Computability of the ergodic decomposition

Mathieu Hoyrup

LORIA, INRIA Nancy-Grand Est, 615 rue du jardin botanique, 54600 Villers-lès-Nancy, France

Abstract

The study of ergodic theorems from the viewpoint of computable analysis is a rich field of investigation. Interactions between algorithmic randomness, computability theory and ergodic theory have recently been examined by several authors. It has been observed that ergodic measures have better computability properties than non-ergodic ones. In a previous paper we studied the extent to which non-ergodic measures inherit the computability properties of ergodic ones, and introduced the notion of an *effectively decomposable* measure. We asked the following question: if the ergodic decomposition of a stationary measure is finite, is this decomposition effective? In this paper we answer the question in the negative.

Keywords: computable analysis, Martin-Löf randomness, ergodic decomposition, Birkhoff's ergodic theorem

1. Introduction

The ergodic decomposition theorem says the following: every stationary process can be decomposed into ergodic processes, such that almost every realization of the original process can be seen as a realization of one of the ergodic processes, chosen at random. Ergodic processes are in a sense the building blocks of all the stationary processes. The question of the effectiveness of many ergodic theorems has received much attention in recent years and it progressively appeared that ergodic measures behave differently from non-ergodic ones. For instance, the speed of convergence of Birkhoff averages is computable in the ergodic case [1] while it is not computable in general [16]; Birkhoff ergodic theorem holds exactly at Schnorr random sequences in the ergodic case [8] and at Martin-Löf random sequences in general [16, 4]. These examples suggest that ergodic measures have better computability properties than non-ergodic ones. In [10] we showed that the sticking point is not really ergodicity but the computability of the ergodic decomposition. While every non-ergodic measure has a unique decomposition into ergodic ones, this decomposition is not always computable. The known examples of non-ergodic measures whose decomposition is non-computable are infinite combinations of ergodic measures ([16, 1]). In [10] we raised the following question: if the decomposition of a non-ergodic measure is finite, is this decomposition computable? In the present paper we solve the problem and show that it is not necessarily true. Before presenting this new result, we review the results obtained in [10] and characterize the effective compact classes of ergodic measures.

The paper is organized as follows. In Section 2 we give the necessary background on computability and randomness. In Section 3 we develop results about randomness and combinations of measures that will be applied in the sequel, but are of independent interest (i.e., outside ergodic theory). We start Section 4 with a reminder on the ergodic decomposition and then relate it to randomness. In Section 5 we study the particular case of effective compact classes of ergodic measures. We finish in Section 6 by our main result: there exist ergodic measures P and Q whose average is not effectively decomposable.

2. Preliminaries

We assume familiarity with algorithmic randomness and computability theory. For more details on computable analysis we refer the reader to [17].

2.1. Computability

A computable metric space is a triple (X, d, S) where (X, d) is a complete separable metric space and S is a countable dense subset together with a fixed numbering such that for all $s, s' \in S$, $d(s, s')$ is a computable real number, uniformly in the indices of s and s'. The basic metric balls $B(s, q)$ with $s \in S$ and $q \in \mathbb{Q}_{>0}$ form a countable basis of the topology induced by the metric d. We fix a canonical effective numbering $(B_i)_{i\in\mathbb{N}}$ of this basis.

Let X be a computable metric space. A ***name*** for $x \in X$ is a sequence $s_n \in S$ such that $d(s_n, x) < 2^{-n}$. A point x is ***computable*** if it has a computable name. A set $U \subseteq X$ is an ***effective open set*** if there is r.e. set $E \subseteq \mathbb{N}$ such that $U = \bigcup_{i\in E} B_i$. A function $f : X \to Y$ is computable if there is a machine that, provided a name for x as oracle, computes a name for $f(x)$. Equivalently, f is computable if the pre-images $f^{-1}(B_i)$ are effective open sets, uniformly in i. Let $A \subseteq X$. A function $f : A \to Y$ is ***computable on*** A if there is a machine that, provided a name for $x \in A$ as oracle, computes a name for $f(x)$. Equivalently, f is computable on A if the pre-images $f^{-1}(B_i) \cap A$ are intersections of uniformly effective open sets with A. A point $y \in Y$ is ***computable relative to*** $x \in X$ if the function $x \mapsto y$ is computable on $\{x\}$. A function $f : X \to [0, +\infty]$ is ***lower semi-computable*** if there is a machine that, provided a name for x as oracle, computes a nondecreasing sequence of rational numbers converging to $f(x)$. Equivalently f is lower semi-computable if the pre-images $f^{-1}(q, +\infty]$ are effective open sets, uniformly in $q \in \mathbb{Q}$. A compact set $K \subseteq X$ is ***effectively compact*** if the set of finite unions of balls covering K is r.e.

We will use the following simple results that are the effective counterparts of basic topological properties.

Fact 1 (Folklore).

