Hierarchies of Randomness Tests

Jan Reimann
Institut für Informatik
Ruprecht-Karls-Universität Heidelberg, Germany
Email: reimann@math.uni-heidelberg.de

Frank Stephan*
Departments of Mathematics and Department of Computer Science
National University of Singapore, Republic of Singapore
Email: fstephan@comp.nus.edu.sg

It is well known that Martin-Löf randomness can be characterized by a number of equivalent test concepts, based either on effective nullsets (Martin-Löf and Solovay tests) or on prefix-free Kolmogorov complexity (lower and upper entropy). These equivalences are not preserved as regards the partial randomness notions induced by effective Hausdorff measures or partial incompressibility. Tadaki [19] and Calude, Staiger and Terwijn [2] studied several concepts of partial randomness, but for some of them the exact relations remained unclear. In this paper we will show that they form a proper hierarchy of randomness notions, namely for any ρ of the form $\rho(x) = 2^{-|x|s}$ with s being a rational number satisfying $0 < s < 1$, the Martin-Löf ρ-tests are strictly weaker than Solovay ρ-tests which in turn are strictly weaker than strong Martin-Löf ρ-tests. These results also hold for a more general class of ρ introduced as unbounded premeasures.

1. Introduction

The correspondence between effective nullsets in the sense of measure theoretic tests and compressible initial segments in terms of (prefix-free) Kolmogorov complexity is one of the cornerstones of algorithmic information

* F. Stephan is supported in part by NUS research grant R252-000-212-112.

theory. Furthermore, the concept of randomness itself appears thereby very robust, as several variants of measure theoretic tests (Martin-Löf and Solovay tests) and complexity theoretic properties all yield the same notion of randomness; that is, a sequence A is Martin-Löf random iff one of the following equivalent conditions hold:

(1) A is not covered by any Martin-Löf test;
(2) A is not covered by any Solovay test;
(3) for some constant c and for all n, $K(A\restriction_n) \geq n - c$;
(4) $\lim_n K(A\restriction_n) - n = \infty$.

On the other hand, the complexity theoretic formulations suggests not only a qualitative, but also a quantitative classification of randomness. For instance, a sequence A for which $K(A\restriction_n) \geq n/2 + c$ for all n and some constant c might be classified as being 1/2-random.

This idea is, in particular, reflected in the study of *relative randomness* initiated by Solovay [16] and later leading to a variety of *reducibility notions* for random sequences. The forthcoming book by Downey and Hirschfeldt [5] will provide for a detailed account.

One may ask whether it is possible to catch "partial" randomness not only within a rather fine-grained hierarchy of relative randomness, but as an absolute notion in terms of measure theoretic tests and their complexity theoretic counterparts.

The problem here is that the property of being an (effective) nullset is rather qualitative. What is therefore needed is a further ramification of the effective Lebesgue nullsets (which constitute the non-random sequences).

Such a ramification can be given in terms of *Hausdorff measures*, which are an essential tool in fractal geometry. In particular, they allow for a definition of a non-integral notion of dimension, *Hausdorff dimension*. Works by Ryabko [13, 14], Staiger [18] and Cai and Hartmanis [1] established a close connection between Hausdorff dimension and the (lower) asymptotic complexity of sequences given as $\liminf_n K(A\restriction_n)/n$.

Later, Lutz [7] used the martingale characterization of nullsets to define an effective variant of Hausdorff measure and dimension. This yields a quantitative classification of randomness for individual sequences in terms of measure.

It is straightforward to transform the martingale approach to Hausdorff measures into a Martin-Löf-style test concept (as done by Reimann and Stephan [10], Tadaki [19] or Calude, Staiger and Terwijn [2]). Besides, one may also define partial randomness notions based on the other randomness criteria (2)-(4). It could be shown that in the generalized framework Martin-Löf tests still coincide with the complexity criterion derived from (3), referred to as *weak Chaitin randomness* by Tadaki [19] and Calude, Staiger and Terwijn [2]. Furthermore, Solovay tests and *strong Chaitin randomness* (4) still remain equivalent [2, 19].

However, it remained unclear whether the full robustness of (Lebesgue) randomness (in the sense of complete equivalence of all test notions) prevailed.

