3.23 and 3.25 for their definitions).

Next, in Section 4, we use Tjur's characterization to show that in a natural setting, the conditioning operator, which maps a distribution to a continuous disintegration, is strongly Weihrauch equivalent to Lim. In particular, in Section 4.1, we use a relativization of a simplified variant of the main construction of Ackerman *et al.* (2010) to obtain a lower bound in the Weihrauch degrees of the conditioning operator.

In Section 5, we also consider those measures for which the disintegration is not continuous, and show that under certain conditions, when the disintegration exists everywhere, the disintegration is strongly Weihrauch reducible to Lim. We also exhibit a specific disintegration that realizes this upper bound.

2. Disintegration

In this section, we define the abstract notion of disintegration rigorously. In doing so, we highlight the fact that disintegrations are, in general, not well-defined at points. We also discuss various conditions that allow us to provide a canonical definition at points. Note that these definitions are in terms of probability measures rather than in terms of random variables, as was the case in Section 1.

For a topological space X, write $\mathcal{B}_X$ to denote the Borel σ-algebra on X. When we speak of the *measurable space* X without mentioning the σ-algebra, the Borel σ-algebra is intended.

Definition 2.1. Let S be a topological space and let $\mathcal{M}_1(S)$ denote the collection of Borel probability measures on S. Then the **weak topology** on $\mathcal{M}_1(S)$ is that generated by sets of the form

$$\{\mu \in \mathcal{M}_1(S) \ : \ |\textstyle\int_S \varphi \, \mathrm{d}\mu - c| < \varepsilon\}, \tag{5}$$

where $c, \varepsilon \in \mathbb{R}$, $\varepsilon > 0$, and $\varphi : S \to \mathbb{R}$ is bounded and continuous.

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

Let S and T be measurable spaces. Given a measurable map $g : S \to T$ and probability measure μ on S, define $g[\mu]$ to be the **pushforward** measure on T given by

$$(g[\mu])(B) := (\mu \circ g^{-1})(B) = \mu\big(g^{-1}(B)\big) \tag{6}$$

for all $B \in \mathcal{B}_T$.

Definition 2.2 (Disintegration along a map). Let S and T be measurable spaces, let μ be a probability measure on S, and let g be a measurable function from $S \to T$. Then a **disintegration of μ along** g is a measurable map $\kappa : T \to \mathcal{M}_1(S)$ such that

$$\mu\big(g^{-1}(B) \cap A\big) = \int_B \kappa(x)(A)\ g[\mu](\mathrm{d}x), \tag{7}$$

for all $A \in \mathcal{B}_S$ and $B \in \mathcal{B}_T$.

We will often consider the case where S is a product of two spaces and g is the projection from S onto one of these spaces.

We have chosen to define a disintegration to be a measurable map $\kappa : T \to \mathcal{M}_1(S)$. As is common in probability theory, one might instead have chosen to work with the 'uncurried' *probability kernel* $\kappa^* : (T \times \mathcal{B}_S) \to [0, 1]$ defined by $\kappa^*(t, A) := \kappa(t)(A)$ and assumed to be a probability measure for every fixed first parameter and a measurable function for every fixed second parameter. For more details on the notion of disintegration, see Kallenberg (2002, Ch. 6).

Any two maps κ, κ' satisfying the definition of a disintegration agree on a $g[\mu]$-measure one set, and so we call every such map a *version*. We can speak of *the* disintegration, which is then only determined up to a μ-measure zero set. However, by adding a continuity requirement we are able to pin down a version uniquely. When we speak about the continuity of disintegrations, the topology on the space of measures will be the weak topology.

Lemma 2.3. Let μ be a distribution on S, let $g : S \to T$ be measurable, let $\nu = g[\mu]$, and let κ and κ' be disintegrations of μ along g. If $b \in T$ is a point of continuity of both κ and κ', and $b \in \mathrm{supp}(\nu)$, then $\kappa(b) = \kappa'(b)$. In particular, if κ and κ' are continuous everywhere, they agree on $\mathrm{supp}(\nu)$.

Proof. Assume, towards a contradiction, that $b \in \mathrm{supp}(\nu)$ is a point of continuity of both κ and κ' such that $\kappa(b) \neq \kappa'(b)$. Then there exists some bounded continuous function f such that the function h defined by

$$h(x) := \left| \int f(y)\ \kappa(x)(\mathrm{d}y) - \int f(y)\ \kappa'(x)(\mathrm{d}y) \right| \tag{8}$$

is such that $h(b) > 0$. Because κ and κ' are both disintegrations of μ along g, we have $h(x) = 0$ for ν-almost all x. But h is continuous at b because κ and κ' are presumed to be continuous at b, and so there is an open neighbourhood N of b such that $0 \notin h(N)$. Finally, as $b \in \mathrm{supp}(\nu)$, we have $\nu(N) > 0$, a contradiction. □

In particular, given a measure μ on S, if $\mathrm{supp}(g[\mu]) = T$, there is at most one disintegration of μ along g that is continuous on all of T.

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

The question of the existence of a disintegration has a long history. One of the most important cases is when S is a Borel space, i.e., there exists a measurable bijection, with a measurable inverse, to some Borel subset of the unit interval. The Disintegration Theorem implies that any measure on S has a disintegration along $g : S \to T$ (see, e.g., Kallenberg 2002, Theorems 6.3 and 6.4). Note, however, that the resulting disintegration is only defined up to a null set.

For countable discrete spaces, the notion of disintegration can be given a concrete definition in terms of the elementary notion of conditioning on positive-measure sets.

Definition 2.4 (Conditioning on positive-measure sets). Given a probability measure ν on a topological space S, and a Borel set A of S such that $\nu(A) > 0$, we write

$$\nu^A(\,\cdot\,) := \frac{\nu(\,\cdot \cap A)}{\nu(A)} \tag{9}$$

for **the distribution** ν **conditioned on (the set)** A.

A key relationship between disintegration and conditioning on positive-measure subsets of a countable discrete space T is that the measurable function $\kappa : T \to \mathcal{M}_1(S)$ defined by

$$\kappa(t) = \mu^{g^{-1}(t)}(\,\cdot\,), \tag{10}$$

for every t in the support of $g[\mu]$, (in particular, defined $g[\mu]$-a.e.) is a version of the disintegration of μ along g.

In general, there is no explicit formula like the one for countable discrete spaces. Special cases, e.g., assuming the existence of joint or conditional densities, admit so-called Bayes' rules, but these hypotheses often exclude infinite-dimensional parameter spaces, which are typical of nonparametric Bayesian statistics.

One might hope that the conditional distribution at a point x would be well-approximated by the conditional distribution given a small positive-measure set 'converging' to x. In general, for various natural notions of convergence, this need not hold, but Tjur (1980, Section 9.7) gave conditions for a given probability measure on a point x implying that all 'reasonable' ways of approximating the conditional distribution by positive-measure sets converging to a point x yield the same conditional distribution in the limit. For a fixed measure, when every point possesses this property — which we call the *Tjur property* — it follows that there is a unique continuous disintegration. Indeed, under some mild additional conditions, a measure satisfies Tjur's property at all points exactly when there exists a unique continuous disintegration.

Tjur's conditions are therefore necessarily rather restrictive. Fraser and Naderi (1996) define a less restrictive notion based on the differentiation of set functions. Here, the Vitali covering property of a class of subsets with respect to the measure of interest allows us to explicitly construct a disintegration pointwise. (See Proposition 2.13 for a definition of the Vitali covering property.)

In order to avoid pathologies, all measures in this paper will be Radon. Recall that when S is a Hausdorff space and $\mu \in \mathcal{M}_1(S)$, μ is a **Radon measure** when it is inner

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

regular, i.e., whenever A is a Borel set on S, then

$$\mu(A) = \sup\{\mu(K) \,:\, K \subseteq A \text{ and } K \text{ is compact}\}. \tag{11}$$

It is a standard result that any Borel probability measure on a separable metric space is a Radon measure.

2.1. *The Tjur property*

The following property, which we have named after Tjur, is equivalent to a property described first by Tjur in an unpublished preprint (Tjur 1975) and later in a monograph (Tjur 1980, Section 9.7).

Definition 2.5 (Tjur property). Let S and T be completely regular Hausdorff spaces, let μ be a Radon probability measure on S, and let $g : S \to T$ be a measurable function. Let $x \in T$ be a point in the support of $g[\mu]$, i.e., for every open neighbourhood V of x, we have $g[\mu](V) > 0$. Let $\mathscr{D}(x)$ denote the set of pairs (V, B) where V is an open neighbourhood of x and B is a measurable subset of V with $g[\mu](B) > 0$, and write $(V, B) \preccurlyeq (V', B')$ when $V \subseteq V'$. Note that this relation is a partial ordering on $\mathscr{D}(x)$ and makes $\mathscr{D}(x)$ a downward directed set. We say that x **has the Tjur property (for μ along g)** when the directed limit of the net $(\mu^{g^{-1}(B)} \,:\, (V, B) \in \mathscr{D}(x))$ exists in the weak topology on the space of probability measures.

Remark 2.6. We will sometimes write μ_g^B for $\mu^{g^{-1}(B)}$ when it simplifies notation. We will also write μ_g^x to denote the directed limit when it exists. By the Portmanteau Lemma,

$$\mu_g^x(A) = \lim_{(V,B)\in\mathscr{D}(x)} \mu_g^B(A) \tag{12}$$

for all μ_g^x-continuity sets A. We will write μ^x when g is clear from context.

The usefulness of the Tjur property is demonstrated by the following result, which is a consequence of Proposition 9.9.1, Corollary 9.9.2, and Proposition 9.10.5 of Tjur (1980):

Lemma 2.7. Let S be a completely regular Hausdorff spaces, T a metric space, μ a Radon probability measure on S, and $g : S \to T$ be a measurable function.

Suppose that there is a $g[\mu]$-measure one set C such that every $x \in C$ has the Tjur property (for μ along g). For every $x \in C$, let $\langle B_n^x \rangle_{n\in\mathbb{N}}$ be a sequence of measurable subsets of T such that each B_n^x is contained in the 2^{-n}-ball around x and satisfies $g[\mu](B_n^x) > 0$.

Then the function $\kappa : C \to \mathcal{M}_1(S)$ given by

$$\kappa(x) := \lim_{n\to\infty} \mu_g^{B_n^x} \tag{13}$$

for every $x \in C$ (where the limit is taken in the weak topology) is continuous and can be extended to a disintegration of μ along g.

Many common properties imply that a point is Tjur. For example, every point of continuity of an absolutely continuous distribution is a Tjur point. Also, any isolated point mass (e.g., a point in the support of a discrete random variable taking values in

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

a discrete space) is a Tjur point. On the other hand, nonisolated point masses are not necessarily Tjur points.

Tjur points sometimes exist even in nondominated settings. Let G be a Dirichlet process with an absolutely continuous base measure H on a computable metric space S. That is, G is a random discrete probability distribution on S such that, for every finite, measurable partition $A_1, \ldots, A_k$ of S, the probability vector $(G(A_1), \ldots, G(A_k))$ is Dirichlet distributed with parameter $(H(A_1), \ldots, H(A_k))$. Conditioned on G, let X be G-distributed (i.e., X is a sample from the random distribution G). Then any point $x \in S$ in the support of H is a Tjur point, yet there does not exist a conditional density of G given X.

