Randomness on Computable Probability Spaces - A Dynamical Point of View

Peter Gacs, Mathieu Hoyrup, Cristobal Rojas

Symposium on Theoretical Aspects of Computer Science 2009 (Freiburg), pp. 469–480
www.stacs-conf.org

RANDOMNESS ON COMPUTABLE PROBABILITY SPACES—A DYNAMICAL POINT OF VIEW

PETER GÁCS [1] AND MATHIEU HOYRUP [2] AND CRISTÓBAL ROJAS [3]

[1] Computer Science Department, Boston University

[2] Département d'Informatique, École Normale Supérieure de Paris

[3] DI École Normale Supérieure and CREA École Polytechnique, Paris.

Abstract. We extend the notion of randomness (in the version introduced by Schnorr) to computable Probability Spaces and compare it to a *dynamical* notion of randomness: typicality. Roughly, a point is *typical* for some dynamic, if it follows the statistical behavior of the system (Birkhoff's pointwise ergodic theorem). We prove that a point is Schnorr random if and only if it is typical for every *mixing* computable dynamics. To prove the result we develop some tools for the theory of computable probability spaces (for example, morphisms) that are expected to have other applications.

1. Introduction

The roots of algorithmic randomness go back to the work of von Mises in the 20th century. He suggested a notion of individual infinite random sequence based on *limit-frequency* properties invariant under the action of *selection* functions from some "acceptable" set. The problem was then to properly define what an "acceptable" selection function could be. Some years later, the concept of *computable* function was formalized, providing a natural class of functions to be considered as acceptable. This gave rise to Church's notion of *computable randomness*. Nevertheless, substantial understanding was achieved only with the works of Kolmogorov [7], Martin-Löf [8], Levin [17] and Schnorr [9] and since then, many efforts have contributed to the development of this theory which is now well established and intensively studied.

There are several different possible definitions, but it is Martin-Löf's one which has received most attention. This notion can be defined, at least, from three different points of view:

(1) *measure theoretic.* This was the original presentation by Martin-Löf ([8]). Roughly, an infinite sequence is random if it satisfies all "effective" probabilistic laws (see definition 3.21).

1998 ACM Subject Classification: Theory of Computation (F.0), Probability and Statistics (G.3), Information Theory (H.1.1).

Key words and phrases: Schnorr Randomness, Birkhoff's ergodic theorem, computable measures.

PARTLY SUPPORTED BY ANR GRANT 05 2452 260 OX

(2) *compressibility.* This characterization of random sequences, due to Schnorr and Levin (see [17, 10]), uses the prefix-free Kolmogorov complexity: random sequences are those which are maximally complex.
(3) *predictability.* In this approach (started by Ville [13] and reintroduced to the modern theory by Schnorr [10]) a sequence is random if, in a fair betting game, no "effective" strategy ("martingale") can win an unbounded amount of money against it.

In [9], a somewhat broader notion of algorithmic randomness (narrower notion of probabilistic law) was proposed: Schnorr randomness. This notion received less attention over the years: Martin-Löfs definition is simpler, leads to universal tests, and many equivalent characterizations (besides, Schnorr's book is not in English...). Recently, Schnorr randomness has begun to receive more attention. The work [2] for instance, characterizes it in terms of Kolmogorov complexity.

In the present paper, first we extend Schnorr randomness to arbitrary computable probability spaces and develop some useful tools. Then, taking a *dynamical systems* point of view, we introduce yet another approach to the definition of randomness: typicality. Roughly, a point is *typical* for some measure-preserving ergodic dynamic, if it follows the statistical behavior of the system (given by Birkhoff's pointwise ergodic theorem) with respect to every bounded continous function used to follow its trajectory (or equivalently, every computable function, see Definition 3.28). We then show that:

Theorem. *In any computable probability space, a point is Schnorr random if and only if it is typical for every* mixing *computable dynamical system.*

The paper is organized as follows: Section 2 presents all needed concepts of computability theory and computable measure theory over general metric spaces. Parts of this section, for example on μ-computable functions, are new and should be of independent interest. Section 3.1 generalizes Schnorr randomness and studies some useful properties, after which we introduce the notion of typicality. Section 3.3 is devoted to the proof of our main result.

2. Computability

In classical recursion theory, a set of natural numbers is called ***recursively enumerable*** (***r.e.*** for short) if it is the range of some partial recursive function. That is if there exists an algorithm listing (or enumerating) the set.

Strictly speaking, recursive functions only work on natural numbers, but this can be extended to the objects (thought of as "finite" objects) of any countable set, once a numbering of its elements has been chosen. We will sometimes use the word ***algorithm*** instead of *recursive function* when the inputs or outputs are interpreted as finite objects. The operative power of an algorithm on the objects of such a numbered set obviously depends on what can be effectively recovered from their numbers.

Examples 2.1.
1 $\mathbb{N}^k$ can be numbered in such a way that the k-tuple of number i can be computed from i and vice versa.
2 The set $\mathbb{Q}$ of rational numbers can be injectively numbered $\mathbb{Q} = \{q_0, q_1, \ldots\}$ in an *effective* way: the number i of a rational a/b can be computed from a and b, and vice versa. We fix such a numbering.

All through this work, we will use recursive functions over numbered sets to define *computability* or *constructivity* notions on *infinite* objects. Depending on the context, these notions will take particulars names (computable, recursively enumerable, r.e. open, decidable, etc...) but the definition will be always of the form: *obect x is* **constructive** *if there exists a recursive $\varphi\colon \mathbb{N} \to D$ satisfying* property *$P(\varphi, x)$* (where D is some numbered set).