In this paper we show that this is indeed not the case. Not only with respect to the usual Hausdorff measures, but also a very wide family of measures given by *unbounded premeasures*, Martin-Löf tests and Solovay tests are not equivalent. Another test notion proposed by Calude, Staiger and Terwijn [2] called *strong Martin-Löf randomness* yields an even stronger notion of randomness.

The paper is structured as follows. In Section 2 we give a detailed introduction of effective tests derived from a general class of outer measures on Cantor space. Section 3 will treat the connection between tests and Kolmogorov complexity. Finally, in Section 4 we will show that the test notions introduced in Section 2 form a proper hierarchy of randomness notions for unbounded premeasures. The latter include the non-integral Hausdorff measures.

Notation: Most notation is standard. $\{0,1\}^\star$ denotes the set of finite binary strings, $\{0,1\}^\omega$ the set of all infinite binary strings. $\sqsubset$ is the partial prefix order on strings, which extends to $\{0,1\}^\star \cup \{0,1\}^\omega$ in a natural way. $x \sqsubseteq y$ holds if either $x \sqsubset y$ or $x = y$. Given a set $V \subseteq \{0,1\}^\star$ and a string x, we write V_x for the set $\{w \in V : x \sqsubseteq w\}$ and V_x^+ for the set $\{w \in V : x \sqsubset w\}$.

We assume the reader is acquainted with the basic definitions and results of Recursion Theory and the theory of Kolmogorov complexity. We refer to the textbooks of Li and Vitányi [6], Odifreddi [11] and Soare [17] for any background on this.

2. Effective Randomness Tests for Outer Measures

There are mainly two ways to devise measures on Cantor space (as on any other suitable topological/metric space). One can start with an additive set function on a (semi)algebra of sets (usually comprising a basis of the topology) and then use Caratheodory's extension theorem, which ensures that there is a unique extension of this set function to a σ-algebra (which includes the Borel sets if the starting (semi)algebra included the basic open sets).

Alternatively, measures can be obtained by restricting outer measures to a suitable family of sets in $\{0,1\}^\omega$. Outer measures are often defined via *premeasures* and *coverings*. A premeasure is a non-negative (possibly infinite) set function ρ on a family $\mathfrak{C}$. In most cases, $\mathfrak{C}$ will consist of the family of *basic open cylinders* which are defined as

$$[\![x]\!] = \{X \in \{0,1\}^\omega : \, x \sqsubset X\}$$

Therefore, it is convenient to regard premeasures as functions

$$\rho : \{0,1\}^\star \to \mathbb{R}_0^+$$

from which one can obtain an outer measure μ_ρ by letting

$$\mu_\rho(\mathcal{X}) = \inf\left\{\sum_i \rho(x_i) : \, \bigcup_i [\![x_i]\!] \supseteq \mathcal{X}\right\}.$$

It is not hard to show that $\mu = \mu_\rho$ is a countably subadditive, monotone set function. If one restricts μ to those sets $\mathcal{A}$ which satisfy

$$(\forall \mathcal{Y})\,[\mu(\mathcal{Y}) = \mu(\mathcal{Y} \cap \mathcal{A}) + \mu(\mathcal{Y} \setminus \mathcal{A})],$$

called the *measurable sets*, the measurable sets form a σ-algebra and μ is an additive set function on this σ-algebra.

If the underlying space is a metric space, the method of passing from a premeasure to an outer measure can be refined in a geometrical way. The standard metric d on $\{0,1\}^\omega$ (which yields a topology compatible with the one generated by the cylinder sets defined above) is defined as

$$d(X,Y) = \inf\{2^{-n} : \, (\forall m < n)[X(m) = Y(m)]\}.$$

That is, if $X \neq Y$ then $d(X,Y) = 2^{-n}$ for the least n with $X(n) \neq Y(n)$. The diameter of a set $\mathcal{X} \subseteq \{0,1\}^\omega$ is $d(\mathcal{X}) = \sup\{d(X,Y) : \, X, Y \in \mathcal{X}\}$. If ρ is a premeasure, we can define

$$\mu_\delta(\mathcal{X}) = \inf\left\{\sum_i \rho(x_i) : \, \bigcup_i [\![x_i]\!] \supseteq \mathcal{X} \,\wedge\, (\forall i)[d([\![x_i]\!]) \leq \delta]\right\}$$

and

$$\mu^\rho(\mathcal{X}) = \sup\{\mu_\delta(\mathcal{X}) : \delta > 0\}.$$

Here it is the "fine covers" that determine the value of μ^ρ. It can be shown that μ^ρ is also an outer measure and that it behaves, in geometric sense, more stable than measures constructed via the first method. Note that $d([\![x_i]\!]) \leq \delta$ if and only if $|x| \geq -\log \delta$.