In Propositions 2.8 and 2.9 we will see that, under some regularity conditions, the existence of an everywhere continuous disintegration is equivalent to every point being Tjur. However, even when every point is Tjur, and thus there exists a continuous disintegration, the resulting disintegration may not be computable. In fact, one of the main constructions of Ackerman *et al.* (2010) is an example of a conditional distribution for which every point is Tjur and yet no disintegration is a computable map, even on large measure set. We will use a related construction in Section 4.

Proposition 2.8 (Tjur 1980, Proposition 9.14.2). Let S and T be completely regular Hausdorff, let μ be a Radon probability measure on S, let $g : S \to T$ be a continuous function, and let $\nu = g[\mu]$ be the pushforward. Suppose

$$\xi : C \to \mathcal{M}_1(S) \tag{14}$$

is a continuous mapping, where C is a subset of $\mathrm{supp}(\nu)$ such that C is a ν-measure one set. Then the following two conditions are equivalent.

i. For all $x \in C$, the conditional distribution μ_g^x is defined and $\mu_g^x = \xi(x)$.
ii. For all $x \in C$, we have $\mathrm{supp}(\xi(x)) \subseteq g^{-1}(x)$, and $\mu = \int_C \xi(x)\, \nu(\mathrm{d}x)$.

We now give conditions for when there is a continuous disintegration on a measure one set.

Proposition 2.9. Let S and T be completely regular Hausdorff, with T second countable, let μ be a Radon probability measure on S, let $g : S \to T$ be a continuous function, and let $\nu := g[\mu]$. Further suppose C is a ν-measure one set. The following are equivalent.

1. Every point in C is a Tjur point of μ along g.
2. $C \subseteq \mathrm{supp}(\nu)$ and there exists a disintegration $\kappa : T \to \mathcal{M}_1(S)$ of μ along g whose restriction to C is continuous.

Proof. Assume (1). Then the fact that there is a continuous disintegration follows from Proposition 2.7.

Next assume that (2) holds and that κ is a continuous disintegration of μ along g. As g is continuous, by Proposition 2.8, it suffices to verify (ii), taking $\xi = \kappa$. First note, by the definition of disintegration,

$$\mu\big(g^{-1}(B) \cap A\big) = \int_B \kappa(x)(A)\; g[\mu](\mathrm{d}x), \tag{15}$$

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

for all $A \in \mathcal{B}_S$ and $B \in \mathcal{B}_T$. But as $\nu(C) = 1$, for all $A \in \mathcal{B}_S$ we have

$$\mu(A) = \mu\big(g^{-1}(C) \cap A\big) = \int_C \xi(y)(A)\ \nu(\mathrm{d}y). \tag{16}$$

In particular, $\mu = \int_C \xi(y)\,\nu(\mathrm{d}y)$, as required. Now let $C^* = \{c \in C\ :\ \mu^c(g^{-1}(c)) = 1\}$. It suffices to show $C^* = C$, as g is continuous.

Assume, towards a contradiction, that there is a $c \in C \setminus C^*$, i.e., such that $\mu^c(g^{-1}(c)) < 1$. As T is second countable, there is a countable decreasing collection of closed neighbourhoods $\langle N_i \rangle_{i\in\omega}$ of c such that $\{c\} = \bigcap_{i\in\omega} N_i$. It follows that $\langle g^{-1}(N_i)\rangle_{i\in\omega}$ is a decreasing countable sequence of sets whose intersection has μ^c-measure strictly less than one, and so, for some j, it holds that $c \in N_j \subseteq T$ and $\mu^c\big(g^{-1}(N_j)\big) < 1$. But as N_j is closed, the map which takes c' and returns $\mu^{c'}\big(g^{-1}(N_j)\big)$ is upper semi-continuous. Therefore there must also be an open set N' with $c \in N' \subseteq T$ such that for all $c' \in N' \cap C$ we have $\mu^{c'}(g^{-1}(N_j)) < 1$. Let N^* be the interior of N_j which is non-empty as N_j is a closed neighbourhood of c. Then

$$N'' := N' \cap N^* \cap C \subseteq C \setminus C^*. \tag{17}$$

Because $N'' \subseteq \mathrm{supp}(\nu)$ is a non-empty open set, $\nu(N'') > 0$. But then

$$\nu(N'') = \int_{N''} \nu(\mathrm{d}y) > \int_{N''} \mu^y\big(g^{-1}(N'')\big)\ \nu(\mathrm{d}y) = \mu(g^{-1}(N'')) = \nu(N''), \tag{18}$$

a contradiction. □

Let $\pi_2 : S \times T \to T$ denote the projection onto T. We now define the space of measures we will use to study the computability of disintegration.

Definition 2.10. For second countable regular Hausdorff spaces S and T, let $\mathcal{C}_{S,T} \subseteq \mathcal{M}_1(S \times T)$ be the subset consisting of those measures μ such that

1. the pushforward of μ along π_2, has full support, i.e., $\mathrm{supp}(\pi_2[\mu]) = T$, and
2. μ admits a (necessarily unique, by Proposition 2.3) continuous disintegration along the projection map $\pi_2 : S \times T \to T$.

It follows from Proposition 2.9 that $\mu \in \mathcal{C}_{S,T}$ if and only if every point of T is a Tjur point for μ along π_2.

Because we are conditioning on a projection map, a disintegration can be identified with a continuous maps $\mathscr{C}(T, \mathcal{M}_1(S))$. Let $\iota_1 : \mathcal{M}_1(S \times T) \to \mathcal{M}_1(S)$ be the map defined by $\iota_1(\mu)(A) = \mu(A \times T)$ for all $\mu \in \mathcal{M}_1(S \times T)$ and $A \in \mathcal{B}_S$. It is easy to check that ι_1 is a continuous map. Let $\mathrm{D}_{S,T} : \mathcal{C}_{S,T} \to \mathscr{C}(T, \mathcal{M}_1(S))$ be defined by $\mathrm{D}_{S,T}(\mu) = \iota_1 \circ \kappa$, where $\kappa\colon T \to \mathcal{M}_1(S \times T)$ is the (unique) continuous disintegration of μ along T. The map ι_1 is injective on the image of $\mathcal{C}_{S,T}$ under $\mathrm{D}_{S,T}$, and so when no confusion can arise, we will also call $\mathrm{D}_{S,T}(\mu)$ the disintegration of μ when $\mu \in \mathcal{C}_{S,T}$.

One often requires that a disintegration be merely a.e. continuous (rather than continuous everywhere). However, for any measure admitting an a.e. continuous disintegration, there is a G_δ subset of T of measure one on which it is continuous everywhere. One could therefore consider all measures on $S \times T$ such that the set of Tjur points contains a G_δ-subset of T of measure one. However one would then have to define the domain

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

of the disintegration map to be the set of pairs consisting of a measure on $S \times T$ along with a G_δ-subset of T. In this context the fundamental results of this paper should still hold with essentially the same proofs; however, the notational complexity would also be greatly increased.

2.2. *Weaker conditions than the Tjur property*

In Section 5 we will consider a distribution which does not admit a continuous disintegration and is thus not in $\mathcal{C}_{S,T}$. To make sense of this we need a weaker notion of disintegration than that given by Tjur. These definitions are based on those from Fraser and Naderi (1996).

Definition 2.11. Suppose T is a separable metric space and ν is a probability measure on T. A sequence of sets $\langle E_n \rangle_{n\in\omega} \in \mathcal{B}_T$ is said to **converge regularly** to x (with respect to ν) if there is a sequence of closed balls $\langle B_n \rangle_{n\in\omega}$ of respective radii $\langle r_n \rangle_{n\in\omega}$ such that

1. $\lim_{n\to\infty} r_n = 0$,
2. $x \in E_n \subseteq B_n$ for all n, and
3. there is an $\alpha > 0$ such that $\mu(E_n) \geqslant \alpha \cdot \mu(B_n)$ for all n.

Intuitively, a sequence of sets converges regularly to x if it is 'close' to a decreasing sequence of closed balls whose intersection is $\{x\}$.

Definition 2.12. For $F \in \mathcal{B}_T$, a class $W \subseteq \mathcal{B}_T$ is a **Vitali cover of** F (with respect to ν) when, for all $x \in F$, there is a sequence of elements of W that converges regularly to x with respect to ν.

Definition 2.13. A class $V \subseteq \mathcal{B}_T$ has the **Vitali covering property** (with respect to ν) when, for every $F \in \mathcal{B}_T$, (1) there is a Vitali cover $W \subseteq V$ of F with respect to ν and (2) for every Vitali cover $W \subseteq V$ of F with respect to ν, there is a collection $\{F_n : n \in \omega\} \subseteq W$ of disjoint sets such that $\mu(F \setminus \bigcup_{n\in\omega} F_n) = 0$.

These definitions allow us to define a notion of disintegration at points:

Proposition 2.14 (Fraser and Naderi 1996, Theorem 2). Let $g : S \to T$, let $\mu \in \mathcal{M}_1(S)$, and let $\nu = g[\mu]$. Suppose that $V \subseteq \mathcal{B}_T$ has the Vitali covering property with respect to ν.

i. For every $A \in \mathcal{B}_S$ there exists $C_A \in \mathcal{B}_T$ with $\nu(C_A) = 1$ such that, for every $x \in C_A$ and every sequence $\langle E_n \rangle_{n\in\omega} \subseteq V$ that converges regularly to x, the limit

$$\lim_{n\to\infty} \frac{\mu(g^{-1}(E_n) \cap A)}{\mu(g^{-1}(E_n))} \tag{19}$$

exists and is independent of the choice of $\langle E_n \rangle_{n\in\omega}$. Let $\mu^x_{g,V}(A)$ denote this limit when it exists and define $\mu^x_{g,V}(A) = \mu(A)$ otherwise.

ii. The map $x \mapsto \mu^x_{g,V}(A)$ is measurable with respect to the ν-completion of $\mathcal{B}_T$ and, for each set E in the ν-completion of $\mathcal{B}_T$,

$$\mu(g^{-1}(E) \cap A) = \int_E \mu^x_{g,V}(A)\ \nu(\mathrm{d}x). \tag{20}$$

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

Remark 2.15. By Proposition 2.7, note that if $x \in T$ is a Tjur point for μ along g, then $\mu^x_{g,V} = \mu^x_g$, and so the left hand side does not depend on V.

Remark 2.16. Because S is assumed to be a separable metric space, we may arrange things so that $A \mapsto \mu^x_{g,V}(A)$ is a probability measure for every x and that $x \mapsto \mu^x_{g,V}$ is a measurable map. (In other words, $(x, A) \mapsto \mu^x_{g,V}(A)$ is a probability kernel.)

The notion of disintegration given by Proposition 2.14 relies heavily on the metric structure of the underlying set. Fortunately, though, there are many situations when there are natural collections with the Vitali property. For example, the class of closed sets in the Lebesgue completion of the Borel subsets of $\mathbb{R}^k$ has the Vitali covering property with respect to Lebesgue measure (see Fraser *et al.* 1995, p. 304 for references to proofs). For our purposes, the following result suffices.

Proposition 2.17. If T is an ultrametric space then the collection of open balls has the Vitali covering property with respect to any measure on T.

Proof. This follows from the fact that every open set is the disjoint union of balls and every open ball is clopen. □

In particular, if the underlying space is a countable product of Cantor space, Baire space, and discrete spaces then the collection of open balls has the Vitali covering property with respect to any measure.