For example, *$E \subset \mathbb{N}$ is* **r.e.** *if there exists a recursive $\varphi\colon \mathbb{N} \to \mathbb{N}$ satisfying $E =$* range(φ).

Each time, a *uniform version* will be implicitly defined: *a sequence $(x_i)_i$ is constructive* **uniformly in i** *if there exists a recursive $\varphi\colon \mathbb{N} \times \mathbb{N} \to D$ satisfying* property *$P(\varphi(i, \cdot), x_i)$ for all i.*

In our example, *a sequence $(E_i)_i$ is* **r.e. uniformly in i** *if there exists $\varphi\colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ satisfying $E_i =$* range$(\varphi(i, \cdot))$ *for all i.*

Let us ilustrate this in the case of reals numbers (computable reals numbers were introduced by Turing in [11]).

Definition 2.2. A real number $x \in \mathbb{R}$ is said to be ***computable*** if there exists a total recursive $\varphi : \mathbb{N} \to \mathbb{Q}$ satisfying $|x - \varphi(n)| < 2^{-n}$ for all $n \in \mathbb{N}$.

Hence by a sequence of reals $(x_i)_i$ computable ***uniformly in i*** we mean that there exists a recursive $\varphi : \mathbb{N} \times \mathbb{N} \to \mathbb{Q}$ satisfying $|x - \varphi(i, n)| < 2^{-n}$ for all $n \in \mathbb{N}$, for all $i \in \mathbb{N}$.

We also have the following notions:

Definition 2.3. Let x be a real number. We say that:
- x is ***lower semi-computable*** if the set $\{i \in \mathbb{N} : q_i < x\}$ is r.e.,
- x is ***upper semi-computable*** if the set $\{i \in \mathbb{N} : q_i > x\}$ is r.e.,

It is easy to see that a real number is computable if and only if it is lower and upper semi-computable.

2.1. Computable metric spaces

We breifly recall the basic of computable metric spaces.

Definition 2.4. A ***computable metric space*** (CMS) is a triple $\mathcal{X} = (X, d, \mathcal{S})$, where
- (X, d) is a separable complete metric space.
- $\mathcal{S} = (s_i)_{i \in \mathbb{N}}$ is a numbered dense subset of X (called ***ideal points***).
- The real numbers $(d(s_i, s_j))_{i,j}$ are all computable, uniformly in i, j.

Some important examples of computable metric spaces:

Examples 2.5.
1 The Cantor space $(\Sigma^{\mathbb{N}}, d, S)$ with Σ a finite alphabet. If $x = x_1 x_2 \ldots$, $y = y_1 y_2 \ldots$, are elements then the distance is defined by $d(x, y) = \sum_{i : x_i \neq y_i} 2^{-i}$. Let us fix some element of Σ denoting it by 0. The dense set S is the set of ultimately 0-stationary sequences.
2 $(\mathbb{R}^n, d_{\mathbb{R}^n}, \mathbb{Q}^n)$ with the Euclidean metric and the standard numbering of $\mathbb{Q}^n$.

For further examples we refer to [15].

The numbered set of ideal points $(s_i)_i$ induces the numbered set of ***ideal balls*** $\mathcal{B} := \{B(s_i, q_j) : s_i \in S, q_j \in \mathbb{Q}_{>0}\}$. We denote by $B_{\langle i,j \rangle}$ (or just B_n) the ideal ball $B(s_i, q_j)$, where $\langle \cdot, \cdot \rangle$ is a computable bijection between tuples and integers.

Definition 2.6 (Computable points)**.** A point $x \in X$ is said to be ***computable*** if the set $E_x := \{i \in \mathbb{N} : x \in B_i\}$ is r.e.

Definition 2.7 (R.e. open sets)**.** We say that the set $U \subset X$ is ***r.e. open*** if there is some r.e. set $E \subset \mathbb{N}$ such that $U = \bigcup_{i \in E} B_i$. If U is r.e. open and $D \subset X$ is an arbitrary set then the set $A := U \cap D$ is called ***r.e. open in D***.

Examples 2.8.
1 If the sequence $(U_n)_n$ is r.e. open uniformly in n, then the union $\bigcup_n U_n$ is an r.e. open set.
2 $U_i \cup U_j$ and $U_i \cap U_j$ are r.e. open uniformly in (i, j). See [5].

Let (X, S_X, d_X) and (Y, S_Y, d_Y) be computable metric spaces. Let $(B_i^Y)_i$ be the collection of ideal balls from Y.

Definition 2.9 (Computable Functions)**.** A function $T : X \to Y$ is said to be ***computable*** if $T^{-1}(B_i^Y)$ is r.e. open uniformly in i.

It follows that computable functions are continuous. Since we will work with functions which are not necessarily continuous everywhere (and hence not computable), we shall consider functions which are computable on some subset of X. More precisely, a function T is said to be ***computable on D*** ($D \subset X$) if $T^{-1}(B_i^Y)$ is r.e. open in D, uniformly in i. The set D is called the ***domain of computability*** of T.