An extensive treatment of constructing measures via premeasures is found in the book by Rogers [12].

It can be shown that every *nullset*, a set for which μ^ρ takes the value zero, is measurable. It was Martin-Löf's groundbreaking idea to use the concept of an *effective nullset* to define a notion of randomness for individual sequences. Basically, a sequence is random with respect to a measure if it is not contained in an effectively presented nullset with respect to the measure. As the nullsets are precisely the sets which have outer measure zero, it suffices to study effective nullsets with respect to premeasures. It is not hard to see that Martin-Löf's approach works for arbitrary (outer) measures which are derived from computable premeasures.

A well-known group of outer measures is obtained from the premeasures $\rho(x) = 2^{-|x|s}$ where $0 \leq s \leq 1$, the s-dimensional Hausdorff measures. For $s = 1$, we obtain the uniform distribution $\rho(x) = 2^{-|x|}$, which generates a measure isomorphic to Lebesgue measure on the unit interval.

We will study a certain class of premeasures. These premeasures can be thought of as "geometrically well behaved". Among the measures they induce are the usual probability measures on $\{0,1\}^\omega$ as well as the family of s-dimensional Hausdorff measures. To be able to effectivize, we will always assume premeasures to be computable.

Definition 2.1: A (*geometrical*) *premeasure* is a computable function $\rho : \{0,1\}^\star \to \mathbb{R}_0^+$ such that $\rho(\epsilon) = 1$ and there are (computable) real numbers p, q with

- $1/2 \leq p < 1$ and $1 \leq q < 2$;
- $(\forall x \in \{0,1\}^\star)\,(\forall i \in \{0,1\})\,[\rho(xi) \leq p\rho(x)]$;
- $(\forall x \in \{0,1\}^\star)\,[q\rho(x) \leq \rho(x0) + \rho(x1)]$.

We will call such ρ a (p,q)-premeasure. ρ is called an *unbounded premeasure* if it is (p,q)-premeasure for some $q > 1$. A premeasure is called *length-invariant* if

$$(\forall x, y)\,[|x| = |y| \;\Rightarrow\; \rho(x) = \rho(y)]$$

and is called *additive* if

$$(\forall x \in \{0,1\}^\star)\,[\rho(x) = \rho(x0) + \rho(x1)].$$

Remarks and Examples: (a) Additive premeasures induce non-atomic *probability measure* on $\{0,1\}^\omega$. Among the probability measures induced by additive premeasures are standard uniform distribution $\rho(x) = 2^{-|x|}$, which we will denote by λ, as well as all non-degenerate *Bernoulli measures* μ_p with $0 < p < 1$. μ_p is the measure obtained by setting

$$\rho(x) = p^{|\{i:x(i)=0\}|}(1-p)^{|\{i:x(i)=1\}|}$$

for all x.

(b) It is not hard to see that, if ρ is unbounded, the corresponding measure μ constructed by the second method mentioned above will be infinite ($\mu(\{0,1\}^\omega) = \infty$). This motivates the term "unbounded" (see Proposition 2.4).

(c) The most common unbounded and length invariant premeasures are the functions of the form $\rho(x) = 2^{-|x|s}$ with $0 < s \leq 1$. The corresponding constants are $p = 2^{-s}$ and $q = 2^{1-s}$. These premeasures give rise to the *s-dimensional Hausdorff measures*. In general, length-invariant premeasures are often called *dimension functions*, as the induce a generalized type of Hausdorff measure.

(d) Other less orthodox examples of premeasures are $\rho(x) = p(|x|)2^{-s|x|}$, where p is a suitable polynomial, or $\rho(x) = 2^{-p|\{i:x(i)=0\}|-(1-p)|\{i:x(i)=1\}|}$, where p is a real number satisfying $0 < p < 1$.

In the following, if we consider measures derived from premeasures, we will always assume they are constructed via the second method. We will be particularly interested in *nullsets*, that is, sets for which μ^ρ takes the value zero. The following proposition states some equivalent characterization of nullsets for a measure μ^ρ. These will be important for the definition of effective nullsets later on. It will be convenient to introduce some further notation: Given $W \subseteq \{0,1\}^\star$, let $\rho(W)$ stand for $\sum_{x\in W} \rho(x)$.