1. The complement of an effective compact set is an effective open set.

2. If K is effectively compact and U effectively open then $K \setminus U$ is effectively compact.

Proof. 1. Let $K \subseteq X$ be effectively compact. Let B be a basic metric ball and $\overline{B}$ be the corresponding closed ball. As the complement of $\overline{B}$ is effectively open so $K \cap \overline{B} = \emptyset$ can be semi-decided. Hence $X \setminus K$ is the r.e. union of all basic balls B such that $K \cap \overline{B} = \emptyset$.

2. $K \setminus U$ is compact and $K \setminus U \subseteq (B_1 \cup \ldots \cup B_n) \iff K \subseteq U \cup (B_1 \cup \ldots \cup B_n)$ which can be semi-decided.

□

Let $K \subseteq X$ be an effective compact set and $f : K \to Y$ a function computable on K.

Fact 2 (Folklore). $f(K)$ is an effective compact set.

Proof. Let $B_1, \ldots, B_n$ be basic balls of Y. $f(K)$ is contained in $B_1 \cup \ldots \cup B_n$ if and only if K is contained in $f^{-1}(B_1 \cup \ldots \cup B_n)$, which is an effective open set. As K is effectively compact the latter inclusion can be semi-decided. □

Fact 3 (Folklore). If f is moreover one-to-one then $f^{-1} : f(K) \to K$ is computable on $f(K)$.

Proof. For the sake of clarity, we denote f^{-1} by g.

Let $B \subseteq X$ be a basic ball. We have to prove that there is an effective open set $V \subseteq Y$ such that $g^{-1}(B) = V \cap f(K)$. The set $C := K \setminus B$ is an effective compact set. $g^{-1}(B) = g^{-1}(K \setminus C) = g^{-1}(K) \setminus g^{-1}(C) = f(K) \setminus f(C)$. As C is an effective compact set, its complement V is an effective open set and $g^{-1}(B) = f(K) \cap V$. As everything is uniform in B, g is computable. □

The product of two computable metric spaces has a natural structure of computable metric space.

Fact 4 (Folklore). If $K \subseteq X$ is an effective compact set and $f : K \times Y \to \overline{\mathbb{R}}$ is lower semi-computable, then the function $g : Y \to \overline{\mathbb{R}}$ defined by $g(y) = \inf_{x \in K} f(x, y)$ is lower semi-computable.

Proof. Let us prove that $g^{-1}(q, +\infty] = \{y : K \times \{y\} \subseteq f^{-1}(q, +\infty]\}$ is an effective open set, uniformly in q. Let q be some fixed rational number. The effective open set $U_q = f^{-1}(q, +\infty]$ can be expressed as an effective union of product balls $U_q = \bigcup_{i \in \mathbb{N}} B_i^X \times B_i^Y$. The set $E_q = \{(i_1, \ldots, i_k) : K \subseteq B_{i_1}^X \cup \ldots \cup B_{i_k}^X\}$ is r.e. and it is easy to prove that $g^{-1}(q, +\infty] = \bigcup_{(i_1, \ldots, i_k) \in E_q} B_{i_1}^Y \cap \ldots \cap B_{i_k}^Y$, which is an effective open set. The argument is uniform in q. □

If X is a computable metric space then the set of Borel probability measures over X can be endowed with a structure of computable metric space (see [6], e.g.) inducing the weak*-topology: measures P_n converge to P if for every bounded continuous $f : X \to \mathbb{R}$, $\int f \,\mathrm{d}P_n$ converge to $\int f \,\mathrm{d}P$. If X is effectively compact then so is $\mathcal{P}(X)$. With this computability structure, a probability measure P is computable if for every lower semi-computable function $f : X \to [0, +\infty]$, $\int f \,\mathrm{d}P$ is lower semi-computable, uniformly in f. Equivalently, P is computable if for every bounded computable function f, $\int f \,\mathrm{d}P$ is computable, uniformly in f.

Notation. The Cantor space of infinite binary sequences will be denoted by $2^{\mathbb{N}}$. It can be made a computable metric space with the metric $d(x, y) = 2^{-\min\{n:x_n \neq y_n\}}$, where $x = x_0x_1x_2\dots$ and $y = y_0y_1y_2\dots$. This space is effectively compact.

If f, g are real-valued functions, $f \overset{*}{\leq} g$ means that there exists $c \geq 0$ such that $f \leq cg$. $f \overset{*}{=} g$ means that $f \overset{*}{\leq} g$ and $g \overset{*}{\leq} f$.

2.2. Effective randomness

Martin-Löf [13] was the first one to define a sound individual notion of random infinite binary sequence. He developed his theory for any computable probability measure on the Cantor space. This theory was then extended to non-computable measures by Levin [12], and later by [6, 11] on general spaces ([9] was an extension to topological spaces, but for computable measures).

We will use the most general theory: we will be consider infinite binary sequences and probability measures that are random w.r.t. non-computable measures, over $2^{\mathbb{N}}$ and $\mathcal{P}(2^{\mathbb{N}})$ respectively. We will use the notion of uniform test of randomness, introduced by Levin [12] and further developed in [6, 7, 11].