3. Computable Probability Spaces

Let us recall some basic concepts of measure theory. Let X be a set. A family $\mathfrak{B}$ of subsets of X is called an ***algebra*** if (i)$X \in \mathfrak{B}$, (ii)$A \in \mathfrak{B} \Rightarrow A^{\mathcal{C}} \in \mathfrak{B}$ and (iii) $A, B \in \mathfrak{B} \Rightarrow A \cup B \in \mathfrak{B}$. We say that $\mathfrak{B}$ is a ***σ-algebra*** if moreover $A_i \in \mathfrak{B}, i \geq 1 \Rightarrow \bigcup_i A_i \in \mathfrak{B}$. If $\mathfrak{B}_0$ is a family of subsets of X, the σ-algebra generated by $\mathfrak{B}_0$ (denoted $\sigma(\mathfrak{B}_0)$) is defined to be the smallest σ-algebra over X that contains $\mathfrak{B}_0$. If $\mathfrak{B}$ is a σ-algebra of subsets of X, we say that $\mu : \mathfrak{B} \to [0, 1]$ is a ***probability measure*** if, for every family $(A_i)_i \subset \mathfrak{B}$ of disjoint subsets of X, the following holds:

$$\mu\Big(\bigcup_i A_i\Big) = \sum_i \mu(A_i). \tag{3.1}$$

If X is a topological space, the ***Borel*** σ-algebra of X is defined as the σ-algebra generated by the family of open sets of X. Sets in the Borel σ-algebra are called Borel sets. In this paper, a ***probability space*** will always refer to the triple $(X, \mathfrak{B}, \mu)$, where $\mathfrak{B}$ is the Borel σ-algebra of X and μ is a probability measure. A set $A \subset X$ has ***measure zero*** if there is a Borel set A_1 such that $A \subset A_1$ and $\mu(A_1) = 0$. We call two sets $A_1, A_2 \subset X$ ***equivalent modulo zero***, and write $A_1 = A_2 \pmod 0$, if the symmetric difference has measure zero. We write $A_1 \subset A_2 \pmod 0$ if A_1 is a subset of A_2 and $A_1 = A_2 \pmod 0$.

When X is a computable metric space, the space of probability measures over X, denoted by $\mathcal{M}(X)$, can be endowed with a structure of computable metric space. Then a computable measure can be defined as a computable point in $\mathcal{M}(X)$.

Example 3.1 (Measure over a Cantor space)**.** As a special example, we can set $X = \mathbb{B}^{\mathbb{N}}$ where $\mathbb{B} = \{0, 1\}$ and $\lambda([x]) = 2^{-|x|}$, where $|x|$ is the length of the binary string $x \in \{0, 1\}^*$.

This is the distribution on the set of infinite binary sequences obtained by tossing a fair coin, and condition (3.1) simplifies to

$$\lambda(x0) + \lambda(x1) = \lambda(x).$$

Let $\mathcal{X} = (X, d, S)$ be a computable metric space. Let us consider the space $\mathcal{M}(X)$ of measures over X endowed with weak topology, that is:

$$\mu_n \to \mu \text{ iff } \mu_n f \to \mu f \text{ for all real continuous bounded } f,$$

where μf stands for $\int f \, \mathrm{d}\mu$.

If X is separable and complete, then $\mathcal{M}(X)$ is separable and complete. Let $D \subset \mathcal{M}(X)$ be the set of those probability measures that are concentrated in finitely many points of S and assign rational values to them. It can be shown that this is a dense subset ([1]).

We consider the Prokhorov metric ρ on $\mathcal{M}(X)$ defined by:

$$\rho(\mu, \nu) := \inf\{\epsilon \in \mathbb{R}^+ : \mu(A) \leq \nu(A^\epsilon) + \epsilon \text{ for every Borel set } A\}$$

where $A^\epsilon = \{x : d(x, A) < \epsilon\}$.

This metric induces the weak topology on $\mathcal{M}(X)$. Furthermore, it can be shown that the triple $(\mathcal{M}(X), D, \rho)$ is a computable metric space (see [3], [5]).

Definition 3.2. A measure μ is computable if it is a computable point of $(\mathcal{M}(X), D, \rho)$

The following result (see [5]) will be intensively used in the sequel:

Lemma 3.3. *A probability measure μ is computable if and only if the measure of finite union of ideal balls $\mu(B_{i_1} \cup \ldots \cup B_{i_k})$ is lower semi-computable, uniformly in $i_1, \ldots, i_k$.*

Definition 3.4. A ***computable probability space (CPS)*** is a pair $(\mathcal{X}, \mu)$ where $\mathcal{X}$ is a computable metric space and μ is a computable Borel probability measure on X.

As already said, a computable function defined on the whole space is necessarily continuous. But a transformation or an observable need not be continuous at every point, as many interesting examples prove (piecewise-defined transformations, characteristic functions of measurable sets,...), so the requirement of being computable everywhere is too strong. In a measure-theoretical setting, the natural weaker condition is to require the function to be computable *almost everywhere*. In the computable setting this is not enough, and a computable condition on the set on which the function is computable is needed:

Definition 3.5 (Constructive G_δ-sets)**.** We say that the set $D \subset X$ is a ***constructive G_δ-set*** if it is the intersection of a sequence of uniformly r.e. open sets.

Definition 3.6 (μ-computable functions)**.** Let $(\mathcal{X}, \mu)$ and $\mathcal{Y}$ be a CPS and a CMS respectively. A function $f : (\mathcal{X}, \mu) \to Y$ is ***μ-computable*** if it is computable on a constructive G_δ-set (denoted as $\mathrm{dom} f$ or D_f) of measure one.