Proposition 2.2: *Given a premeasure ρ and a set $\mathcal{X} \subseteq \{0,1\}^\omega$, the following are equivalent:*

(a) $\mu^\rho(\mathcal{X}) = 0$;

(b) *For every $n \in \mathbb{N}$ there exists a set $U_n \subseteq \{0,1\}^\star$ such that*

$$\mathcal{X} \subseteq [\![U_n]\!] \quad \text{and} \quad \rho(U_n) \leq 2^{-n};$$

(c) *There exists a set $W \subseteq \{0,1\}^\star$ such that $\rho(W) < \infty$ and for any $X \in \mathcal{X}$ there are infinitely many $w \in W$ such that $w \sqsubset X$.*

Proof: By definition, $\mu^\rho(\mathcal{X})$ is the infimum of all numbers $\rho(W)$ with $\mathcal{X} \subseteq [\![W]\!]$. This gives the equivalence of (a) and (b). Furthermore, taking W to be the union of all U_n gives the direction from (b) to (c).

For the missing direction from (c) to (b), let $w_0, w_1, \ldots$ be an one-one enumeration of W. Now one defines $U_n = \{w_m, w_{m+1}, \ldots\}$ for the first m such that $\sum_{k \geq m} \rho(w_k) < 2^{-n}$. This m exists as $\sum_{k \geq 0} \rho(w_k)$ is finite. As for every $X \in \mathcal{X}$ there are infinitely many k with $w_k \sqsubset X$, it holds that $X \in [\![U_n]\!]$ for all n. This completes the direction from (c) to (b). □

We now introduce the effective variants of measure zero sets and the corresponding randomness concepts. This generalizes earlier work. The underlying tests are defined as below but were mostly restricted to $\rho(x) = 2^{-|x|}$ (or some additive premeasure, i.e. probability measures on $\{0,1\}^\omega$). Martin-Löf [8] introduced the notion of effective tests now named after him. Solovay [16] showed that $X \in \{0,1\}^\omega$ is Martin-Löf random iff there is no r.e. set W of strings such that $\rho(W) < \infty$ and $w \sqsubset X$ for infinitely many $w \in W$. Schnorr [15] contrasted Martin-Löfs general condition to more restrictive randomness-tests where X is random iff there is no uniformly enumerable sequence $(U_n)_{n\in\mathbb{N}}$ of sets with $X \in [\![U_n]\!]$ and $\rho(U_n) = 2^{-n}$ for all n. This condition had also many natural characterizations and is now known as Schnorr randomness. Based on this work, Lutz [7] initiated the study effectivizations of Hausdorff measures via modified martingales. Building on these notions, further investigations were carried out by Tadaki [19], Reimann [9], Calude, Staiger and Terwijn [2].

Definition 2.3: Let $\rho : \{0,1\}^\star \to \mathbb{R}_0^+$ be a geometrical premeasure.

(a) A *Martin-Löf ρ-test* is a uniformly enumerable sequence $(U_n)_{n\in\mathbb{N}}$ of sets of strings such that

$$(\forall n)[\rho(U_n) \leq 2^{-n}].$$

The test $(U_n)_{n\in\mathbb{N}}$ *covers* a set $\mathcal{X} \subseteq \{0,1\}^\omega$ if

$$\mathcal{X} \subseteq \bigcap_{n\in\mathbb{N}} [\![U_n]\!].$$

In this case $\mathcal{X}$ is called *Martin-Löf ρ-null*. A sequence $A \in \{0,1\}^\omega$ is called *Martin-Löf ρ-random* if $\{A\}$ is not covered by any Martin-Löf ρ-test.

(b) A *strong Martin-Löf ρ-test* is a uniformly enumerable sequence $(U_n)_{n\in\mathbb{N}}$ such that

$$(\forall n)(\forall V \subseteq U_n)[V \text{ prefix-free } \Rightarrow \rho(V) \leq 2^{-n}].$$

Again, the test $(U_n)_{n\in\mathbb{N}}$ *covers* $\mathcal{X} \subseteq \{0,1\}^\omega$ if

$$\mathcal{X} \subseteq \bigcap_{n\in\mathbb{N}} [\![U_n]\!].$$

Accordingly, a $\mathcal{X}$ is called *strongly Martin-Löf ρ-null* and A is called *strongly Martin-Löf ρ-random* if $\{A\}$ is not covered by any strong Martin-Löf ρ-test.