On a computable metric space X endowed with a probability measure P, there is a set ML_P of P-random elements satisfying $P(\mathsf{ML}_P) = 1$, together with a canonical decomposition (coming from the universal P-test) $\mathsf{ML}_P = \bigcup_n \mathsf{ML}_P^n$ where ML_P^n are uniformly effective compact sets relative to P, $\mathsf{ML}_P^n \subseteq \mathsf{ML}_P^{n+1}$ and $P(\mathsf{ML}_P^n) > 1-2^{-n}$. The sets $X \setminus \mathsf{ML}_P^n$ constitute a universal Martin-Löf test. A test is a function $t : \mathcal{P}(X) \times X \to [0, +\infty]$ which is lower semi-computable and such that $\int t_P \,\mathrm{d}P \leq 1$ for all $P \in \mathcal{P}(X)$, where $t_P(x)$ is a notation for $t(P, x)$.

A function $f : X \to Y$ is ***P-layerwise computable*** if there is an oracle machine that, given n as input and a name of $x \in \mathsf{ML}_P^n$ as an oracle, outputs a name of $f(x)$. Nothing is required to the machine when x is not P-random. In other words, f is P-layerwise computable if it is computable on each ML_P^n, uniformly in n. When f is P-layerwise computable, for every P-random x, $f(x)$ is computable relative to x in a way that is not fully uniform, but uniform on each set ML_P^n.

Lemma 2.1. *Let P be a computable measure, $f : X \to Y$ a P-layerwise computable function and $Q = f_*P$ the push-forward of P under f.*

1. *Q is computable and $f : \mathsf{ML}_P \to \mathsf{ML}_Q$ is onto.*

2. *If $f : X \to Y$ is moreover one-to-one then $f : \mathsf{ML}_P \to \mathsf{ML}_Q$ is one-to-one and f^{-1} is Q-layerwise computable.*

3. Randomness and continuous combination of measures

The material developed here will be used to investigate the algorithmic content of the ergodic decomposition.

Given a *countable* class of probability measures P_i over $2^{\mathbb{N}}$ and real numbers $\alpha_i \in [0,1]$ such that $\sum_i \alpha_i = 1$, the convex combination $P = \sum_i \alpha_i P_i$ is again a probability measure. This can be generalized to *continuous* classes of measures, as we briefly recall now.

Let m be a probability measure over $\mathcal{P}(2^{\mathbb{N}})$. The set function P defined by $P(A) = \int Q(A)\,\mathrm{d}m(Q)$ for measurable sets $A \subseteq 2^{\mathbb{N}}$ is a probability measure over $2^{\mathbb{N}}$, called the barycenter of m. It satisfies

$$\int f(x)\,\mathrm{d}P(x) = \int \Big(\int f(x)\,\mathrm{d}Q(x) \Big)\,\mathrm{d}m(Q) \tag{1}$$

for $f \in L^1(2^{\mathbb{N}}, P)$. We can think of P as the measure describing the following process: first pick some measure Q at random according to m; then run the process with distribution Q.

Probabilistically, picking a sequence according to P or decomposing into these two steps are equivalent. We are interested in whether the algorithmic theory of randomness fits well with this intuition: are the P-random sequences the same as the sequences that are Q-random for some m-random Q?

The answer is positive when m is computable. Observe that in this case P is also computable: (1) gives a formula to compute P knowing m. Actually, the function which maps m to its barycenter P is itself computable.

Theorem 3.1. *Let $m \in \mathcal{P}(\mathcal{P}(2^{\mathbb{N}}))$ be computable, and P be the barycenter of m. For $x \in X$, the following are equivalent:*

1. *x is P-random,*

2. *x is Q-random for some m-random Q.*

In other words,

$$\mathsf{ML}_P = \bigcup_{Q \in \mathsf{ML}_m} \mathsf{ML}_Q.$$

4. Randomness and ergodic decomposition

4.1. Background from ergodic theory

We consider the ***shift transformation*** $T : 2^{\mathbb{N}} \to 2^{\mathbb{N}}$ defined by $T(x_0x_1x_2\ldots) = x_1x_2x_3\ldots$. A measure P over $2^{\mathbb{N}}$ is ***stationary***, or shift-invariant, if $P(T^{-1}(A)) = P(A)$ for all Borel sets $A \subseteq 2^{\mathbb{N}}$. A stationary measure P is ***ergodic*** if for every Borel set A satisfying $T(A) \subseteq A$, $P(A) = 0$ or 1.

A sequence $x \in \{0,1\}^{\mathbb{N}}$ is ***generic*** if for each $w \in \{0,1\}^*$, the frequency of occurrences of w in x converges. If x is generic, we denote by Q_x the set function which maps each cylinder $[w]$ to the limit frequency of occurrences of w in x. Q_x extends to a probability measure over the Cantor space, which we also denote by Q_x. If x is generic then Q_x is stationary. Birkhoff's ergodic theorem states that given a stationary measure P, P-almost every x is generic. If P is moreover ergodic then $Q_x = P$ for P-almost every x.