It is worth contrasting Proposition 2.5 with Proposition 2.11. In particular, the notion of a Tjur point can be thought of as requiring that the conditional measures converge at each point irrespective of the manner in which we approximate the point, and specifically does not require the measurable sets in the approximating sequence to actually contain the point they are approximating, so long as each is contained in an open neighbourhood that does. In contrast, the use of regular convergence by Fraser and Naderi is in terms of a fixed class of sets having the Vitali covering property with respect to the given measure, and does require that the sets one is conditioning on contain the point being approximated.

3. Represented spaces and Weihrauch reducibility

We now introduce the notion of a space equipped with a representation and describe some special cases in detail, including spaces of probability measures. We then define some key operators and introduce the notion of a Weihrauch degree in order to compare the computational complexity of operators. For background on represented spaces, see Pauly (2016). A similar treatment can be found in Collins (2009, 2010). Both build off the framework of synthetic topology Escardó (2004). For background on Weihrauch reducibility and Weihrauch degrees, see, e.g., Brattka and Gherardi (2011b), or the introduction of Brattka *et al.* (2015).

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

3.1. *Represented spaces*

In order to define a notion of computability beyond that of the classical setting of Baire space, $\mathbb{N}^{\mathbb{N}}$, we adopt the framework of represented spaces, which is itself a distillation of core ideas from the computable analysis framework connecting back to Turing's original work on computable real numbers. The following definitions are taken from Pauly (2016).

Definition 3.1. A **represented space** is a set X along with a partial surjective function $\delta_X :\subseteq \mathbb{N}^{\mathbb{N}} \to X$, called a **representation**. A point $x \in \mathbb{N}^{\mathbb{N}}$ in the domain of δ_X is said to be a **name** for its image $\delta_X(x)$ in X.

In general, every point in X will have many names: indeed, each represented space (X, δ_X) defines an equivalence relation $\equiv_X$ on $\delta_X^{-1}(X) \subseteq \mathbb{N}^{\mathbb{N}}$ given by $p \equiv_X q$ iff $\delta_X(p) = \delta_X(q)$. For each equivalence relation on a subset of $\mathbb{N}^{\mathbb{N}}$, there is a corresponding represented space.

Functions between represented spaces can be understood as functions mapping names in the domain to names in the codomain.

Definition 3.2. Let (X, δ_X) and (Y, δ_Y) be represented spaces and let $f :\subseteq X \rightrightarrows Y$ be a multi-valued partial function. Then a **realizer** for f is a (single-valued) partial function $F :\subseteq \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}$ such that for all p in the domain of $f \circ \delta_X$,

$$\delta_Y\big(F(p)\big) \in f\big(\delta_X(p)\big). \tag{21}$$

When this holds, we write $F \vdash f$.

Note that a function F is a realizer for a single-valued function if and only if it preserves the equivalence relations associated to the represented spaces.

In what follows we will give explicit representations for only a few fundamental spaces. It is worth mentioning that what is important is not the specific representations but their defining properties: if the defining properties are computable in the appropriate senses, then any two representations that satisfy the defining properties will be computably isomorphic. As such, after introducing only those basic representations that we need to get started, we will aim to avoid mentioning or using the details of a given representation whenever possible. For more on this approach to represented spaces see Pauly (2016) and references therein.

3.1.1. *Continuous functions.* A map between represented spaces is called continuous (computable) if it has a continuous (computable) realizer. The notions of continuous maps between represented spaces and continuous maps between topological spaces are, in general, distinct, and should not be conflated. Once we have introduced the notion of an admissible represented space, however, Proposition 3.8 relates these two notions precisely.

Definition 3.3. Let X and Y be represented spaces. Then the space $\mathscr{C}(X, Y)$ of continuous maps between X and Y can itself be made into a represented space. One such representation of $\mathscr{C}(X, Y)$, which we will use in this paper, is given by $\delta(0^n 1p) = f$ iff the n'th oracle Turing machine computes a realizer for f from oracle $p \in 2^{\omega}$.

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

Intuitively speaking, the name of a map f describes an oracle Turing machine that translates names for inputs in X to names for outputs in Y. Natural operations (evaluation, currying, uncurrying, composition, etc.) on continuous functions are computable (see Pauly 2016, Proposition 3.3).

3.1.2. *The naturals, Sierpinski space, and spaces of open sets.* The representation of sets and relations requires that we introduce two special represented spaces, namely $\mathbb{N} = (\mathbb{N}, \delta_{\mathbb{N}})$ and $\mathbb{S} = (\{0,1\}, \delta_{\mathbb{S}})$. In particular, $\delta_{\mathbb{N}}(0^n 1 0^\omega) = n$ for all $n \in \mathbb{N}$, $\delta_{\mathbb{S}}(0^\omega) = 0$, and $\delta_{\mathbb{S}}(p) = 1$ for $p \neq 0^\omega$. Note that $\mathbb{S}$ should not be confused with the represented space of binary digits $\mathbb{B} = (\{0,1\}, \delta_{\mathbb{B}})$ where $\delta_{\mathbb{B}}(0^\omega) = 0$ and $\delta_{\mathbb{B}}(1^\omega) = 1$. Indeed, it is common to define $\mathbb{S}$ on the two-point space $\{\bot, \top\}$ instead of $\{0,1\}$, however, the latter has the advantage, as we will see, that the EC operator becomes an identity map.

The following basic operations on $\mathbb{S}$ are computable (see Pauly 2016, Proposition 4.1 for details):

1. finite and (i.e., min), $\wedge : \mathbb{S} \times \mathbb{S} \to \mathbb{S}$;
2. finite or (i.e., max), $\vee : \mathbb{S} \times \mathbb{S} \to \mathbb{S}$;
3. countable or (i.e., max), $\bigvee : \mathscr{C}(\mathbb{N}, \mathbb{S}) \to \mathbb{S}$.

Given Sierpinski space, we can define (computable) open sets as those for which membership $\in : X \to \mathbb{S}$ is continuous (computable). Note that, in light of the representation of $\mathbb{S}$, this corresponds to semi-decidability relative to the name of the point in X and the open set.

Definition 3.4. Let X be a represented space. Then the represented space $\mathscr{O}(X)$ of open sets of X is identified with the represented space $\mathscr{C}(X, \mathbb{S})$. In particular, a map $f \in \mathscr{C}(X, \mathbb{S})$ is taken to represent the inverse image $f^{-1}(1)$.

In order to connect this notion with classical computability, note that the computable elements of $\mathscr{O}(\mathbb{N})$ are the computably enumerable subsets of $\mathbb{N}$.

Lemma 3.5 (Pauly 2016, Proposition 4.2). Let X and Y be represented space. Then the following operations are computable:

1. finite intersection, $\cap : \mathscr{O}(X) \times \mathscr{O}(X) \to \mathscr{O}(X)$;
2. finite union, $\cup : \mathscr{O}(X) \times \mathscr{O}(X) \to \mathscr{O}(X)$;
3. countable union, $\bigcup : \mathscr{C}(\mathbb{N}, \mathscr{O}(X)) \to \mathscr{O}(X)$;
4. inverse image, $^{-1} : \mathscr{C}(X, Y) \to \mathscr{C}(\mathscr{O}(Y), \mathscr{O}(X))$;
5. membership, $\in : X \times \mathscr{O}(X) \to \mathbb{S}$;
6. product, $\times : \mathscr{O}(X) \times \mathscr{O}(Y) \to \mathscr{O}(X \times Y)$.

The notion of admissibility connects topological continuity and continuity in the sense of represented spaces. For more details on represented spaces and admissible representations, see Weihrauch (2000, Ch. 3) or Pauly (2016, Section 9), which is based on earlier work in Schröder (2002a, Section 4.3) and Schröder (2002b).

By part (4) of Proposition 3.5, the map $f \mapsto f^{-1}$ is computable. Admissibility can be characterized in terms of the inverse of this map.

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

Definition 3.6. A represented space X is **admissible (computably admissible)** when the map $f \mapsto f^{-1} : \mathscr{C}(Y, X) \to \mathscr{C}(\mathscr{O}(X), \mathscr{O}(Y))$ has a well-defined and continuous (computable) partial inverse for any represented space Y.

Remark 3.7. The previous definition makes use of the following characterization of admissibility, which can be found in Pauly (2016, Theorem 9.11).

Proposition 3.8 (Pauly 2016, Theorem 9.11). A represented space X is admissible if and only if any topologically continuous function $f : Y \to X$ (i.e., one for which the map $f^{-1} : \mathscr{O}(X) \to \mathscr{O}(Y)$ is well-defined) is continuous as a function between represented spaces (i.e., $f \in \mathscr{C}(Y, X)$).

Every representation that we use in this paper is computably admissible.

3.1.3. *Some standard spaces as represented spaces.* We are interested in studying operators defined on certain common mathematical spaces. We first describe how these spaces can be made into computably admissible represented spaces. In particular, the Naturals $\mathbb{N}$, Sierpinski space $\mathbb{S}$, the binary digits $\mathbb{B}$, but also the Reals $\mathbb{R}$ can all be made into computably admissible spaces with respect to their standard topologies. Finite or countable products of these spaces are again computably admissible. (The details are not important for our presentation, but the interested reader can find a thorough account in Weihrauch 2000.)

For the represented space $\mathbb{R}$ of reals, we will require that the map $\mathrm{id}_{\mathbb{Q}} : \mathbb{B} \times \mathbb{N} \times \mathbb{N} \to \mathbb{R}$ defined by $s, a, b \mapsto (-1)^s \cdot a/b$ is computable, as well as the $<$ and $>$ comparison operators $\mathbb{R} \times \mathbb{R} \to \mathbb{S}$. These requirements on the represented space $\mathbb{R}$ already determine several other key properties, including the computability of standard operations on the reals such as addition, subtraction, multiplication, division by a nonzero, and exponentiation.

We will need to define two slightly more exotic represented spaces. We define $\mathbb{R}_<$ to be the represented space with underlying set $\mathbb{R} \cup \{\infty\}$, such that the comparison map $< : \mathbb{R} \times \mathbb{R}_< \to \mathbb{S}$ is computable, and which is computably admissible with respect to the right-order topology generated by the rays $\{(x, \infty] : x \in \mathbb{R}\}$. Similarly we define $\mathbb{R}_>$ to be the represented space with underlying set $\mathbb{R} \cup \{-\infty\}$, such that the comparison map $> : \mathbb{R} \times \mathbb{R}_> \to \mathbb{S}$ is computable, and which is computably admissible with respect to the left-order topology generated by the rays $\{[-\infty, x) : x \in \mathbb{R}\}$. Note that an element of $\mathbb{R}_<$ is a computable point precisely when it is a lower semi-computable real. We will also use $\mathbb{R}^+, \mathbb{R}^+_<, \mathbb{R}^+_>$ for the subspaces with underlying set $(0, \infty]$ and similarly $\mathbb{R}^-, \mathbb{R}^-_<, \mathbb{R}^-_>$ for the subspaces with underlying set $[-\infty, 0)$.