(c) A *Solovay ρ-test* is a recursively enumerable set W such that

$$\rho(W) < \infty.$$

The test W *covers* $\mathcal{X} \subseteq \{0,1\}^\omega$ if, for any $X \in \mathcal{X}$, W contains infinitely many prefixes of X. $A \in \{0,1\}^\omega$ is called *Solovay ρ-random* if it is not covered by any Solovay ρ-test.

Note that the name "strong Martin-Löf test" might be misleading at first, since every Martin-Löf test is also a strong Martin-Löf test. In fact, the use of "strong" makes more sense from the viewpoint of random sequences since every strongly Martin-Löf ρ-random sequence is also Martin-Löf ρ-random.

Also note that for additive premeasures, the notions of Martin-Löf ρ-randomness, Solovay ρ-randomness and strong Martin-Löf ρ-randomness coincide, yielding an effective analog of Proposition 2.2. If ρ is additive, any Solovay ρ-test W can be converted effectively into a Martin-Löf ρ-test (V_n) by letting x enter V_n if and only if 2^n proper prefixes of x have been enumerated into W prior to x. Furthermore, we can pass from a strong Martin-Löf ρ-test (U_n) to an ordinary Martin-Löf ρ-test (V_n) covering the same set of sequences by the following effective procedure:

If x is enumerated into U_n at some stage, check whether a prefix of x has already been enumerated into V_n. If so, discard x. Otherwise check whether some extensions of x have already been enumerated into V_n. If not, enumerate x into V_n. Otherwise assume W is the finite set of strings extending x and already in V_n. Enumerate into V_n a finite, prefix-free set $\widetilde{W}$ such that $[\![\widetilde{W}]\!] \cap [\![W]\!] = \emptyset$ and $[\![\widetilde{W}]\!] \cup [\![W]\!] = [\![x]\!]$.

The so constructed V_n covers the same set of sequences as a (maximal, with respect to covering) prefix-free subset V of U_n. If ρ is additive, $\rho(V_n) = \rho(V) \leq 2^{-n}$.

It is not clear how these procedures may be transferred to unbounded premeasures, since a string may not be substituted by a cover-equivalent set of longer strings of the same ρ-measure. In fact, we will show later that this is in general not possible.

We conclude this section by showing that unbounded, length-invariant premeasures induce measures that are incompatible with the uniform distribution λ.

Proposition 2.4: *For every unbounded premeasure ρ there exists a set $\mathcal{X}$ such that some Martin-Löf λ-test covers all sequences $X \in \mathcal{X}$ but $\mathcal{X}$ does not have μ^ρ-measure zero.*

Proof: Assume ρ is an unbounded (p, q)-premeasure. Let $R_n(x) = \{xy : y \in \{0,1\}^n\}$. It follows from an easy induction that,

$$\rho(R_n(x)) \geq q^n \rho(x) \quad \text{and} \quad (\forall w \in R_n(x))\,[\rho(w) \leq p^n \rho(x)]. \tag{2.1}$$

Let $U_0 = \{\epsilon\}$. Given U_n, we can use (2.1) and the fact that ρ is computable to (effectively) find prefix-free sets V_n^0, V_n^1 such that $[\![V_n^i]\!] \subset [\![U_n]\!]$, $[\![V_n^0]\!] \cap [\![V_n^1]\!] = \emptyset$ and $\rho(V_n^i) \geq 1$. But obviously, for some i we must have $\lambda([\![V_n^i]\!]) \leq \frac{1}{2}\lambda([\![U_n]\!])$. Pick such i and let $U_{n+1} = V_n^i$. Then, for $\mathcal{U} = \bigcap [\![U_n]\!]$, by choice of the sets $U_0, U_1, \ldots$, the test $(U_n)_{n\in\mathbb{N}}$ is a Martin-Löf λ-test and covers $\mathcal{U}$ but $\mu^\rho(\mathcal{U}) \neq 0$. □

As a corollary we get that the (weak) randomness notions with respect to λ on the one hand and with respect to unbounded measure functions on the other hand differ.