There is a geometrical way of describing the ergodic measures. The set of stationary measures is a compact convex subset of $\mathcal{P}(2^{\mathbb{N}})$ whose extremal points are exactly the ergodic measures. A theorem of Choquet from convex analysis (see [15], e.g.) can then be applied to get the ergodic decomposition theorem: every stationary measure can be uniquely decomposed into a convex combination of ergodic measures. Formally, for every stationary measure P there exists a unique probability measure m_P over $\mathcal{P}(2^{\mathbb{N}})$ such that (i) m_P gives full weight to the set of ergodic measures and (ii) m_P is the barycenter of m_P as defined in (1). We will call m_P the ***Choquet measure*** associated to P.

The measure m_P can be obtained in the following way: let ϕ be the P-almost-everywhere defined function which maps x to Q_x. m_P is the push-forward measure $\phi_* P$, i.e. $m_P(\mathcal{A}) = P(\{x : Q_x \in \mathcal{A}\})$ for all Borel sets $\mathcal{A} \subseteq \mathcal{P}(2^{\mathbb{N}})$. As m_P is concentrated on the ergodic measures, Q_x is ergodic for P-almost every x.

We will need the following effective topological properties of the set of stationary measures. The class of stationary measures is an effective compact subset of $\mathcal{P}(2^{\mathbb{N}})$. The class of ergodic stationary measures is an effective G_δ-set, i.e. an intersection of uniformly effective open sets, which is dense in the set of stationary measures: every stationary measure can be approached by ergodic measures with finite memory, also called Markov measures (see [14] for instance).

4.2. Randomness and ergodic theorems

An algorithmic version of Birkhoff's ergodic theorem was eventually proved by V'yugin [16]: given a stationary measure P, every P-random sequence is generic, and if P is moreover ergodic then $Q_x = P$ for every P-random sequence x (it was proved for computable measures, but it still works for non-computable measures). The proof was not immediate to obtain from the classical proof of Birkhoff's theorem, which is in a sense not constructive. In this paper we are interested in an algorithmic version of the ergodic decomposition theorem, which again cannot be proved directly.

More precisely, given a stationary measure P, we are interested in the following questions:

- if x is P-random, is Q_x ergodic?
- if x is P-random, is x also Q_x-random?
- if x is P-random, is Q_x an m_P-random measure?

- does any converse implication hold?

We give positive partial answers to these questions, leaving the general problem open. We will use the following lemmas (the first one was proved in [16]).

Lemma 4.1. *Let P be an ergodic stationary probability measure. For every $x \in \mathsf{ML}_P$, $Q_x = P$.*

Lemma 4.2. *Let P be a stationary probability measure and m_P the associated Choquet measure. Every m_P-random measure is ergodic and stationary.*

4.3. Effective decomposition

A stationary probability measure P is always computable relative to its associated Choquet measure m_P. The converse does not always hold (V'yugin [16] constructed a counter-example).

Definition 4.1. A stationary probability measure P is ***effectively decomposable*** if its Choquet measure is computable relative to P.

When P is computable. As an application of Theorem 3.1, we directly get a result when P is computable and effectively decomposable (i.e. when $m := m_P$ is computable).

Corollary 4.1. *Let P be a computable stationary probability measure that is effectively decomposable. For $x \in X$, the following are equivalent:*

1. *x is P-random,*
2. *x is Q-random for some m-random Q.*

In other words, the following are equivalent:

1. *x is P-random,*
2. *x is generic, Q_x-random and Q_x is m-random.*

We also have the following characterization. For $f \in L^1(X, P)$, we denote by f^* the limit of the Birkhoff averages of f.

Theorem 4.1. *Let P be a computable stationary probability measure. The following are equivalent.*

1. *P is effectively decomposable,*
2. *the function $X \to \mathcal{P}(X), x \mapsto Q_x$ is P-layerwise computable,*
3. *the function $L^1(X, P) \to L^1(X, P), f \mapsto f^*$ is computable.*

When P is not computable. If P is not computable but still effectively decomposable, one implication in Corollary 4.1 remains, with the same proof.

Theorem 4.2. *Let P be a stationary probability measure that is effectively decomposable. For every P-random x, Q_x is m_P-random, hence ergodic, and x is Q_x-random.*

The converse implication does not hold in general, as illustrated by the following counter-example. Let x be a sequence that is random w.r.t. the uniform measure λ. Let p_x be the real number whose binary expansion is $0.x$ and B_x be the Bernoulli measure with parameter p_x. Let $P = \frac{1}{2}(\lambda + B_x)$. P is not computable as x, which is not computable, is computable relative to P: $x = 2P[1] - 1/2$. P is effectively decomposable: indeed, $m_P = \frac{1}{2}(\delta_\lambda + \delta_{B_x})$ is computable relative to x which is computable relative to P. Now, x is λ-random and λ is m_P-random, but x is not P-random as it is computable relative to P and $P(\{x\}) = 0$.