The following standard operations are computable:

1. identity, $\mathrm{id} : \mathbb{R} \to \mathbb{R}_<$ and $\mathrm{id} : \mathbb{R} \to \mathbb{R}_>$, and also $\mathrm{id} : \mathbb{R}_< \times \mathbb{R}_> \to \mathbb{R}$ defined on the diagonal (x, x);
2. comparison, $> : \mathbb{R}_< \times \mathbb{R}_> \to \mathbb{S}$ and $< : \mathbb{R}_> \times \mathbb{R}_< \to \mathbb{S}$;
3. addition, $+ : \mathbb{R}_< \times \mathbb{R}_< \to \mathbb{R}_<$;
4. negation, $- : \mathbb{R}_< \to \mathbb{R}_>$ as well as $- : \mathbb{R}_> \to \mathbb{R}_<$;
5. multiplication by a positive real, $\times_+ : \mathbb{R}_< \times \mathbb{R}^+ \to \mathbb{R}_<$;
6. multiplication by a negative real, $\times_- : \mathbb{R}_< \times \mathbb{R}^- \to \mathbb{R}_>$;

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

7. reciprocation, $^{-1} : \mathbb{R}^+_< \to \mathbb{R}^+_>$, $^{-1} : \mathbb{R}^-_< \to \mathbb{R}^-_>$, $^{-1} : \mathbb{R}^+_> \to \mathbb{R}^+_<$, and $^{-1} : \mathbb{R}^-_> \to \mathbb{R}^-_<$ for nonzero quantities;
8. infinite sums, $\sum : \mathscr{C}(\mathbb{N}, \mathbb{R}^+_<) \to \mathbb{R}_<$.

Besides these concrete spaces, we are interested in admissible representations of complete separable metric spaces.

Definition 3.9. A **computable metric space** is a triple $((S, \delta_S), d_S, s)$ such that (S, d_S) is a complete separable metric space, (S, δ_S) is an admissible represented space, $s \in \mathscr{C}(\mathbb{N}, S)$ is a computable and dense sequence in S, and the following are computable:

1. the distance function, $d_S : S \times S \to \mathbb{R}$;
2. the limit operator $\mathscr{C}(\mathbb{N}, S) \to S$ defined on the set of rapidly converging Cauchy sequences, i.e., defined on $\langle r_n \rangle_{n\in\mathbb{N}} \in \mathscr{C}(\mathbb{N}, S)$ satisfying $|d_S(r_n) - d_S(r_{n+1})| < 2^{-n}$ for all $n \in \mathbb{N}$;
3. the map that takes a sequence $\langle (s_n, q_n) \rangle_{n\in\mathbb{N}} \in \mathscr{C}(\mathbb{N}, \mathbb{N} \times \mathbb{Q})$ to the element $\bigcup_{n\in\mathbb{N}} B(s_n, q_n) = \{x \in S : (\exists n \in \mathbb{N})\ d_S(x, s_n) < q_n\} \in \mathscr{O}(S, \delta_S)$ is computable with a computable multi-valued inverse.

We will omit mention of δ_S and refer to (S, δ_S) by S when no confusion can arise.

Notice that we do not insist that the limit map in the definition of a space be computable on the collection of all Cauchy sequences, but only those sequences which are rapidly converging. In particular, all such rapidly converging Cauchy sequences have a concrete, and hence computable, bound (which tends to zero) on the distance between each element of the sequence and the limit point. We will reconsider the general limit operator on Cauchy sequences in Sections 3.3.2 and 3.3.3.

Note that $(\mathbb{R}, d, \mathbb{Q})$, where $d(x, y) = |x - y|$, is a computable metric space. Another important computable metric space is Cantor space, 2^ω, by which we mean the computable metric space consisting of the space of functions

$$\mathbb{N} \to \{0, 1\} \tag{22}$$

with the usual ultrametric distance, and with dense set consisting of the eventually zero functions enumerated lexicographically.

3.1.4. *Representing probability measures.*

Definition 3.10. Let (S, d_S) be a metric space. For $\varepsilon > 0$ and $A \subseteq S$, let $A^\varepsilon := \{x : d_S(x, A) < \varepsilon\}$. The **Prokhorov metric** $d_{\mathcal{M}_1}$ on $\mathcal{M}_1(S)$ is defined by

$$d_{\mathcal{M}_1}(\mu, \nu) := \inf\{\varepsilon > 0 : \mu(A) \leqslant \nu(A^\varepsilon) + \varepsilon \text{ for every Borel set } A\} \tag{23}$$

for $\mu, \nu \in \mathcal{M}_1(S)$.

Note that the Prokhorov metric generates the weak topology on $\mathcal{M}_1(S)$.

The following result shows how to make the space $\mathcal{M}_1(S)$ of Borel probability measures on a computable metric space S into a computable metric space itself.

Proposition 3.11 (Gács 2005, Appendix B.6). Let (S, d_S, s) be a computable metric space. Then the space of Borel probability measures $\mathcal{M}_1(S)$ is itself a computable metric space

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

under the Prokhorov metric and with dense set

$$\left\{\sum_{i=0}^{k} q_i \delta_{s(m_i)} : k, m_0, \ldots, m_k \in \mathbb{N},\ q_0, \ldots, q_k \in \mathbb{Q}^+ \text{ and } \sum_{i=0}^{k} q_i = 1\right\}, \tag{24}$$

where $\mathbb{Q}^+ := \{q \in \mathbb{Q} : q > 0\}$. An alternative representation of a Borel probability measure is in terms of its restriction to the class of open sets. Such a set function is known as a *valuation*. A valuation on a separable metric space uniquely determines a Borel probability measure (Schröder 2007, Section 3.1).

Definition 3.12. Let X be an admissible represented space. The represented space $\mathcal{V}(X)$ of **valuations on** X is the subspace of $\mathscr{C}(\mathscr{O}(X), \mathbb{R}_<)$ corresponding to restrictions of Borel probability measures to the open sets.

By Schröder (2007, Theorem 3.3), $\mathcal{V}(X)$ is computably admissible when X is. The next result relates the representation of Borel measures as points in the Prokhorov metric to their representation as points in the space of valuations.

Lemma 3.13. Let X be a computable metric space. The identity map from $\mathcal{M}_1(X)$ to $\mathcal{V}(X)$ is computable and so is its inverse.

In other words, a probability measure is computable if and only if, uniformly in the name for an open set, we can lower semi-compute the measure of that open set. See Bosserhoff (2008, Section 2.4) or Hoyrup and Rojas (2009, Theorem 4.2.1) for more on computable measures.

Let μ be a probability measure on a topological space S, let E be a Borel subset of S, and let ∂E denote the boundary of E, i.e., the difference between its closure and its interior. The set E is a **μ-continuity set** when $\mu(\partial E) = 0$.

For a computable metric space X and Borel probability measure $\mu \in \mathcal{M}_1(X)$, let $\mathscr{O}_\mu(X)$ be the class of open sets $U \in \mathscr{O}(X)$ that are also μ-continuity sets. Concretely, one representation for $\mathscr{O}_\mu(X)$, which we will use here, is the map $\delta : \mathbb{N}^\mathbb{N} \to \mathscr{O}_\mu(X)$ such that $\delta(p) = U$ if and only if $p = (u, v)$, where u is a name for U as an element of $\mathscr{O}(X)$ and v is a name an open set V, as an element of $\mathscr{O}(X)$, where $V \subseteq X \setminus U$ and $\mu(V) = \mu(X \setminus U)$. A more elegant approach, which we do not take here, is to define $\mathscr{O}_\mu(X)$ in terms of the extremal property that containment (i.e., the $\in$ relation) is computable on a μ-measure one set. We now describe several computable operations:

Lemma 3.14. Let X and Y be computable metric spaces, and let $\mu, \mu_X \in \mathcal{M}_1(X)$ and $\mu_Y \in \mathcal{M}_1(Y)$. The following maps are computable:

1. $\cap : \mathscr{O}_\mu(X) \times \mathscr{O}_\mu(X) \to \mathscr{O}_\mu(X)$;
2. $\cup : \mathscr{O}_\mu(X) \times \mathscr{O}_\mu(X) \to \mathscr{O}_\mu(X)$;
3. $\times : \mathscr{O}_{\mu_X}(X) \times \mathscr{O}_{\mu_Y}(Y) \to \mathscr{O}_{\mu_X \otimes \mu_Y}(X \times Y)$;
4. $^{\circ\mu} : \mathscr{O}_\mu(X) \rightrightarrows \mathscr{O}_\mu(X)$, which takes an open μ-continuity set U to the set of open μ-continuity sets V contained within, and of the same μ-measure as, $X \setminus U$;
5. $\in : X \times \mathscr{O}_\mu(X) \to \mathbb{S}$;
6. $\mathrm{id} : \mathscr{O}_\mu(X) \to \mathscr{O}(X)$.

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

Proof. That (1), (2), and (5) are computable follows immediately from Proposition 3.5, parts (1), (2), and (5), respectively. The computability of (6) follows trivially from the definition of the representation for $\mathscr{O}_\mu(X)$. The computability of (4) follows from the fact that, if $p = (u, v)$ is a name for U, then (v, u) is a name for an element in $U^{\circ\mu}$. To see that the map (3) is computable let U and V be a μ_X- and μ_Y-continuity set, respectively, and compute $U^* \in U^{\circ\mu_X}$ and $V^* \in V^{\circ\mu_Y}$. Then $(U^* \times V) \cup (U \times V^*) \cup (U^* \times V^*) \in (U \times V)^{\circ(\mu_X \otimes \mu_Y)}$ and is a computable element of $\mathscr{O}(X \times Y)$ by parts (2) and (6) of Proposition 3.5. □

Lemma 3.15. Let X be a computable metric space and let $\mu \in \mathcal{M}_1(X)$. The map taking μ and $U \in \mathscr{O}_\mu(X)$ to $\mu(U) \in \mathbb{R}$ is computable.

Proof. Let $U \in \mathscr{O}_\mu(X)$. Then U and some $V \in U^{\circ\mu}$, are elements of $\mathscr{O}(X)$ computable from U by Proposition 3.14, parts (4) and (6). Therefore $\mu(U) \in \mathbb{R}_\prec$ and $1 - \mu(V) \in \mathbb{R}_\succ$ are computable from U and V, respectively. But $\mu(U) = 1 - \mu(V)$, hence $\mu(U) \in \mathbb{R}$ is computable from U. □

Lemma 3.16 (Bosserhoff 2008, Lemma 2.15 and Hoyrup and Rojas 2009, Lemma 5.1.1). Let $(S, d_S, \langle s_i \rangle_{i\in\mathbb{N}})$ be a computable metric space, let $\mu \in \mathcal{M}_1(S)$ be a Borel probability measure on S, and let B_μ be the set of all sequences $\langle \epsilon_i \rangle_{i\in\mathbb{N}}$ of positive reals such that

$$\{B(s_j, \epsilon_i)\}_{i,j\in\mathbb{N}} \tag{25}$$

forms a subbasis for S comprised of μ-continuity sets. Then the multi-valued map from $\mathcal{M}_1(S) \rightrightarrows \mathscr{C}(\mathbb{N}, \mathbb{R}_+)$ taking μ to B_μ is computable.

Definition 3.17. Let X be a computable metric space, let μ be a Borel probability measure on X, and let $\mathscr{B}_\mu(X) \subseteq \mathscr{C}(\mathbb{N}, \mathscr{O}_\mu(X))$ be the represented space of enumerations of countable bases of X composed of μ-continuity sets, where $\delta(p) = B$ if and only $p = (p', p'')$ where p' is a name for B as an element of $\mathscr{C}(\mathbb{N}, \mathscr{O}_\mu(X))$ and p'' is a name for the (multi-valued) right inverse of the map

$$\bigcup{}^B : \mathscr{C}(\mathbb{N}, \mathbb{N}) \to \mathscr{O}(X) \tag{26}$$

taking U to $\bigcup_{n\in\mathbb{N}} B(U(n))$. Elements of $\mathscr{B}_\mu(X)$ are called μ-continuous bases.