Example 3.7. Let m be the Lebesgue measure on $[0, 1]$. The binary expansion of reals defines a function from non-dyadic numbers to infinite binary sequences which induces a m-computable function from $([0, 1], m)$ to $\{0, 1\}^{\mathbb{N}}$.

Remark 3.8. Given a uniform sequence of μ-computable functions $(f_i)_i$, any computable operation $\odot_{i=0}^n f_i$ (adition, multiplication, composition, etc...) is μ-computable, uniformly in n.

We recall that $F : (\mathcal{X}, \mu) \to (\mathcal{Y}, \nu)$ is measure-preserving if $\mu(F^{-1}(A)) = \nu(A)$ for all Borel sets A.

Definition 3.9 (morphisms of CPS's)**.** A ***morphism of CPS's*** $F : (\mathcal{X}, \mu) \to (\mathcal{Y}, \nu)$, is a μ-computable measure-preserving function $F : D_F \subseteq X \to Y$.

An ***isomorphism of CPS's*** $(F, G) : (\mathcal{X}, \mu) \rightleftarrows (\mathcal{Y}, \nu)$ is a pair (F, G) of morphisms such that $G \circ F = \mathrm{id}$ on $F^{-1}(D_G)$ and $F \circ G = \mathrm{id}$ on $G^{-1}(D_F)$.

Example 3.10. Let $(\mathbb{B}^{\mathbb{N}}, \lambda)$ the probability space introduced in Example 3.1 with the coin-tossing distribution λ over the infinite sequences. The binary expansion (see example 3.7) creates an isomorphism of CPS's between the spaces $([0, 1], m)$ and $(\mathbb{B}^{\mathbb{N}}, \lambda)$.

Remark 3.11. To every isomorphism of CPS's (F, G) one can associate the canonical invertible morphism of CPS's $\varphi = F|_{D_\varphi}$ with $\varphi^{-1} = G|_{D_{\varphi^{-1}}}$, where $D_\varphi = F^{-1}(G^{-1}(D_F))$ and $D_{\varphi^{-1}} = G^{-1}(D_F)$. Of course, (φ, φ^{-1}) is an isomorphism of CPS's as well.

The next proposition is a direct consequence of theorem 5.1.1 from [5]:

Proposition 3.12. *Every computable probability space is isomorphic to the Cantor space with an appropiate computable measure.*

Definition 3.13. A set $A \subset X$ is said to be ***almost decidable*** if the function $1_A : X \to \{0, 1\}$ is μ-computable.

It is easy to see that a set A is almost decidable iff there is a constructive G_δ set D of measure one and two r.e. open sets U and V such that:

$$U \cap D \subset A, \quad V \cap D \subseteq A^{\mathcal{C}}, \quad \mu(U) + \mu(V) = 1.$$

Remarks 3.14.

1 The collection of almost decidable sets is an algebra.
2 An almost decidable set is always a continuity set.
3 Ideal balls with zero boundary measure are always almost decidable.
4 Unless the space is disconnected (i.e. has non-trivial clopen subsets), no set can be *decidable*, i.e. semi-decidable (r.e.) and with a semi-decidable complement (such a set must be clopen[1]). Instead, a set can be decidable *with probability* 1: there is an algorithm which decides if a point belongs to the set or not, for almost every point. This is why we call it *almost decidable*.

Ignoring computability, the existence of open sets with zero boundary measure directly follows from the fact that the collection of open sets is uncountable and μ is finite. The problem in the computable setting is that there are only countable many open r.e. sets. Fortunately, there still always exists a basis of almost decidables balls.

Lemma 3.15. *Let X be $\mathbb{R}$ or $\mathbb{R}^+$ or $[0, 1]$. Let μ be a computable probability measure on X. Then there is a sequence of uniformly computable reals $(x_n)_n$ which is dense in X and such that $\mu(\{x_n\}) = 0$ for all n.*

[1] In the Cantor space for example (which is totally disconnected), every cylinder (ball) is a decidable set. Indeed, to decide if some infinite sequence belongs to some cylinder it suffices to compare the finite word defining the cylinder to the corresponding finite prefix of the infinite sequence.

Proof. Let I be a closed rational interval. We construct $x \in I$ such that $\mu(\{x\}) = 0$. To do this, we construct inductively a nested sequence of closed intervals J_k of measure $< 2^{-k+1}$, with $J_0 = I$. Suppose $J_k = [a, b]$ has been constructed, with $\mu(J_k) < 2^{-k+1}$. Let $m = (b - a)/3$: one of the intervals $[a, a + m]$ and $[b - m, b]$ must have measure $< 2^{-k}$, and since their measure is upper-computable, we can find it effectively—let it be J_{k+1}.

From a constructive enumeration $(I_n)_n$ of all the dyadic intervals, we can construct $x_n \in I_n$ uniformly. ∎

Corollary 3.16. *Let $(\mathcal{X}, \mu)$ be a CPS and $(f_i)_i$ be a sequence of uniformly computable real valued functions on X. Then there is a sequence of uniformly computable reals $(x_n)_n$ which is dense in $\mathbb{R}$ and such that $\mu(\{f_i^{-1}(x_n)\}) = 0$ for all i, n.*

Proof. Consider the uniformly computable measures $\mu_i = \mu \circ f_i^{-1}$ and define $\nu = \sum_i 2^{-i}\mu_i$. By Lemma 3.3, ν is a computable measure and then, by Lemma 3.15, there is a sequence of uniformly computable reals $(x_n)_n$ which is dense in $\mathbb{R}$ and such that $\nu(\{x_n\}) = 0$ for all n. Since $\nu(A) = 0$ iff $\mu_i(A) = 0$ for all i, we get $\mu(\{f_i^{-1}(x_n)\}) = 0$ for all i, n. ∎

The following result will be used many times in the sequel.