The effectivity of the ergodic decomposition enables one to extend results from ergodic systems to non-ergodic ones. Let us illustrate it. It was proved in [2] that when P is an ergodic measure, every P-random sequence eventually visits every effective compact set of positive measure under shift iterations. When the decomposition is effective, this theorem can be generalized to non-ergodic measures, giving a version of Poincaré recurrence theorem for random sequences.

Corollary 4.2. *Let P be a stationary measure that is effectively decomposable. Let F be an effective compact set such that $P(F) > 0$. Every P-random $x \in F$ falls infinitely often in F under shift iterations.*

The result actually holds as soon as for every P-random x, Q_x is ergodic and x is Q_x-random.

5. Effective compact classes of ergodic measures

When restricting to some classes of ergodic measures, as the Bernoulli measures, the ergodic decomposition is computable.

Proposition 5.1. *Let P be a stationary probability measure. If m_P is supported on an effective compact class of ergodic measures, then P is effectively decomposable.*

The above proposition implies the computability of De Finetti measures on the Cantor space (see [5]).

Example 1. Let m be a computable probability measure over the real interval $[0, 1]$. Pick a real number p at random according to m, and then generate an infinite binary sequence tossing a coin with probability of heads p. As an application of the preceding proposition, we get that the function which maps a random sequence generated by the process to the number p that was picked is

P-layerwise computable: it can be computed from the observed outcomes with high probability.

We also learn that the algorithmic theory of randomness fits well with this example: obviously, we expect a sequence that is random w.r.t. the measure underlying the whole process to be random for some Bernoulli measure B_p, which is not immediate.

In Section 2.2, we defined P-layerwise computable functions when P is a computable probability measure. This can be extended to any effective compact class of measures $\mathscr{C}$. The class $\mathscr{C}$ admits a universal test, which gives a canonical decomposition of the set of points that are random w.r.t. to some measure in $\mathscr{C}$: $\mathsf{ML}_{\mathscr{C}} = \bigcup_n \mathsf{ML}_{\mathscr{C}}^n$ where $\mathsf{ML}_{\mathscr{C}}^n = \bigcup_{P\in\mathscr{C}} \mathsf{ML}_P^n$. The effective compactness of $\mathscr{C}$ implies the effective compactness of all the sets $\mathsf{ML}_{\mathscr{C}}^n$. A function $f : X \to Y$ is $\mathscr{C}$-layerwise computable if it is computable on each $\mathsf{ML}_{\mathscr{C}}^n$, uniformly in n. It means that one can compute $f(x)$ if x is random for some measure $P \in \mathscr{C}$, with probability of error bounded by 2^{-n}, whatever P is (as long as it is in $\mathscr{C}$), and for any n.

From Proposition 5.1 and Corollary 4.1 we know that every point that is random w.r.t. some measure in $\mathcal{I}_{\mathscr{C}}$ is already random w.r.t. to some ergodic measure in $\mathscr{C}$, namely Q_x. In other words, $\mathsf{ML}_{\mathscr{C}} = \mathsf{ML}_{\mathcal{I}_{\mathscr{C}}}$ while $\mathscr{C} \subsetneq \mathcal{I}_{\mathscr{C}}$ in general. We now prove a quantitative version of this fact. We recall that if $\mathcal{A}$ is an effective compact class of measures, $t_{\mathcal{A}} := \inf_{P\in\mathcal{A}} t_P$ is a universal $\mathcal{A}$-test, i.e. (i) it is lower semi-computable, (ii) $\int t_{\mathcal{A}}\,\mathrm{d}P \leq 1$ for every $P \in \mathcal{A}$ and (iii) $t_{\mathcal{A}}$ multiplicatively dominates every function satisfying (i) and (ii) (see [7] for more details about such class tests). We will consider the class tests $t_{\mathscr{C}}$ and $t_{\mathcal{I}_{\mathscr{C}}}$.

Theorem 5.1. *Let $\mathscr{C}$ be an effective compact class of stationary ergodic probability measures. One has:*

1. $t_{\mathscr{C}}(x) \stackrel{*}{=} t_{\mathcal{I}_{\mathscr{C}}}(x)$
2. *The function $x \mapsto Q_x$ is $\mathcal{I}_{\mathscr{C}}$-layerwise computable and $\mathscr{C}$-layerwise computable.*

Observe that for generic sequences x, $t_{\mathscr{C}}(x) = t_{Q_x}(x)$. Indeed, $t_{\mathscr{C}}(x) = \inf_{P\in\mathscr{C}} t_P(x) = t_{Q_x}(x)$ as $t_P(x) = +\infty$ for every $P \in \mathscr{C} \setminus \{Q_x\}$.

Theorem 5.1 tells us in particular that if the ergodic measure P belongs to an effective compact class of ergodic measures then it is computable relative to its random points. In the language of [3], P is *learnable*. Given an ergodic measure P it is not clear how to check whether it can be embedded in such an effective compact class of ergodic measures: Theorem 5.2 below identifies the property of the measure that makes it possible.