Note that the map $\bigcup^B$ is necessarily computable from B by Proposition 3.5, part (3). We have the following important corollary, first proved by Hoyrup and Rojas (2009, Corollary 5.2.1) in a non-relativized form.

Corollary 3.18. Let X be a computable metric space and let $B \in \mathscr{B}_\mu(X)$ be a μ-continuous basis. Then μ is computable from $\langle \mu(\bigcup_{n\in I} B(n)) \rangle_{I\subseteq\mathbb{N}, |I|<\omega}$ and B.

Proof. By a relativized version of Hoyrup and Rojas (2009, Theorem 4.2.1) it suffices to show that for any open set $U \in \mathscr{O}(X)$, we can lower semi-compute $\mu(U)$ from $\langle \mu(\bigcup_{n\in I} B(n)) \rangle_{I\subseteq\mathbb{N}, |I|<\omega}$, U and B. By definition, we can compute, from U and B, a function $f_U \in \mathscr{C}(\mathbb{N}, \mathbb{N})$ such that $U = \bigcup_{n\in\mathbb{N}} B(f_U(n))$. By the continuity of measures, $\mu(U) = \sup_{n\in\mathbb{N}} \mu(\bigcup_{i\leqslant n} B(f_U(i))$, which is an increasing sequence that is computable from $\langle \mu(\bigcup_{n\in I} B(n)) \rangle_{I\subseteq\mathbb{N}, |I|<\omega}$, B, and U. □

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

Lemma 3.19. Let T be a computable metric space. The multi-valued map taking a Borel probability measure $\mu \in \mathcal{M}_1(T)$ to the set $\mathscr{B}_\mu(T)$ of μ-continuous bases of T is computable.

Proof. Let d_T denote the computable metric on T. By Lemma 3.16, we can compute, from μ, a dense collection $\langle \epsilon_k \rangle_{k\in\mathbb{N}}$ of positive reals such that $\langle B(t_i, \epsilon_k)\rangle_{i,k\in\mathbb{N}}$ is a subbasis composed of μ-continuity sets. Let $V(i,k)$ be the interior of the complement of the open ball $B(t_i, \epsilon_k)$. For all $b \in T$, it holds that $b \in V(i,k)$ if and only if there are $j, k' \in \mathbb{N}$ such that

$$d_T(t_j, b) < \epsilon_{k'} \quad \text{and} \quad d_T(t_{p_0(n)}, t_j) > \epsilon_{p_1(n)} + \epsilon_{k'}, \tag{27}$$

i.e., $b \in B(t_j, \epsilon_{k'}) \subseteq V(i,k)$. The set of all such pairs is computably enumerable from μ, uniformly in i,k, and so $V(i,k)$, viewed as an element of $\mathscr{O}(T)$, is computable from μ, uniformly in i,k. Therefore, $B(t_i, \epsilon_k)$, viewed as an element of $\mathscr{O}_\mu(T)$, is computable from μ, uniformly in i,k. Finite intersection is a computable operation on $\mathscr{O}_\mu(T)$ (by part (1) of Proposition 3.14), and so we can compute an enumeration $B \in \mathscr{C}(\mathbb{N}, \mathscr{O}_\mu(T))$ of a basis for T from μ.

Finally, by the definition of a computable metric space, from every $U \in \mathscr{O}(T)$, we can compute a collection of balls of rational radii whose union is U, and as $\langle \epsilon_k \rangle_{k\in\mathbb{N}}$ is dense and computable from μ, it is clear that the right-inverse of $\bigcup^B$ is computable from μ, and so B, viewed as an element of $\mathscr{B}_\mu(T)$, is computable from μ. □

3.2. *Weihrauch reducibility*

Let $\langle \cdot, \cdot \rangle : \mathbb{N}^\mathbb{N} \times \mathbb{N}^\mathbb{N} \to \mathbb{N}^\mathbb{N}$ be a computable bijection with a computable inverse, and let

$$\mathrm{id} : \mathbb{N}^\mathbb{N} \to \mathbb{N}^\mathbb{N} \tag{28}$$

denote the identity map on $\mathbb{N}^\mathbb{N}$. For $f, g : \mathbb{N}^\mathbb{N} \to \mathbb{N}^\mathbb{N}$, write $\langle f, g\rangle$ to denote the function mapping $x \in \mathbb{N}^\mathbb{N}$ to $\langle f(x), g(x)\rangle \in \mathbb{N}^\mathbb{N}$.

Definition 3.20. Let $\mathcal{F}$ and $\mathcal{G}$ be sets of partial functions from $\mathbb{N}^\mathbb{N}$ to $\mathbb{N}^\mathbb{N}$. Then $\mathcal{F}$ is **Weihrauch reducible** to $\mathcal{G}$, written $\mathcal{F} \leqslant_\mathrm{W} \mathcal{G}$, when there are computable partial functions $H, K :\subseteq \mathbb{N}^\mathbb{N} \to \mathbb{N}^\mathbb{N}$ such that

$$(\forall G \in \mathcal{G})(\exists F \in \mathcal{F})\ F = H \circ \langle \mathrm{id}, G \circ K\rangle. \tag{29}$$

The set $\mathcal{F}$ is **strongly Weihrauch reducible** to $\mathcal{G}$, written $\mathcal{F} \leqslant_\mathrm{sW} \mathcal{G}$ when there are computable partial functions $H, K :\subseteq \mathbb{N}^\mathbb{N} \to \mathbb{N}^\mathbb{N}$ such that

$$(\forall G \in \mathcal{G})(\exists F \in \mathcal{F})\ F = H \circ G \circ K. \tag{30}$$

Recall that $F \vdash f$ means that $F :\subseteq \mathbb{N}^\mathbb{N} \to \mathbb{N}^\mathbb{N}$ is a realizer for a map f between represented spaces.

Definition 3.21. Let f, g be multi-valued functions on (not necessarily the same) represented spaces. Then f is **Weihrauch reducible** to g, written $f \leqslant_\mathrm{W} g$, when

$$\{F : F \vdash f\} \leqslant_\mathrm{W} \{G : G \vdash g\}. \tag{31}$$

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

Two functions f and g are Weihrauch equivalent (written $f \equiv_{\mathrm{W}} g$) when $f \leqslant_{\mathrm{W}} g$ and $g \leqslant_{\mathrm{W}} f$. Weihrauch reducibility is transitive, and the **Weihrauch degree** of a function is the $\equiv_{\mathrm{W}}$-class of the function.

Similarly, $f \leqslant_{\mathrm{sW}} g$ when $\{F : F \vdash f\} \leqslant_{\mathrm{sW}} \{G : G \vdash g\}$, and the **strong Weihrauch degree** of a function is its $\equiv_{\mathrm{sW}}$-class.

When a function f has the property that $\mathrm{id} \times f \equiv_{\mathrm{sW}} f$, where id is the identity on $\mathbb{N}^{\mathbb{N}}$, the function f is called a **cylinder**. We will make use of the following fact on several occasions.

Lemma 3.22 (Brattka and Gherardi 2011b, Corollary 3.6). Let f and g be multi-valued functions on represented spaces, and suppose that f is a cylinder. If $g \leqslant_{\mathrm{W}} f$, then $g \leqslant_{\mathrm{sW}} f$.

3.3. *Operators*

3.3.1. *The disintegration operator.* For computable topological spaces S and T, recall the disintegration map $\mathrm{D}_{S,T} : \mathcal{C}_{S,T} \to \mathscr{C}(T, \mathcal{M}_1(S))$. In what follows, we will write D for $\mathrm{D}_{[0,1],[0,1]}$. One of our main theorems is a characterization of the strong Weihrauch degree of D, showing that $\mathrm{D} \equiv_{\mathrm{sW}} \mathrm{Lim}$. Before formally introducing the operator Lim, we first describe an equivalent operator EC.

3.3.2. *The* EC *operator.* The EC operator can be thought of as taking an enumeration of a set of natural numbers to its characteristic function.

Definition 3.23. EC is the identity operator

$$\mathrm{EC} : \mathscr{C}(\mathbb{N}, \mathbb{S}) \to \mathscr{C}(\mathbb{N}, \mathbb{B}) \tag{32}$$

that maps a function $\mathscr{C}(\mathbb{N}, \mathbb{S})$ to itself in $\mathscr{C}(\mathbb{N}, \mathbb{B})$.

The following result is folklore.

Lemma 3.24. The operator EC is a cylinder.

Proof. By Brattka and Gherardi (2011b, Lemma 6.3), the operator EC is strongly equivalent to the parallelization of LPO, which by Brattka and Gherardi (2011b, Proposition 6.5) is a cylinder. □

Note that EC applied to the enumeration of some set yields an output that is computable from the halting problem relative to that enumeration. Furthermore, every Turing degree contains an enumeration such that applying EC yields a set in the Turing jump of that degree.

3.3.3. *The* Lim *operator.* Limit operators are fundamental in the program of calibrating the difficulty of analytic problems. For several problems within analysis that are equivalent to a limit operator, see Pauly and Fouché (2014, Section 2.3).

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

Furthermore, the limit operator on a sufficiently complicated space is equivalent to EC (see Lemma 3.28), and hence plays a role among the Weihrauch degrees somewhat analogous to that of the halting set within Turing computability.

Definition 3.25. Let S be a metric space, and define $\mathrm{Lim}_S :\subseteq \mathscr{C}(\mathbb{N}, S) \to S$ to be the partial map taking Cauchy sequences to their limit points. Write Lim to denote Lim_{2^ω}.

We will make use of the following result in Section 5.

Lemma 3.26 (Brattka 2005, Proposition 9.1). Suppose that 2^ω computably embeds into a computable metric space S. Then $\mathrm{Lim}_S \equiv_{\mathrm{sW}} \mathrm{Lim}$.

In both Sections 4 and 5, we will use the strong Weihrauch equivalence of EC and Lim.

Lemma 3.27 (Brattka *et al.* 2012, Fact 3.5). The operator Lim is a cylinder.

Lemma 3.28. $\mathrm{EC} \equiv_{\mathrm{sW}} \mathrm{Lim}$.

Proof. By Brattka (2005, Proposition 9.1), we have $\mathrm{EC} \equiv_{\mathrm{W}} \mathrm{Lim}$. (For how this aligns with the definitions above, see Brattka (2005, Definition 7.7) in the case $k = 1$.) But EC and Lim are both cylinders by Lemmas 3.24 and 3.27, and so by Lemma 3.22 we have $\mathrm{EC} \equiv_{\mathrm{sW}} \mathrm{Lim}$. □

There are two key ingredients in the proof of the main theorem of Section 4: $\mathrm{EC} \leqslant_{\mathrm{sW}} \mathrm{D}_{\mathbb{N},[0,1]}$ and $\mathrm{D}_{S,T} \leqslant_{\mathrm{sW}} \mathrm{EC}$ for arbitrary computable metric spaces S and T. In Section 5, we make use of reductions involving both EC and $\mathrm{Lim}_{\mathcal{M}_1(S)}$ for a particular space S.