Corollary 3.17. *There is a sequence of uniformly computable reals $(r_n)_{n\in\mathbb{N}}$ such that $(B(s_i, r_n))_{i,n}$ is a basis of almost decidable balls.*

Proof. Apply Corollary 3.16 to $(f_i)_i$ defined by $f_i(x) = d(s_i, x)$. ∎

We remark that every ideal ball can be expressed as a r.e. union of almost decidable balls, and vice-versa. So the two bases are constructively equivalent.

Definition 3.18. A computable probability space is a ***computable Lebesgue space*** if it is isomorphic to the computable probability space $([0, 1], m)$ where m is the Lebesgue measure.

Theorem 3.19. *Every computable probability space with no atoms is a computable Lebesgue space.*

Proof. We first prove the result for $I = ([0, 1], \mu)$.

Lemma 3.20. *The interval endowed with a non-atomic computable probability measure is a computable Lebesgue space.*

Proof. We define the morphism of the CPS as $F(x) = \mu([0, x])$. As μ has no atom and is computable, F is computable and surjective. As F is surjective, it has right inverses. Two of them are $G_<(y) = \sup\{x : F(x) < y\}$ and $G_>(y) = \inf\{x : F(x) > y\}$, and satisfy $F^{-1}(\{y\}) = [G_<(y), G_>(y)]$. They are increasing and respectively left- and right-continuous. As F is computable, they are even lower- and upper semi-computable respectively. Let us define $D = \{y : G_<(y) = G_>(y)\}$: every $y \in D$ has a unique pre-image by F, which is then injective on $F^{-1}(D)$. The restriction of F on $F^{-1}(D)$ has a left-inverse, which is given by the restriction of $G_<$ and $G_>$ on D. Let us call it $G : D \to I$. By lower and upper semi-computability of $G_<$ and $G_>$, G is computable. Now, D is a constructive G_δ-set: $D = \bigcap_n \{y : G_>(y) - G_<(y) < 1/n\}$. We show that $I \setminus D$ is a countable set. The family $\{[G_<(y), G_>(y)] : y \in I\}$ indexed by I is a family of disjoint closed intervals, included in $[0, 1]$. Hence, only countably many of them have positive length. Those intervals correspond to points y belonging to $I \setminus D$, which is then countable. It follows that D has Lebesgue measure one (it is even dense). (F, G) is then an isomorphism between (I, μ) and (I, m). ∎

Now, we know from Theorem 3.12 that every CPS $(\mathcal{X}, \mu)$ has a binary representation, which is in particular an isomorphism with the Cantor space $(\mathbb{B}^{\mathbb{N}}, \mu')$. As mentioned in Example 3.10, the latter is isomorphic to (I, μ_I) where μ_I is the induced measure. If μ is non-atomic, so is μ_I. By the previous lemma, (I, μ_I) is isomorphic to (I, m). ∎

3.1. Randomness and typicality

3.1.1. *Algorithmic randomness.*

Definition 3.21. A ***Martin-Löf test*** (ML-test) is an uniform sequence $(A_n)_n$ of r.e. open sets such that $\mu(A_n) \leq 2^{-n}$. We say that x ***fails*** the ML-test if $x \in A_n$ for all n. A point x is called ***ML-random*** if it fails no ML-test.

Definition 3.22. A ***Borel-cantelli test*** (BC-test) is a uniform sequence $(C_n)_n$ of r.e. open sets such that $\sum_n \mu(C_n) < \infty$. We say that x ***fails*** the BC-test if $x \in C_n$ infinitly often (i.o.).

It is easy to show that:

Proposition 3.23. *x fails a ML-test iff x fails a* BC*-test.*

Definition 3.24. A ***Schnorr test*** (Sch-test) is a ML-test $(A_n)_n$ such that the sequence of reals $(\mu(A_n))_n$ is uniformly computable. We say that x fails the Sch-test if $x \in A_n$ for all n. A point x is called Sch-***random*** if it fails no Sch-test.

Definition 3.25. A ***strong*** BC-***test*** is a BC-test $(C_n)_n$ such that $\sum_n \mu(C_n)$ is computable.

Proposition 3.26. *An element x fails a* Sch*-test if and only if x fails a strong* BC*-test.*