Corollary 2.5: *For every unbounded premeasure ρ there exists a sequence X such that X is Martin-Löf ρ-random but not Martin-Löf λ-random.*

Proof: Any set of non-zero μ^ρ-measure has to contain a Martin-Löf ρ-random sequence, so the set $\mathcal{U}$ constructed above has to contain one, but cannot contain a Martin-Löf λ-random sequence, for it is covered by some Martin-Löf λ-test. □

3. Nullsets and Kolmogorov complexity

In this section we will study to what extent the correspondence between effective nullsets and compressibility extends from the well-known characterization of randomness with respect to the uniform distribution to nullsets with respect to (length-invariant) premeasures.

For this purpose, we generalize a definition by Chaitin [3] to arbitrary premeasures. For the case $\rho(x) = 2^{-|x|s}$, this was first done by Tadaki [19].

Definition 3.1: Let ρ be a premeasure.

(a) A sequence $A \in \{0,1\}^\omega$ is *weakly Chaitin ρ-random* if there exists a constant c such that

$$(\forall n)\,[K(A\restriction_n) \geq -\log\rho(A\restriction_n) - c].$$

(b) A set $A \in \{0,1\}^\omega$ is *strongly Chaitin ρ-random*[a] if

$$(\forall c)(\overset{\infty}{\forall} n)[K(A\restriction_n) \geq -\log\rho(A\restriction_n) + c].$$

The well-known Kolmogorov characterization for random sequences using prefix-free complexity K generalizes to a characterization for Martin-Löf ρ-nullsets. For one direction, though, it seems one has to presuppose length-invariance.

In the case $\rho(x) = 2^{-|x|s}$, the following two propositions were shown by Tadaki [19]. Reimann [9] obtained related results for the more general case of computable *dimension functions*.

Proposition 3.2: *Let ρ be a premeasure. If $A \in \{0,1\}^\omega$ is weakly Chaitin ρ-random then it is also Martin-Löf ρ-random.*

Proof: Assume A is not Martin-Löf ρ-random. Thus there exists a computable sequence $C_1, C_2, C_3, \ldots$ of enumerable sets of strings such that for all n

$$(\exists w \in C_n)\,[w \sqsubset A] \quad \text{and} \quad \rho(C_n) \leq 2^{-n}. \tag{3.1}$$

Define functions $m_n : \{0,1\}^\star \to \mathbb{Q}$ by

$$m_n(w) = \begin{cases} n\rho(w) & \text{if } w \in C_n, \\ 0 & \text{otherwise,} \end{cases}$$

and let

$$m(w) = \sum_{n=1}^{\infty} m_n(w).$$

Obviously, all m_n and thus m are enumerable from below. Furthermore,

$$\sum_{w\in\{0,1\}^\star} m(w) = \sum_{w\in\{0,1\}^\star}\sum_{n=1}^{\infty} m_n(w) = \sum_{n=1}^{\infty} n \sum_{w\in C_n}\rho(w) \leq \sum_{n=1}^{\infty}\frac{n}{2^n} < \infty.$$

[a] The notion of strong Chaitin ρ-randomness should not be confused with strong Chaitin randomness as defined in [5], meaning $(\exists^\infty n)[K(A\restriction_n) \geq n + K(n) - \mathcal{O}(1)]$.

It follows from the fundamental *Coding Theorem*, due to Levin (see the book by Li and Vitányi [6]), that there exists a constant c_m such that $m(w) \leq c_m 2^{-K(w)}$ for almost every w. Now let $c > 0$ be any constant. If we set $k = \lceil c \rceil + 1$, then, by (3.1), there is some $w \in C_k$ with $w \sqsubset A$, say $w = A\restriction_n$. This implies $m(A\restriction_n) \geq k\rho(A\restriction_n) > c\rho(A\restriction_n)$ and therefore $\limsup m(A\restriction_n)/\rho(A\restriction_n) = \infty$. Hence A is not weakly Chaitin ρ-random. □

For length-invariant premeasures, we can prove the converse of Proposition 3.2.

Proposition 3.3: *Let ρ be a length-invariant premeasure. If $A \in \{0,1\}^\omega$ is Martin-Löf ρ-random, then it is weakly Chaitin ρ-random.*

Proof: The proof is an adaptation of a standard proof that every Marti-Löf random sequence is incompressible with respect to prefix-free Kolmogorov complexity (see for instance the book by Downey and Hirschfeldt [5]). It is based on a fundamental result by Chaitin [4] which establishes that for any l,

$$|\{x \in \{0,1\}^n : K(x) \leq n + K(n) - l\}| \leq 2^{n+c-l}, \tag{3.2}$$

where c is a constant independent of n, l. Remember that the natural numbers are identified with their binary representation.