4. The disintegration operator on distributions admitting a unique continuous disintegration

We now present lower bounds (Section 4.1) and upper bounds (Section 4.2) for the conditioning operator in the context of measures admitting unique continuous disintegrations. The latter makes use of key results in Section 2 pertaining to Tjur points.

4.1. *Lower bound*

The following proposition amounts to a relativized version of the main results of Ackerman *et al.* (2010), which demonstrates that the conditional distribution of a computable random variable given another need not be computable. We thank an anonymous referee for suggesting a construction that greatly simplified the following proof.

Proposition 4.1. $\mathrm{EC} \leqslant_{\mathrm{sW}} \mathrm{D}_{\mathbb{N},[0,1]}$.

Proof. Let $x \in \mathscr{C}(\mathbb{N}, \mathbb{S})$. From x, we can compute a function $y \in \mathscr{C}(\mathbb{N} \times \mathbb{N}, \mathbb{B})$ such that $y(m, \cdot)$ is nondecreasing and, for every $m \in \mathbb{N}$, we have $x(m) = 1$ if and only if there exists an $n \in \mathbb{N}$ such that $y(m, n) = 1$.

Define the function $\iota : \mathbb{N} \to \mathbb{N} \cup \{\infty\}$ by $\iota(m) = \inf\{k < \infty : y(m, k) = 1\}$, with the convention that $\inf \varnothing = \infty$.

Let $f_k, f_\infty : [0, 1] \to [0, 2]$, for $k \in \mathbb{N}$, be defined by $f_\infty(z) = 1$ and $f_k(z) = 1 + \cos(2^{k+1}\pi z)$.

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

Claim 4.2. The measure μ_k with density $f_{\iota(k)}$ with respect to Lebesgue measure on $[0,1]$ is computable from y, uniformly in k.

Proof of Claim. By Proposition 3.18, it suffices to prove that μ_k is computable from y on a Lebesgue-almost-decidable basis which is computably closed under finite unions uniformly in k and for which the union map is computable. Let $i, j, m \in \mathbb{N}$ with $i < j \leqslant 2^m$, let

$$J_{i,j,m} := \left(\frac{i}{2^m}, \frac{j}{2^m}\right) \tag{33}$$

be an open interval with dyadic end points, and let $\mathcal{J}$ be the collection of all finite disjoint unions of such sets, indexed in the obvious way. Then $\mathcal{J}$ is a basis of continuity sets with respect to any measure absolutely continuous with respect to Lebesgue measure. It is straightforward to show that $\mathcal{J}$ is moreover a Lebesgue-almost-decidable basis. Uniformly in $m \in \mathbb{N}$, the functions f_m, f_∞ are computable and so their definite integrals over $\mathcal{J}$ are uniformly computable. It follows that $\mu_k(J_{i,j,m})$ is computable, uniformly in i, j, k, m, for every m in the set $\{m \in \mathbb{N} : m \geqslant \iota(k)\}$, which is itself a computable element of $\mathscr{O}(\mathbb{N})$ from y, uniformly in k. It suffices to show that $\mu_k(J_{i,j,m})$ is computable, uniformly in i, j, k, m for every m in the set $\{m \in \mathbb{N} : m < \iota(k)\}$, which is also a computable element of $\mathscr{O}(\mathbb{N})$ from y, uniformly in k. To see this, note that, for $m < \iota(k)$,

$$\int_{J_{i,j,m}} f_{\iota(k)}(z)\,\mathrm{d}z = 2^{-m}(j-i), \tag{34}$$

completing the proof of the claim. □

We now define μ as the measure on $\mathbb{N} \times [0,1]$ such that $\mu(\{n\} \times \cdot)$ is the measure having density g_n with respect to Lebesgue measure, where

$$g_{2m} := 2^{-m-2} f_{\iota(m)} \quad \text{and} \quad g_{2m+1} := 2^{-m-2}(2 - f_{\iota(m)}). \tag{35}$$

By construction $\sum_{n\in\mathbb{N}} g_n(x) = 1$ for every $x \in [0,1]$, and so μ is a probability measure with $\mathrm{supp}(\pi_2[\mu]) = [0,1]$. Note that these measures are easily seen to be uniformly computable in y from Proposition 4.2, and so, again by Proposition 3.18, it follows that μ is computable.

It follows from an application of Bayes' rule that the disintegration of μ along its second coordinate agrees almost everywhere with the map $\kappa : [0,1] \to \mathcal{M}_1(\mathbb{N})$ given by

$$\kappa(z)\{n\} = g_n(z). \tag{36}$$

Clearly, κ is continuous, and because $\pi_2[\mu]$ has full support, κ is the unique continuous disintegration.

In summary, we have described a computable map $\mathscr{C}(\mathbb{N}, \mathbb{S}) \to \mathcal{C}_{\mathbb{N},[0,1]}$ sending a function x to a measure μ admitting a unique continuous disintegration. Take K to be a realizer for this map. We now show how to compute x as an element in $\mathscr{C}(\mathbb{N}, \mathbb{B})$ from κ, which gives a realizer H such that $H \circ G \circ K \vdash \mathrm{EC}$ for any $G \vdash \mathcal{C}_{\mathbb{N},[0,1]}$.

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

Note that, by Equation (36), the function g_k is computable from κ, uniformly in k. It follows that the functions $f_{\iota(k)}$ are computable from κ, uniformly in k. Observe that

$$x(k) = f_{\iota(k)}(0) - 1 = \begin{cases} 1, & \text{if } \iota(k) < \infty; \\ 0, & \text{if } \iota(k) = \infty. \end{cases} \tag{37}$$

It follows that $x(k)$ is computable from κ, uniformly in k, as desired. □

Recall that D denotes the operator $\mathrm{D}_{[0,1],[0,1]}$. We thank an anonymous referee for suggestions that simplified the following proof.

Corollary 4.3. $\mathrm{D}_{\mathbb{N},[0,1]} \leqslant_{\mathrm{sW}} \mathrm{D}$ and so $\mathrm{EC} \leqslant_{\mathrm{sW}} \mathrm{D}$.

Proof. Let $\alpha : \mathbb{N} \to [0,1]$ be some canonical computable injective map. In particular, assume there exists a computable map $\phi : \mathbb{N} \to \mathscr{O}([0,1])$ such that $\alpha^{-1}[\phi_n] = \{n\}$. Given $\mu \in \mathcal{M}_1(\mathbb{N} \times [0,1])$, define $\nu \in \mathcal{M}_1([0,1]^2)$ to be the pushforward of μ via $\alpha \times \mathrm{id}$. Then $\mathrm{D}_{\mathbb{N},[0,1]}(\mu)(t)\{n\} = \mathrm{D}(\nu)(t)(\phi_n)$, the latter being an element of $\mathbb{R}_<$, computable from $\mathrm{D}_{\mathbb{N},[0,1]}(\mu)$ and t, uniformly in n. This demonstrates $\mathrm{D}_{\mathbb{N},[0,1]} \leqslant_{\mathrm{sW}} \mathrm{D}$. Then $\mathrm{EC} \leqslant_{\mathrm{sW}} \mathrm{D}$ follows from Proposition 4.1. □

4.2. *Upper bound*

We now make use of results about Tjur points from Section 2.1. Recall that for a computable metric space S and $\mu \in \mathcal{M}_1(S)$, we have defined $\mathscr{O}_\mu(S)$ to be the collection of μ-continuity sets in $\mathscr{O}(S)$.

Lemma 4.4. For any computable metric space S, the function taking $\mu \in \mathcal{M}_1(S)$ and $H \in \mathscr{O}_\mu(S)$, satisfying $\mu(H) > 0$, to the probability measure μ^H is computable.

Proof. Let $A \in \mathscr{O}(S)$. Because $H \in \mathscr{O}(S)$, we have $A \cap H \in \mathscr{O}(S)$ and so $\mu(A \cap H) \in \mathbb{R}_<$. But, by Proposition 3.15, $\mu(H) \in \mathbb{R}$ is computable from H, and so $\mu^H(A) = \frac{\mu(A \cap H)}{\mu(H)}$, viewed as an element of $\mathbb{R}_<$, is is computable from H and A, because division of an element of $\mathbb{R}_<$ by an element of $\mathbb{R} \setminus \{0\}$ is computable. □

Recall that $\pi_2 : S \times T \to T$ denotes the projection map.

Corollary 4.5. Let S and T be computable metric spaces, let $\mu \in \mathcal{M}_1(S \times T)$, let μ_T be its projection onto T, and let $U \in \mathscr{O}_{\mu_T}(T)$ be a μ_T-continuity set with $\mu_T(U) > 0$. Then $\mu^U_{\pi_2} = \mu^{S \times U}$ is computable from μ and U.

Proof. We have $\mu(S \times U) = \mu_T(U) > 0$. Moreover, $\mu(\partial(S \times U)) = \mu_T(\partial U) = 0$, and so $S \times U \in \mathscr{O}_\mu(S \times T)$. It follows that $\mu^{S \times U}$ is computable by Proposition 4.4. □

Proposition 4.6. For any two computable metric spaces S and T, we have

$$\mathrm{D}_{S,T} \leqslant_{\mathrm{sW}} \mathrm{EC}. \tag{38}$$

Proof. By Lemma 3.24, the operator EC is a cylinder. Hence by Lemma 3.22, it suffices to show that $\mathrm{D}_{S,T} \leqslant_{\mathrm{W}} \mathrm{EC}$. Let $\mu \in \mathcal{C}_{S,T}$, let μ_T be its projection onto T, and let $d_{\mathcal{M}_1}$ denote the (computable) metric on $\mathcal{M}_1(S)$.

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

By Lemma 3.19, we can compute a μ-continuous basis B from μ. By Proposition 4.5, $\mu_{\pi_2}^{B(n)} \in \mathcal{M}_1(S)$ is computable from μ, uniformly in n, and so, the set F of those triples (m,n,k) such that $d_{\mathcal{M}_1}(\mu_{\pi_2}^{B(m)}, \mu_{\pi_2}^{B(n)}) > 2^{-k}$ is also computably enumerable from μ. Define $\xi \in \mathscr{C}(\mathbb{N}\times\mathbb{N}, \mathbb{S})$ by

$$\xi(n,k) = \begin{cases} 1, & \text{if } \exists m \in \mathbb{N},\ m \geqslant n \text{ and } (m,n,k) \in F, \\ 0, & \text{otherwise.} \end{cases} \tag{39}$$

As an element of $\mathscr{C}(\mathbb{N}\times\mathbb{N}, \mathbb{S})$, note that ξ is computable from I and F.

Let K be the map taking μ to ξ, and define $\hat{\xi} \in \mathscr{C}(\mathbb{N}\times\mathbb{N}, \mathbb{B})$ by $\hat{\xi} := (\mathrm{EC}\circ K)(\mu)$. Because $\mu \in \mathcal{C}_{S,T}$, by Proposition 2.9 every point of T is a Tjur point for μ along π_2. Let $t \in T$. As t is a Tjur point, for every $k \in \mathbb{N}$, there is some $n_k \in \mathbb{N}$ such that $\hat{\xi}(n_k, k) = 0$ and $t \in B(n_k)$. Further, for such an n_k we have $d_{\mathcal{M}_1}(\mu_{\pi_2}^{B(n_k)}, \mu^t) \leqslant 2\cdot 2^{-k}$.