Proof. Let $(C_n)_n$ be a strong BC-test. Let c be such that $2^c > \sum_n \mu(C_n)$. Define the r.e. open set $A_k := \{x : |\{n : x \in C_n\}| \geq 2^{k+c}\}$. Then $\mu(A_k) < 2^{-k}$. Observe that A_k is the union of all the (2^{k+c})-intersections of C_n's. Since $\mu(C_k) = \sum_n \mu(C_n) - \sum_{n \neq k} \mu(C_n)$ and the C_n's are r.e. we have that $\mu(C_n)$ is computable (uniformly in n). We choose a basis $(B^i)_i$ of almost decidable balls to work with. Recall that finite unions or intersections of almost decidable sets are almost decidable too and that the measure of an almost decidable set is computable. Now we show that $\mu(A_k)$ is computable uniformly in k. Let $\epsilon > 0$ be rational. Let n_0 be such that $\sum_{n \geq n_0} \mu(C_n) < \frac{\epsilon}{2}$. Then $\mu(\bigcup_{n \geq n_0} C_n) < \frac{\epsilon}{2}$. For each C_n with $n < n_0$ we construct an almost decidable set $C_n^\epsilon \subset C_n$ (a finite union of almost decidable balls) such that $\mu(C_n) - \mu(C_n^\epsilon) < \frac{1}{n_0}\frac{\epsilon}{2}$. Then $\sum_{n<n_0}[\mu(C_n) - \mu(C_n^\epsilon] < \frac{\epsilon}{2}$. Define A_k^ϵ to be the union of the (2^{k+c})-intersections of the C_n^ϵ's for $n < n_0$. Then A_k^ϵ is almost decidable and then has a computable measure. Moreover $A_k \subset A_k^\epsilon \cup (\bigcup_{n \geq n_0} C_n) \cup (\bigcup_{n<n_0} C_n \setminus C_n^\epsilon)$, hence $\mu(A_k) - \mu(A_k^\epsilon) < \epsilon$. ∎

The following result is an easy modification of a result from [5], so we omit the proof.

Proposition 3.27. *Morphisms of computable probability spaces are defined (and computable) on Schnorr random points and preserve* Sch*-randomness.*

3.2. Dynamical systems and typicality

Let X be a metric space, let $T : X \mapsto X$ be a Borel map. Let μ be an invariant Borel measure on X, that is: $\mu(A) = \mu(T^{-1}(A))$ holds for each measurable set A. A set A is called T-invariant if $T^{-1}(A) = A$ modulo a set of measure 0. The system (T, μ) is said to be ergodic if each T-invariant set has total or null measure. In such systems the famous Birkhoff ergodic theorem says that time averages computed along μ-typical orbits coincide with space averages with respect to μ. More precisely, for any $f \in L^1(X)$ it holds

$$\lim_{n\to\infty} \frac{S_n^f(x)}{n} = \int f \, \mathrm{d}\mu, \tag{3.2}$$

for μ-almost each x, where $S_n^f = f + f \circ T + \ldots + f \circ T^{n-1}$.

If a point x satisfies equation (3.2) for a certain f, then we say that x is ***typical*** with respect to the ***observable*** f.

Definition 3.28. If x is typical w.r. to every bounded continuous function $f : X \to \mathbb{R}$, then we call it a T-***typical point***.

Remark 3.29. The proof of our main theorem will show as a side result that the definition would not change if we replaced "continuous" with "computable" in it.

In [14] is proved that ML-random infinite binary sequences are typical w.r. to any computable f. In [4], this is generalized via effective symbolic dynamics to computable probability spaces and μ-computable observables.

To have the result for Sch-random points it seems that a certain "mixing" property or "loss of memory" of the system has to be required. This is naturally expressed by means of the ***correlation functions***. For measureable functions f, g let

$$\begin{aligned} C(f, g) &= \mu(f \cdot g) - \mu f \cdot \mu g, \\ C_n(f, g) &= C(f \circ T^n, g). \end{aligned}$$

For events A, B with indicator functions $1_A, 1_B$ let

$$C_n(A, B) = C_n(1_A, 1_B),$$

which measures the dependence between the events A and B at times $n \gg 1$ and 0 respectively. Note that $C_n(A, B) = 0$ corresponds, in probabilistic terms, to $T^{-n}(A)$ and B being independent events.

Let us say that a family of Borel sets $\mathcal{E}$ is ***essential***, if for every open set U there is a sequence $(E_i)_i$ of borel sets in $\mathcal{E}$ such that $\cup_i E_i \subset U \pmod 0$ (see Section 3).

Definition 3.30. We say that a system (X, T, μ) is (polynomially) ***mixing*** if there is $\alpha > 0$ and an essential family $E = \{E_1, E_2, ...\}$ of almost decidable events such that for each i, j there is $c_{i,j} > 0$ computable in i, j such that

$$|C_n(E_i, E_j)| \leq \frac{c_{i,j}}{n^\alpha} \qquad \text{for all } n \geq 1.$$

We say that the system is ***independent*** if all correlation functions $C_n(E_i, E_j)$ are 0 for sufficiently large n.

Examples of non-mixing but still ergodic systems are given for instance by irrational circle rotations with the Lebesgue measure. Examples of mixing but not independent sytems are given by piecewise expandings maps or uniformly hyperbolic systems which have a

distinguished ergodic measure (called SRB measure and which is "physical" in some sense) with respect to which the correlations decay exponentially (see [12]). An example of a mixing system for which the decrease of correlations is only polynomial and not exponential, is given by the class of *Manneville-Pomeau* type maps (non uniformly expanding with an indifferent fixed point, see [6]). For a survey see [16].

3.3. Proof of the main result

Now we prove our main theorem.

Theorem 3.31. Let $(\mathcal{X}, \mu)$ be a computable probability space with no atoms The following properties of a point $x \in X$ are equivalent.

(i) x is Schnorr random.
(ii) x is T-typical for every mixing endomorphism T.
(iii) x is T-typical for every independent endomorphism T.