It was proved in [1] that when P is an ergodic measure, the convergence of the Birkhoff averages is computable *relative to* P. There exist ergodic measures P for which the convergence is plainly computable, i.e. for which the oracle P can be avoided. We prove that these measures are precisely the members of effective compact classes of ergodic measures. To state and prove this characterization, we consider an effective enumeration f_i of a class of bounded computable functions from X to $\mathbb{R}$, that determine the probability measures: if $\int f_i \, \mathrm{d}P = \int f_i \, \mathrm{d}P'$ for all i then $P = P'$.

Theorem 5.2. *Let P be an ergodic probability measure. The following are equivalent:*

1. *P belongs to an effective compact class of ergodic measures,*
2. *the convergence of $A_n^{f_i}$ to $\int f_i \, \mathrm{d}P$ is effective, uniformly in i.*

Proof. 1. ⇒ 2. As mentioned above it was proved in [1] that the convergence of Birkhoff averages is effective relative to P when P is ergodic. More precisely, there exists an upper semi-computable function $n(P, i, \delta, \epsilon)$ defined on ergodic measures P, such that for all ergodic measures P, all positive rationals δ, ϵ and all $i \in \mathbb{N}$,

$$P\Big\{x : \sup_{n \geq n(P,i,\delta,\epsilon)} |A_n^{f_i}(x) - \int f_i \, \mathrm{d}P| > \delta\Big\} \leq \epsilon.$$

Let $\mathscr{C}$ be an effective compact class of ergodic measures and let $m(i, \delta, \epsilon) = \max_{P \in \mathscr{C}} n(P, i, \delta, \epsilon)$ (which is finite as $\mathscr{C}$ is compact and n upper semi-continuous). The function m is upper semi-computable and for all $P \in \mathscr{C}$,

$$P\Big\{x : \sup_{n \geq m(i,\delta,\epsilon)} |A_n^{f_i}(x) - \int f_i \, \mathrm{d}P| > \delta\Big\} \leq \epsilon. \tag{2}$$

2. ⇒ 1. Let P_0 be a measure satisfying 2.. There exists a computable function $m(i, \delta, \epsilon)$ such that (2) holds for P_0. Let $\mathscr{C}$ the class of measures P satisfying (2) for all $i, \delta > 0$ and $\epsilon > 0$. We prove that the complement of $\mathscr{C}$ is effectively open, which implies that $\mathscr{C}$ is effectively compact. Let i, δ, ϵ be fixed. The function $f(x, P) = \sup_{n \geq m(i,\delta,\epsilon)} |A_n^{f_i}(x) - \int f_i \, \mathrm{d}P|$ is lower semi-computable so the set $\{x : f(x, P) > \delta\}$ is effectively open in P, hence the function $P \mapsto P\{x : f(x, P) > \delta\}$ is lower semi-computable. As a result, the set of measures P for which (2) does not hold is effectively open, and this is uniform in i, δ, ϵ. The complement of $\mathscr{C}$ is the union over i, δ, ϵ of these uniformly effective open set, so it is effectively open.

Now, if $P \in \mathscr{C}$ then P is ergodic, as for each i, the Birkhoff averages $A_n^{f_i}$ converge P-almost everywhere to $\int f_i \, \mathrm{d}P$. By assumption, P_0 belongs to $\mathscr{C}$. □

6. Finitely decomposable measures

V'yugin [16] constructed a non-effectively decomposable measure, given by an infinite convex combination of ergodic measures. We now prove that finitely but non-effectively decomposable measures also exist, which settles a problem left open in [10].

The set of probability measures is endowed with the weak*-topology, which is induced by the following metric:

$$d(P, Q) = \sum_{w \in \{0,1\}^*} 2^{-|w|} |P[w] - Q[w]|.$$

This metric makes the set of probability measures a complete metric space, hence a Baire space.

The subset of stationary measures is closed in this topology, as P is stationary if and only if for every finite string w, $P[w] = P[0w] + P[1w]$. As a result, the metric subspace $\mathcal{S}$ of stationary measures is also complete and is also a Baire space. The set of ergodic measures is a dense G_δ-set in the subspace of stationary measures (see [14] for a proof). We endow $\mathcal{S} \times \mathcal{S}$ with the product topology.

Theorem 6.1. *There exist ergodic measures P, Q such that neither P nor Q is computable relative to $P + Q$.*

The set of such pairs (P, Q) is even co-meager in $\mathcal{S} \times \mathcal{S}$.

Intuitively, the existence of such measures is possible because P and Q do not depend continuously on $P + Q$: even a very good approximation of $P + Q$ does not give much information about P and Q. In particular a machine M cannot uniformly compute P from $P + Q$ for all stationary measures P, Q, as it can only compute continuous functions. The following lemma tells us much more: in the sense of Baire category, the set of pairs (P, Q) such that the machine M computes P from $P + Q$ is small, i.e. nowhere dense.