Define a sequence $\nu \in \mathscr{C}(\mathbb{N}, \mathcal{M}_1(S))$ by $\nu(k) = \mu_{\pi_2}^{B(n_k)}$. By construction of $\hat{\xi}$, the sequence ν is a rapidly converging Cauchy sequence converging to μ^t and so μ^t is computable from ν by the definition of a computable metric space.

In summary, there is a computable map taking the tuple $((\mu, \hat{\xi}), t)$ to μ^t. Let H be the corresponding 'curried' map taking the pair $(\mu, \hat{\xi})$ to the map $t \mapsto \mu^t$, which is also computable. Then H satisfies $\mathrm{D}_{S,T} = H \circ \langle \mathrm{id}, \mathrm{EC}\circ K\rangle$. □

4.3. *Equivalence*

The following result is an immediate consequence of Propositions 4.1, 4.3 and 4.6, as well as the fact that $\mathrm{Lim} \equiv_{\mathrm{sW}} \mathrm{EC}$ by Lemma 3.28.

Theorem 4.7. $\mathrm{D} \equiv_{\mathrm{sW}} \mathrm{D}_{\mathbb{N},[0,1]} \equiv_{\mathrm{sW}} \mathrm{Lim}$. □

5. The disintegration of specific distributions

In the previous section, we considered the disintegration operator on the space of measures for which the disintegration operator is continuous (and hence computable from a single real). In this section, we consider the possible Weihrauch degrees of disintegrations which may not be continuous. We will show that there exist individual disintegrations that are themselves strongly Weihrauch equivalent to Lim.

5.1. *Definitions*

In order to study disintegration in the context of probability measures that do not necessarily admit continuous disintegrations, we work with the notions developed by Fraser–Naderi based on the Vitali covering property, as described in Section 2.2.

Definition 5.1. Let T be a computable metric space and let μ be a probability measure on T. A class $V \subseteq \mathcal{B}_T$ of μ-continuity sets has the **strong Vitali covering property** with respect to μ when (1) it has the Vitali covering property with respect to μ and (2) there exists a multi-valued map $U : T \rightrightarrows \mathscr{C}(\mathbb{N}, \mathscr{O}_\mu(T))$, computable relative to μ, such that, for

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

every $x \in T$, there is a sequence $\langle E_i \rangle_{i \in \mathbb{N}}$ converging regularly to x with respect to μ such that

$$U(x)(n) \subseteq E_n \text{ and } \mu(U(x)(n)) = \mu(E_n), \tag{40}$$

for every $n \in \mathbb{N}$.

The following is an important example of a collection of sets with the computable Vitali covering property.

Lemma 5.2. If T is a computable ultrametric space then the collection of open balls has the computable Vitali covering property with respect to any measure on T.

Proof. By Proposition 2.17, we know that the collection of open balls has the Vitali covering property. However, if T is a computable metric space, then we can computably find a map that takes each element x in T to a sequence of open balls that converges regularly to x. □

Similarly, for $\mathbb{R}^n$, the collection of closed sets also has the computable Vitali covering property with respect to Lebesgue measure.

Definition 5.3. Let S and T be computable metric spaces, let $\mu \in \mathcal{M}_1(S \times T)$, let μ_T be its projection onto T, and let $V \subseteq \mathcal{B}_T$ be a class of μ_T-continuity sets having the computable Vitali covering property with respect to μ_T. Define $\mathrm{D}^\star_{S,T}(\mu, V)$ to be the map sending $t \in T$ to $\iota_1(\mu^t_{\pi_2,V})$, where $\mu^t_{\pi_2,V}$ is as defined in Proposition 2.14 and ι_1 is as defined after Proposition 2.10.

Remark 5.4. Observe that, by Proposition 2.14, $\mathrm{D}^\star_{S,T}(\mu, V)$ is equal to the composition of the map ι_1 with a disintegration of μ along π_2.

Remark 5.5. By Proposition 2.7, we see that, for every $\mu \in \mathcal{C}_{S,T}$ and class $V \subseteq \mathcal{B}_T$ of $\pi_2[\mu]$-continuity sets having the computable Vitali covering property, $\mathrm{D}^\star_{S,T}(\mu, V) = \mathrm{D}_{S,T}(\mu)$.

5.2. *Upper bound*

Proposition 5.6. Let S and T be computable metric spaces, let $\mu \in \mathcal{M}_1(S \times T)$ be computable, let μ_T be its projection onto T, let $V \subseteq \mathcal{B}_T$ be a class of μ_T-continuity sets with the strong Vitali covering property, and assume that $\mathrm{D}^\star_{S,T}(\mu, V)$ is total. Then there exists a computable map $K : T \to \mathscr{C}(\mathbb{N}, \mathcal{M}_1(S))$ such that

$$\mathrm{D}^\star_{S,T}(\mu, V) = \mathrm{Lim}_{\mathcal{M}_1(S)} \circ K. \tag{41}$$

In particular,

$$\mathrm{D}^\star_{S,T}(\mu, V) \leqslant_{\mathrm{sW}} \mathrm{Lim}_{\mathcal{M}_1(S)}. \tag{42}$$

Proof. Note that, if $\mu_T(H) = \mu_T(H') > 0$ and $H' \subseteq H$, then $\mu^H_{\pi_2} = \mu^{H'}_{\pi_2}$. Now, let $t \in T$. By the strong Vitali covering property, we may compute from t an element $U \in \mathscr{C}(\mathbb{N}, \mathscr{O}_\mu(T))$ such that $v(n) := \mu^{E_n}_{\pi_2} = \mu^{U_n}_{\pi_2}$ for all $n \in \mathbb{N}$, where $\langle E_i \rangle_{i \in \mathbb{N}}$ is some sequence of elements of V converging regularly to t with respect to μ_T. By Proposition 4.5, $v(n)$ is computable

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

from t, uniformly in n. By Definition 5.3, the limit of v exists and is $\mathrm{D}^{\star}_{S,T}(\mu, V)(t)$. Take K to be the map sending t to v. □

5.3. *Lower bound*

Let $\mathbb{I} := [0,1] \setminus \{\frac{a+1}{2^b} : a, b \in \mathbb{N}\}$. There is then a computable injection $i : \mathbb{I} \to 2^{\omega}$ which takes each element to its binary representation. Further $\mathbb{I}$ is a Lebesgue measure one subset of $[0,1]$ and Lebesgue measure on 2^{ω} is the pushforward of Lebesgue measure on $\mathbb{I}$ along i. Let $\mathcal{C} = 2^{\omega} \times 2^{\omega}$, which is an computable ultrametric space (i.e. a computable metric space where the metric is an ultrametric). Then $\mathbb{N} \times \mathcal{C}$ can be viewed as a computable ultrametric space as well. Let $V_{\mathcal{C}}$ be the collection of open balls in $\mathcal{C}$. Note that, by Proposition 2.17, the class $V_{\mathcal{C}}$ has the Vitali covering property with respect to any measure. Because every open ball in an ultrametric space is a continuity sets with respect to every measure, it is also immediate that $V_{\mathcal{C}}$ has the strong Vitali covering property with respect to every measure.

Write $a_n \sim b_n$ when $\frac{a_n}{b_n} \to 1$ as $n \to \infty$. The proof of the next proposition follows a suggestion from an anonymous referee and is significantly simpler than the original one presented.

Proposition 5.7. There is a computable distribution $\mu \in \mathcal{M}_1(\mathbb{N} \times \mathcal{C})$ such that its disintegration $\mathrm{D}^{\star}_{\mathbb{N},\mathcal{C}}(\mu, V_{\mathcal{C}})$ (along the projection $\pi : \mathbb{N} \times \mathcal{C} \to \mathcal{C}$) exists everywhere and satisfies $\mathrm{EC} \leqslant_{\mathrm{sW}} \mathrm{D}^{\star}_{\mathbb{N},\mathcal{C}}(\mu, V_{\mathcal{C}})$.

Proof. In Proposition 4.1, each $x \in \mathscr{C}(\mathbb{N}, \mathbb{S})$ is mapped to a measure μ_x over $\mathbb{N} \times [0,1]$ admitting a total continuous disintegration along the projection $\mathbb{N} \times [0,1] \to [0,1]$. Further, for each $x \in \mathscr{C}(\mathbb{N}, \mathbb{S})$ we have $\mu_x(\mathbb{I}) = 1$. This implies that the map $\mu_x \to \mu_x^{\mathbb{I}}$ is computable and also that for each $x \in \mathscr{C}(\mathbb{N}, \mathbb{S})$, the measure $\mu_x^{\mathbb{I}}$ admits a total continuous disintegration along the projection $\mathbb{N} \times \mathbb{I} \to \mathbb{I}$ with $\mu_x^y = (\mu_x^{\mathbb{I}})^y$ for all $y \in \mathbb{I}$.

Let $\eta_x = i[\mu_x^{\mathbb{I}}]$, be the pushforward of $\mu_x^{\mathbb{I}}$ along i. Note there is a map α that takes an open ball B in 2^{ω} to an open ball $\alpha(B) \subseteq i^{-1}(B) \subseteq \mathbb{I}$ such that for all $x \in 2^{\omega}$, $\nu_x(B) = \mu_x^{\mathbb{I}}(\alpha(B))$. As such, the map $\mu_x^{\mathbb{I}} \mapsto \eta_x$ is computable and so the map $x \mapsto \eta_x$ is also computable. Further it is immediate that for every $y \in \mathbb{I}$, $i(y)$ is a Tjur point (for the projection map onto 2^{ω}) and $\eta_x^{i(y)} = \mu_x^y$.

For each x, let $\nu_x \in \mathcal{M}_1(\mathbb{N})$ be $\eta_x^{i(0)}$ $(= \mu_x^0)$, i.e., the image of $i(0)$ under the continuous disintegration of η_x along the projection. From Equations (35) and (36), we see that, for every x and $k \in \mathbb{N}$,

$$x(k) = 2^{k+2}\, \nu_x\{2k\}. \tag{43}$$

Hence, x is a computable element of $\mathscr{C}(\mathbb{N}, \mathbb{B})$ from ν_x. Let H be a realizer for the map from $\mathcal{M}_1(\mathbb{N})$ to $\mathscr{C}(\mathbb{N}, \mathbb{B})$ taking ν_x to x.

Consider the map ρ from 2^{ω} to $\mathscr{C}(\mathbb{N}, \mathbb{S})$ such that $\rho(s) = x$ if and only if, for all $k \in \mathbb{N}$,

$$x(k) = 1 \iff \exists n \in \mathbb{N} \text{ s.t. the string } 0(1^k0)^n \text{ appears in } s \text{ at position } n. \tag{44}$$

Call an element $s \in 2^{\omega}$ *ρ-faithful* when, for all $k \in \mathbb{N}$, it holds that s contains the substring 01^k0 if and only if $\rho(s)(k) = 1$. It is straightforward to establish that ρ is a

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

total computable surjective map admitting a computable multi-valued inverse realized by a map $\bar{\rho}$ such that $\bar{\rho}(x)$ is ρ-faithful for all x. Now endow 2^ω with the uniform measure λ and define $\mu \in \mathcal{M}_1(\mathbb{N} \times 2^\omega \times 2^\omega)$ as the mixture given by

$$\mu(A \times B) = \int_B \eta_{\rho(s)}(A)\, \lambda(\mathrm{d}s), \tag{45}$$

for all measurable subsets $A \subseteq \mathbb{N} \times 2^\omega$ and $B \subseteq 2^\omega$, and let $\kappa = \mathrm{D}^\star_{\mathbb{N},\mathcal{C}}(\mu, V_{\mathcal{C}})$ be the disintegration of μ along the projection $\pi : \mathbb{N} \times \mathcal{C} \to \mathcal{C}$. Let G be a realizer for κ.