Remark 3.32. If the measure μ is atomic, it is easy to see that:

(1) $(\mathcal{X}, \mu)$ admits a mixing endomorphism if and only if $\mu = \delta_x$ for some x. In this case the theorem still holds, the only random point being x.
(2) $(\mathcal{X}, \mu)$ admits an ergodic endomorphism if and only if $\mu = \frac{1}{n}(\delta_{x_1} + ... + \delta_{x_n})$ (where $x_i \neq x_j$, for all $i \neq j$). In this case, a point x is Schnorr random if and only if it is typical for every ergodic endomorphism if and only if it is an atom.

Proof. Let us first prove a useful lemma. Let $E \subset X$ be a Borel set. Denote by 1_E its indicator function. The ergodic theorem says that the following equality holds for almost every point:

$$\lim_n \frac{1}{n} \sum_{i=0}^{n-1} 1_E \circ T^i(x) = \mu(E). \tag{3.3}$$

Lemma 3.33. *Let $\mathcal{E}$ be an essential family of events. If x satisfies equation (3.3) for all $E \in \mathcal{E}$ then x is a T-typical point.*

Proof. We have to show that equation (3.3) holds for any bounded continuous observable f. First, we extend equation (3.3) to every continuity open set C. Let $(E_i)_i$ be a sequence of elements of $\mathcal{E}$ such that $\bigcup_i E_i \subseteq \mathrm{Int}(C)$ and $\mu(\bigcup_i E_i) = \mu(C)$. Define $C_k = \bigcup_{i \leq k} E_i$. Then $\mu(C_k) \nearrow \mu(C)$. For all k:

$$\liminf_n \frac{1}{n} \sum_{i=0}^{n-1} 1_C \circ T^i(x) \geq \lim_n \frac{1}{n} \sum_{i=0}^{n-1} 1_{C_k} \circ T^i(x) = \mu(C_k)$$

so $\liminf_n \frac{1}{n} \sum_{i=0}^{n-1} 1_C \circ T^i(x) \geq \mu(C)$. Applying the same argument to $X \setminus C$ gives the result.

Now we extend the result to bounded continuous functions. Let f be continuous and bounded ($|f| < M$) and let $\epsilon > 0$ be a real number. Then, since the measure μ is finite, there exist real numbers $r_1, \ldots, r_k \in [-M, M]$ (with $r_1 = -M$ and $r_k = M$) such that $|r_{i+1} - r_i| < \epsilon$ for all $i = 1, \ldots, k-1$ and $\mu(f^{-1}(\{r_i\})) = 0$ for all $i = 1, \ldots, k$. It follows that for $i = 1, \ldots, k-1$ the sets $C_i = f^{-1}(]r_i, r_{i+1}[)$ are all continuity open sets.

Hence the function $f_\epsilon = \sum_{i=1}^{k-1} r_i 1_{C_i}$ satisfies $\|f - f_\epsilon\|_\infty \leq \epsilon$ and then the result follows by density. ∎

We are now able to prove that (i) ⇒ (ii).

Let $E \in \mathcal{E}$. Put $f = 1_E$. Observe that f is μ-computable. For $\delta > 0$, define the deviation sets:

$$A_n^f(\delta) = \left\{ x \in X : \left| \frac{S_n^f(x)}{n} - \int f \, \mathrm{d}\mu \right| > \delta \right\}.$$

By Corollary 3.16 we can choose δ such that $A_n^f(\delta)$ is almost decidable. Then their measures are computable, uniformly in n.

By the Chebychev inequality, $\mu(A_n^f(\delta)) \leq \frac{1}{\delta^2} \left\| \frac{S_n^f(x)}{n} - \int f \, \mathrm{d}\mu \right\|_{L^2}^2$. Let us change f by adding a constant to have $\int f \, \mathrm{d}\mu = 0$. This does not change the above quantity. Then, by invariance of μ we have

$$\left\| \frac{S_n^f(x)}{n} - \int f \, \mathrm{d}\mu \right\|_{L^2}^2 = \int \left(\frac{S_n^f(x)}{n} \right)^2 \mathrm{d}\mu = \frac{1}{n^2} \int n f^2 \, \mathrm{d}\mu + \frac{2}{n^2} \int \Big(\sum_{i<j<n} f \circ T^{j-i} f \Big) \, \mathrm{d}\mu$$

and hence

$$\delta^2 \mu(A_n^f(\delta)) \leq \frac{\|f\|_{L^2}^2}{n} + \frac{2}{n} \sum_{k<n} |C_k(f,f)| \leq \frac{\|f\|_{L^2}^2}{n} + \frac{2c_{f,f}}{(1-\alpha)n^\alpha}.$$

(Observe that α can be replaced by any smaller positive number, so we assume $\alpha < 1$.) Hence, $\mu(A_n^f(\delta)) \leq Cn^{-\alpha}$ for some constant C. Now, it is easy to find a sequence $(n_i)_{i\in\mathbb{N}}$ such that the subsequence $(n_i^{-\alpha})_i$ is effectively summable and $\frac{n_i}{n_{i+1}} \to 1$ (take for instance $n_i = i^\beta$ with $\alpha\beta > 1$). This shows that the sequence $A_{n_i}^f(\delta)$ is a strong BC-test. Therefore, if x is Sch-random then x belongs to only finitely many $A_{n_i}^f(\delta)$ for any δ and hence the subsequence $\frac{S_{n_i}^f(x)}{n_i}$ converges to $\int f \, \mathrm{d}\mu = \mu(E)$. To show that for such points the whole sequence $\frac{S_n^f(x)}{n}$ converges to $\int f \, \mathrm{d}\mu = \mu(E)$, observe that if $n_i \leq n < n_{i+1}$ and $\beta_i := \frac{n_i}{n_{i+1}}$ then we have:

$$\frac{S_{n_i}^f}{n_i} - 2(1-\beta_i)M \leq \frac{S_n^f}{n} \leq \frac{S_{n_{i+1}}^f}{n_{i+1}} + 2(1-\beta_i)M,$$

where M is a bound of f. To see this, for any k, l, β with $\beta \leq k/l \leq 1$:

$$\frac{S_k^f}{k} - \frac{S_l^f}{l} = \left(1 - \frac{k}{l}\right) \frac{S_k^f}{k} - \frac{S_{l-k}^f \circ T^{l-k}}{l} \leq (1-\beta)M + \frac{(l-k)M}{l} = 2(1-\beta)M.$$

Taking $\beta = \beta_i$ and $k = n_i$, $l = n$ first and then $k = n$, $l = n_{i+1}$ gives the result. Thus, we have proved that a Schnorr random point x satisfies equation (3.3) for any $E \in \mathcal{E}$. Lemma 3.33 allows to conclude.

The (ii) ⇒ (iii) part follows since any independent dynamic is in particular mixing.

To prove the (iii) ⇒ (i) part we will need the following proposition which is a strengthening of a result of Schnorr in [9]. The proof is somewhat technical, for lack of space we do not included here (see appendix).

Proposition 3.34. *If the infinite binary string $\omega \in (\mathbb{B}^\mathbb{N}, \lambda)$ is not Schnorr random (w.r. to the uniform measure), then there exists an isomorphism $\Phi : (\mathbb{B}^\mathbb{N}, \lambda) \to (\mathbb{B}^\mathbb{N}, \lambda)$ such that $\Phi(\omega)$ is not typical for the shift transformation σ.*

Now we are able to finish the proof of our main result: suppose that x is not Schnorr random. We construct a dynamic T for which x is not T-typical. From Proposition 3.12 and Theorem 3.19 we know that there is an isomorphism $\eta : (\mathcal{X}, \mu) \to (\mathbb{B}^{\mathbb{N}}, \lambda)$ (here, λ denotes the uniform measure). If $x \notin \mathrm{dom}(\eta)$, we can take any independent endomorphism and modify it in order to be the identity on x. It is cleary still an independent endomorphism (maybe with a smaller domain of computability) and x, being a fixed point, can't be T-typical. So let $x \in \mathrm{dom}(\eta)$. Then $\eta(x)$ is not Schnorr random in $(\mathbb{B}^{\mathbb{N}}, \lambda)$, since η as well as its inverse preserve Schnorr randomness. Then, by Proposition 3.34, $\Phi(\eta(x))$ is not σ-typical, where σ is the shift which is clearly independent (cylinders being the essential events). Put $\psi = \Phi \circ \eta$. Define the dynamics T on X by $T = \psi^{-1} \circ \sigma \circ \psi$. It is easy to see that T is independent for events of the form $E = \psi^{-1}[w]$. Since $\{\psi^{-1}[w] : w \in 2^*\}$ form an essential family of almost decidable events, T is independent too. As $\psi(x)$ is not σ-typical, x is not T-typical either. ∎

References

[1] Patrick Billingsley. *Convergence of Probability Measures*. John Wiley, New York, 1968.

[2] Rodney G. Downey and Evan J. Griffiths. Schnorr randomness. *Electr. Notes Theor. Comput. Sci.*, 66(1), 2002.

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

[4] Stefano Galatolo, Mathieu Hoyrup, and Cristóbal Rojas. Effective symbolic dynamics, random points, statistical behavior, complexity and entropy. *Submitted*, 2007.

[5] Mathieu Hoyrup and Cristóbal Rojas. Computability of probability measures and Martin-Löf randomness over metric spaces. *Information and Computation*, 2009. To appear.

[6] Stefano Isola. On systems with finite ergodic degree. *Far east journal of dynamical systems*, 5:1, 2003.

[7] Andrey N. Kolmogorov. Three approaches to the quantitative definition of information. *Problems in Information Transmission*, 1:1–7, 1965.

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

[9] Claus-Peter Schnorr. *Zufälligkeit und Wahrscheinlichkeit. Eine algorithmische Begründung der Wahrscheinlichkeitstheorie*, volume 218 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin-New York, 1971.

[10] Claus-Peter Schnorr. The process complexity and effective random tests. In *STOC*, pages 168–176, 1972.

[11] Alan Turing. On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, 2, 42:230–265, 1936.

[12] Marcelo Viana. Stochastic dynamics of deterministic systems. *Lecture Notes XXI Braz. Math. Colloq. IMPA Rio de Janeiro*, 1997.

[13] J. Ville. *Etude Critique de la Notion de Collectif*. Gauthier-Villars, Paris, 1939.

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

[15] Klaus Weihrauch. Computability on computable metric spaces. *Theoretical Computer Science*, 113:191–210, 1993. Fundamental Study.

[16] Lai-Sang Young. What are SRB measures, and which dynamical systems have them? *J. Stat. Phys.*, 108:733–754, 2002.

[17] A.K. Zvonkin and L.A. Levin. The complexity of finite objects and the development of the concepts of information and randomness by means of the theory of algorithms. *Russian Mathematics Surveys*, 256:83–124, 1970.