Let us first recall that M computes P from $P + Q$ if for every name of $P + Q$ provided as an oracle to M, it computes a name for P, which equivalently means that on inputs $w \in \{0, 1\}^*$ and $\delta \in \mathbb{Q}_{>0}$, the machine outputs a rational q such that $|q - P[w]| < \delta$. We will say for short that M^{P+Q} computes P.

Lemma 6.1. *Let M be a machine. In $\mathcal{S} \times \mathcal{S}$, the interior of the set*

$$C_M := \{(P, Q) \in \mathcal{S} \times \mathcal{S} : M^{P+Q} \text{ does not compute } P\}$$

is dense.

Proof. We prove that the interior of C_M intersects every non-empty basic open set $U \times V$ of the product topology on $\mathcal{S} \times \mathcal{S}$. We can assume that U is disjoint from V, otherwise we replace them by disjoint open subsets $U' \subseteq U$ and $V' \subseteq V$ (no stationary measure is isolated in $\mathcal{S}$ so neither U nor V is a singleton).

If $U \times V$ is contained in C_M then we are done. Otherwise there exist stationary measures $P_0 \in U$ and $Q_0 \in V$ such that $M^{P_0+Q_0}$ computes P_0. As U is disjoint from V, $P_0 \neq Q_0$. Let $w \in \{0, 1\}^*$ and $\delta > 0$ be such that $|P_0[w] - Q_0[w]| > \delta$. As U and V are open there exists $r > 0$ such that $B(P_0, r) \subseteq U$ and $B(Q_0, r) \subseteq V$. Let $\eta \in (0, 1)$ be such that $\eta d(P_0, Q_0) < r$.

We define a pair (P_1, Q_1) of stationary measures lying in $U \times V$ and in the interior of C_M. This pair is defined as

$$P_1 = (1 - \eta)P_0 + \eta Q_0,$$
$$Q_1 = (1 - \eta)Q_0 + \eta P_0.$$

First, $d(P_0, P_1) = d(Q_0, Q_1) = \eta d(P_0, Q_0) < r$ so $P_1 \in U$ and $Q_1 \in V$. Let $\epsilon < \eta\delta$ be a positive rational number. Let us consider the open set

$$W = \{(P, Q) : M^{P+Q}(w, \epsilon/2) \text{ halts and outputs some } q \text{ with } |q - P[w]| > \epsilon/2\}.$$

More precisely, (P, Q) belongs to W if there exists a representation of $P+Q$ on which the machine behaves as specified. First observe that W is contained in C_M. We claim that $(P_1, Q_1) \in W$. As $P_1+Q_1 = P_0+Q_0$, $M^{P_1+Q_1}$ computes P_0 so on input $(w, \epsilon/2)$ it halts and outputs some q with $|q - P_0[w]| < \epsilon/2$. As $|P_0[w]-P_1[w]| = \eta|P_0[w]-Q_0[w]| > \eta\delta > \epsilon$, $|q-P_1[w]| > \epsilon/2$, so $(P_1, Q_1) \in W$. As a result, $U \times V$ intersects W which is contained in the interior of C_M. □

Proof of Theorem 6.1. On $\mathcal{S}$, the set of ergodic measures is a dense G_δ-set, so on $\mathcal{S}\times\mathcal{S}$ the set of pairs of ergodic measures is also a dense G_δ-set. From the preceding lemma, the set of pairs (P, Q) such that P is not computable relative to $P+Q$ contains a dense G_δ-set, namely the intersection of the interiors of the sets C_M, for M varying among all the machines. As a result, the intersection of these two sets is co-meager in $\mathcal{S}\times\mathcal{S}$. By symmetry, the set of pairs of ergodic measures (P, Q) such that Q is not computable relative to $P+Q$ is also co-meager in $\mathcal{S}\times\mathcal{S}$. Therefore the intersection of the three sets is co-meager in $\mathcal{S}\times\mathcal{S}$. □

6.1. Positive results

Let P be a finite combination of ergodic measures. Even if Theorem 6.1 shows that its decomposition may not be computable, its finite character still have interesting consequences.

Proposition 6.1. *Let P be a stationary measure such that m_P is supported on a closed set $\mathscr{C}$ of stationary ergodic measures. For every P-random x, Q_x is ergodic.*

To prove it we use the following lemma.

Lemma 6.2. *Let X, Y be computable metric spaces. Let $f_n : X \to Y$ be uniformly computable functions that converge P-a.e. to a function f. Let $A \subseteq Y$ be a closed set such that $f(x) \in A$ for P-a.e. x. For every P-random x, $\lim f_n(x) \in A$.*

Proof. It is already known if f is constant P-almost everywhere. Let x_0 be a P-random point such that $\lim f_n(x_0) \notin A$. Let $B(y, r)$ be a ball with computable center and radius, containing $\lim f_n(x_0)$ and disjoint from A. Let $g_n(x) = \max(0, r - d(f_n(x), y))$. For P-almost every x, the sequence $g_n(x)$ converges to 0, but $\lim g_n(x_0) = r - d(\lim f_n(x_0), y) > 0$, which is impossible. □