For a finite string s', write $[s']$ for the clopen subset of 2^ω comprised of all strings with prefix s'. For $n \geqslant 0$ and $s \in 2^\omega$, let s_n be the length-n prefix of s, let $B^s_n = [0^n] \times [s_n]$, and note that $B^s_n \in V_{\mathcal{C}}$ and that $B^s_1, B^s_2, \ldots$ converges regularly to $(0^\omega, s)$ with respect to Lebesgue measure. It follows that, for every $s \in 2^\omega$ and subset $A \subseteq \mathbb{N}$,

$$\kappa((0^\omega, s))(A) = \lim_{n\to\infty} \frac{\mu(A \times B^s_n)}{\mu(\mathbb{N} \times B^s_n)}, \tag{46}$$

provided the limit exists.

Claim 5.8. For every ρ-faithful $s \in 2^\omega$, we have $\lim_{n\to\infty} \frac{\mu(A\times B^s_n)}{\mu(\mathbb{N}\times B^s_n)} = \nu_{\rho(s)}(A)$.

Proof of Claim. Let $s \in 2^\omega$ be ρ-faithful, let $x \in \mathscr{C}(\mathbb{N}, \mathbb{S})$ be the image of s under ρ, and let $k \in \mathbb{N}$. For every $n \in \mathbb{N}$, let $E_n = \{s' \in [s_n] \,:\, \rho(s')(k) = x(k)\}$ and note that

$$\mu(\{k\} \times [0^n] \times E_n) = \eta_x(\{k\} \times [0^n]). \tag{47}$$

If $x(k) = 1$ then there exists $n_0 \in \mathbb{N}$, such that $E_n = [s_n]$ for every $n \geqslant n_0$. If $x(k) = 0$, then, for $n \geqslant k + 1$,

$$\frac{\lambda(E_n)}{\lambda([s_n])} \geqslant 1 - \sum_{i=0}^{\infty} 2^{-(k+1)(n+i)-1} \geqslant 1 - \sum_{i=0}^{\infty} 2^{-n-i} = 1 - 2^{-n+1}, \tag{48}$$

where the first inequality follows from the ρ-faithfulness of s and a union bound on the event that $0(1^k0)^{n+i}$ appears at position $n + i$ for some $i \in \mathbb{N}$. Therefore, in general,

$$\lambda(E_n) \sim \lambda([s_n]), \text{ for } n \to \infty, \tag{49}$$

and so

$$\mu(\{k\} \times [0^n] \times [s_n]) \sim \mu(\{k\} \times [0^n] \times E_n), \text{ for } n \to \infty. \tag{50}$$

But then

$$\lim_{n\to\infty} \frac{\mu(\{k\} \times B^s_n)}{\mu(\mathbb{N} \times B^s_n)} = \lim_{n\to\infty} \frac{\eta_x(\{k\} \times [0^n])}{\eta_x(\mathbb{N} \times [0^n])} = \nu_x(\{k\}), \tag{51}$$

where the final equality follows from the fact, established in Proposition 4.1, that 0^ω is a Tjur point of η_x (along the projection onto 2^ω), completing the proof of the claim. □

Now let $x \in \mathscr{C}(\mathbb{N}, \mathbb{S})$. From x, we can, by assumption, compute a ρ-faithful element $s \in 2^\omega$ such that $\rho(s) = x$. Let K be a realizer for the map taking x to $(0^\omega, s) \in \mathcal{C}$. By the above claim, the image of $(0^\omega, s)$ under κ is ν_x, and so the composition $G \circ K$ is a realizer

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

for the map taking x to v_x. Finally, the map realized by H above takes v_x to x as an element of $\mathscr{C}(\mathbb{N}, \mathbb{B})$, completing the proof. □

Remark 5.9. The particular choice of ρ is essential here: different surjections from 2^ω to $\mathscr{C}(\mathbb{N}, \mathbb{S})$ do not necessarily lead to a reduction. For example, consider the map ρ' taking a binary sequence to the set of all natural numbers k such that 01^k0 appears in the sequence. Like ρ, the map ρ' is a total computable surjective map admitting a computable multi-valued inverse whose images are ρ'-faithful. However, Equation (49) fails to hold for ρ'. The source of the failure is the fact that, for λ-almost all $s \in 2^\omega$, we have $\rho'(s) = c_1$ where $c_1(k) = 1$ for all $k \in \mathbb{N}$! Had we used ρ', then the disintegration of μ obtained by taking limits would have been the constant map $(u, s) \mapsto v_{x'}$. It is surprising, though easy to verify, that the disintegration $(u, s) \mapsto v_{\rho'(s)}$ is a version of this disintegration, but it does not arise from the limiting construction of the disintegration when μ is constructed from ρ'.

Theorem 5.10. There are computable ultrametric spaces S and T and a computable distribution $\mu \in \mathcal{M}_1(S \times T)$ such that if V is the collection of open balls in T then the disintegration $\mathrm{D}^\star_{S,T}(\mu, V)$ (along the projection $\pi : S \times T \to T$) exists everywhere and satisfies $\mathrm{Lim} \equiv_{\mathrm{sW}} \mathrm{D}^\star_{S,T}(\mu, V)$.

Proof. Let $S = \mathbb{N}$, $T = \mathcal{C}$, and $V = V_{\mathcal{C}}$ with $\mathcal{C}, V_{\mathcal{C}}$, and μ as in Proposition 5.7. Then $\mathrm{Lim} \leqslant_{\mathrm{sW}} \mathrm{D}^\star_{S,T}(\mu, V)$ follows immediately from Proposition 5.7 and Proposition 3.28. Finally, $\mathrm{D}^\star_{S,T}(\mu, V) \leqslant_{\mathrm{sW}} \mathrm{Lim}$ follows immediately from Proposition 5.6 and Proposition 3.26. □

Acknowledgements

The authors would like to thank Arno Pauly for helpful conversations regarding Weihrauch reducibility (and in particular the argument for Proposition 3.24), and the anonymous referees for detailed comments on a draft, including suggestions that greatly simplified the proofs of Propositions 4.1, 4.3 and 5.7. The authors would also like to thank Persi Diaconis for pointing us to Tjur's work, Laurent Bienvenu for useful discussions, and Quinn Culver, Bjørn Kjos-Hanssen, and Geoff Patterson for comments on an earlier version of this material.

Work on this publication was made possible through the support of ARO grant W911NF-13-1-0212, ONR grant N00014-13-1-0333, NSF grants DMS-0901020 and DMS-0800198, DARPA Contract Award Numbers FA8750-14-C-0001 and FA8750-14-2-0004, and grants from the John Templeton Foundation and Google. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation or the U.S. Government.

This paper was partially written while CEF and DMR were participants in the program *Semantics and Syntax: A Legacy of Alan Turing* at the Isaac Newton Institute for the Mathematical Sciences. DMR was partially supported by graduate fellowships from the National Science Foundation and MIT Lincoln Laboratory, by a Newton International

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

Fellowship, Emmanuel Research Fellowship, and by an NSERC Discovery Grant and Connaught Award.

References

Ackerman, N.L., Freer, C.E. and Roy, D.M. (2010). On the computability of conditional probability. Preprint, arXiv:1005.3014.

Ackerman, N.L., Freer, C.E. and Roy, D.M. (2011). Noncomputable conditional distributions. In: *Proceedings of the 26th Annual IEEE Symposium on Logic in Computer Science (LICS 2011)*, IEEE Computer Society 107–116.

Bosserhoff, V. (2008). Notions of probabilistic computability on represented spaces. *Journal of Universal Computer Science* **14** (6) 956–995.

Brattka, V. (2005). Effective Borel measurability and reducibility of functions. *Mathematical Logic Quarterly* **51** (1) 19–44.

Brattka, V. and Gherardi, G. (2011a). Effective choice and boundedness principles in computable analysis. *Bulletin of Symbolic Logic* **17** (1) 73–117.

Brattka, V. and Gherardi, G. (2011b). Weihrauch degrees, omniscience principles and weak computability. *Journal of Symbolic Logic* **76** (1) 143–176.

Brattka, V., Gherardi, G. and Marcone, A. (2012). The Bolzano–Weierstrass Theorem is the jump of Weak Kőnig's Lemma. *Annals of Pure Applied Logic* **163** (6) 623–655.

Brattka, V., Hölzl, R. and Gherardi, G. (2015). Probabilistic computability and choice. *Information and Computation* **242** 249–286.

Collins, P. (2009). A computable type theory for control systems. In: *Proceedings of the 48th IEEE Conference on Decision and Control* 5583–5543.

Collins, P. (2010). Computable analysis with applications to dynamical systems. Technical Report MAC-1002, Centrum Wiskunde & Informatica.

Escardó, M. (2004). Synthetic topology. *Electronic Notes in Theoretical Computer Science* **87** 21–156.

Fraser, D.A.S., McDunnough, P., Naderi, A. and Plante, A. (1995). On the definition of probability densities and sufficiency of the likelihood map. *Probability and Mathematical Statistics* **15** 301–310.

Fraser, D.A.S. and Naderi, A. (1996). On the definition of conditional probability. In *Research Developments in Probability and Statistics*, VSP, Utrecht 23–26.

Gács, P. (2005). Uniform test of algorithmic randomness over a general space. *Theoretical Computer Science* **341** (1-3) 91–137.

Hoyrup, M. and Rojas, C. (2009). Computability of probability measures and Martin-Löf randomness over metric spaces. *Information and Computation* **207** (7) 830–847.

Hoyrup, M. and Rojas, C. (2011). Absolute continuity of measures and preservation of randomness. Preprint, `http://www.loria.fr/~hoyrup/abscont.pdf`.

Hoyrup, M., Rojas, C. and Weihrauch, K. (2012). Computability of the Radon–Nikodym derivative. *Computability* **1** (1) 3–13.

Kallenberg, O. (2002). *Foundations of Modern Probability*, 2nd ed. Springer.

Kolmogorov, A.N. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Springer.

Pauly, A. (2016). On the topological aspects of the theory of represented spaces. *Computability*.

Pauly, A. and Fouché, W. (2014). How constructive is constructing measures? arXiv:1409.3428v1.

Pfanzagl, J. (1979). Conditional distributions as derivatives. *Annals of Probability* **7** (6) 1046–1050.

Schröder, M. (2002a). *Admissible Representations for Continuous Computations*. Ph.D. Thesis, Fachbereich Informatik, FernUniversität Hagen.

Schröder, M. (2002b). Effectivity in spaces with admissible multirepresentations. *Mathematical Logic Quarterly* **48** (suppl. 1)78–90.

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press

Schröder, M. (2007). Admissible representations for probability measures. *Mathematical Logic Quarterly* **53** (4-5) 431–445.

Tjur, T. (1975). *A Constructive Definition of Conditional Distributions.* Preprint 13. Institute of Mathematical Statistics, University of Copenhagen.

Tjur, T. (1980). *Probability Based on Radon Measures*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons.

Weihrauch, K. (2000). *Computable Analysis: An Introduction*, Springer.

https://doi.org/10.1017/S0960129516000098 Published online by Cambridge University Press