Proof of Proposition 6.1. For every n, define $Q_n : X \to \mathcal{P}(X)$ by $Q_n(x) = \frac{1}{n}(\delta_x + \ldots + \delta_{T^{n-1}x})$. A sequence x is generic if and only if $Q_n(x)$ is weakly convergent, and in that case Q_x is the limit of $Q_n(x)$. The functions Q_n are uniformly computable. As $Q_x \in \mathscr{C}$ for P-almost every x, $Q_x \in \mathscr{C}$ for every P-random x by Lemma 6.2. □

As a direct application of Proposition 6.1, if P has a finite decomposition, i.e. if $P = \sum_{i=1}^n \alpha_i P_i$ where $\alpha_i \in [0, 1]$, $\sum_i \alpha_i = 1$ and all P_i are ergodic, then regardless of the computability of P, α_i, P_i, for every P-random x, $Q_x \in \{P_1, \ldots, P_n\}$ as the latter set is closed. In this particular case, Q_x is always m_P-random as m_P is concentrated on the P_i's.

7. Open questions

As mentioned in the introduction, ergodic measures have better computability properties than non-ergodic ones. Theorem 6.1 shows that finite combinations of ergodic measures may not be effectively decomposable, but Proposition 6.1 shows that they still have some of the interesting properties of effectively decomposable measures. Many other computability properties of finitely decomposable measures could be investigated. For instance, if P and Q are ergodic and x is Martin-Löf random w.r.t. $\frac{P+Q}{2}$, is x random w.r.t. P or Q?

Another open question is whether Theorem 6.1 has a constructive version making $P + Q$ computable.

Acknowledgements

We wish to thank Jeremy Avigad, Peter Gács and Cristóbal Rojas for useful discussions and suggestions, and the anonymous referee for helpful comments.

References

[1] Jeremy Avigad, Philipp Gerhardy, and Henry Towsner. Local stability of ergodic averages. *Transactions of the American Mathematical Society*, 362:261–288, 2010.

[2] Laurent Bienvenu, Adam Day, Ilya Mezhirov, and Alexander Shen. Ergodic-type characterizations of algorithmic randomness. In Fernando Ferreira, Benedikt Löwe, Elvira Mayordomo, and Luís Mendes Gomes, editors, *CiE*, volume 6158 of *Lecture Notes in Computer Science*, pages 49–58. Springer, 2010.

[3] Laurent Bienvenu and Benoît Monin. Von Neumann's biased coin revisited. In *LICS*, pages 145–154. IEEE, 2012.

[4] Johanna Franklin and Henry Towsner. Randomness and non-ergodic systems. Submitted, 2012.

[5] Cameron E. Freer and Daniel M. Roy. Computable exchangeable sequences have computable de Finetti measures. In Klaus Ambos-Spies, Benedikt Löwe, and Wolfgang Merkle, editors, *CiE*, volume 5635 of *Lecture Notes in Computer Science*, pages 218–231. Springer, 2009.

[6] Peter Gács. Uniform test of algorithmic randomness over a general space. *Theoretical Computer Science*, 341:91–137, 2005.

[7] Peter Gács. Lecture notes on descriptional complexity and randomness. Technical report, Boston University, 2008. http://www.cs.bu.edu/faculty/gacs/papers/ait-notes.pdf.

[8] Peter Gács, Mathieu Hoyrup, and Cristobal Rojas. Randomness on computable probability spaces - A dynamical point of view. In Susanne Albers and Jean-Yves Marion, editors, *STACS*, LIPIcs, pages 469–480. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, Germany, 2009.

[9] Peter Hertling and Klaus Weihrauch. Random elements in effective topological spaces with measure. *Information and Computation*, 181(1):32–56, 2003.

[10] Mathieu Hoyrup. Randomness and the ergodic decomposition. In Benedikt Löwe, Dag Normann, Ivan N. Soskov, and Alexandra A. Soskova, editors, *CiE*, volume 6735 of *Lecture Notes in Computer Science*, pages 122–131. Springer, 2011.

[11] Mathieu Hoyrup and Cristobal Rojas. Computability of probability measures and Martin-Löf randomness over metric spaces. *Inf. Comput.*, 207(7):830–847, 2009.

[12] Leonid A. Levin. On the notion of a random sequence. *Soviet Mathematics Doklady*, 14:1413–1416, 1973.

[13] Per Martin-Löf. The definition of random sequences. *Information and Control*, 9(6):602–619, 1966.

[14] K.R. Parthasarathy. On the category of ergodic measures. *Ill. J. Math.*, 5:648–656, 1961.

[15] R. R. Phelps. *Lectures on Choquet's Theorem.* Springer, Berlin, 2001.

[16] Vladimir V. V'yugin. Effective convergence in probability and an ergodic theorem for individual random sequences. *SIAM Theory of Probability and Its Applications*, 42(1):39–50, 1997.

[17] Klaus Weihrauch. *Computable Analysis.* Springer, Berlin, 2000.