Let ρ be a length-invariant premeasure and assume A is not weakly Chaitin ρ-random. As ρ is length-invariant, there exists a computable function $h : \mathbb{N} \to \mathbb{R}_0^+$ such that

$$h(n) = \alpha \iff (\forall x)[|x| = n \Rightarrow \rho(x) = \alpha].$$

Now choose a c for which (3.2) holds. Define

$$V_n = \{x \in \{0,1\}^\star : K(x) \leq -\log h(x) - n - c\}.$$

Then each V_n covers A, since for every l there is some prefix x of A such that $K(x) \leq -\log \rho(x) - l$. Furthermore, each V_n is r.e., since K is enumerable from above. Finally, using (3.2), we have for each n,

$$\begin{aligned}\sum_{w \in V_n} \rho(w) = \sum_{k=0}^{\infty} \sum_{\substack{w \in V_n \\ |w|=k}} \rho(w) &= \sum_{k=0}^{\infty} h(k) |\{0,1\}^k \cap V_n| \\ &\leq 2^{-n} \sum_{k=0}^{\infty} 2^{-K(k)} \leq 2^{-n}\end{aligned}$$

and this bound completes the proof. □

It is also possible to characterize Solovay randomness in terms of complexity. Namely, Solovay and strong Chaitin randomness coincide. The following proposition generalizes a result by Tadaki [19] and Calude, Staiger and Terwijn [2].

Proposition 3.4: *Given a premeasure ρ, a sequence $A \in \{0,1\}^\omega$ is Solovay ρ-random if and only if it is strongly Chaitin ρ-random.*

Proof: Assume A is not strongly Chaitin ρ-random. Then there is a constant c such that for infinitely many n, $\rho(A\restriction_n) < 2^{c-K(A\restriction_n)}$; without loss of generality c is a natural number. Then the set $W = \{x : \rho(x) < 2^{c-K(x)}\}$ is recursively enumerable since ρ is computable, K is enumerable from above and thus $2^{c-K(x)}$ is enumerable from below. Furthermore, $\rho(W) < \sum_{x\in W} 2^{c-K(x)} < 2^c$. So W is a Solovay ρ-test. By choice of c, W covers every A. So A is not Solovay ρ-random.

The converse direction can easily be seen by using the Kraft-Chaitin Theorem (see for instance Li and Vitányi's book [6]). Assume A is covered by a Solovay ρ-test W. Then $\sum_{x\in W} \rho(x)$ is finite and one can consider $\{\langle x, -\log \rho(x)\rangle : x \in W\}$ as a Kraft-Chaitin axiom set. Thus there is a constant c with $K(x) < c - \log \rho(x)$. Since A is covered by W, there are infinitely many n with $K(A\restriction_n) < c - \log \rho(A\restriction_n)$. $\square$

4. A Hierarchy of Randomness Tests

Since a Martin-Löf ρ-test (V_n) can be transformed into a Solovay ρ-test W covering all the sets covered by (V_n) by letting

$$W = \bigcup_{n\in\mathbb{N}} V_n,$$

one obtains that every Martin-Löf ρ-nullset is contained in a Solovay ρ-nullset. Thus every Solovay ρ-random set is also Martin-Löf ρ-random.

Proposition 4.1: *For every premeasure ρ, every Martin-Löf ρ-nullset is covered by a Solovay ρ-test.*

In the case that $\rho = \lambda$, the converse direction of above proposition is also true. The next result shows that this is not the case when ρ is an unbounded premeasure.

Theorem 4.2: *For every unbounded premeasure ρ there exists a sequence A which is Solovay ρ-null but not Martin-Löf ρ-null.*

Proof: Assume ρ is an unbounded (p,q)-premeasure. Let $F(x) = K(x) + \log \rho(x)$. Note that for all $x \in \{0,1\}^*$ and $i \in \{0,1\}$, $(q-p)\rho(x) \leq \rho(xi) \leq p\rho(x)$ and $q - p > 0$. Thus there is a constant bounding the absolute value of the difference $\rho(xi) - \rho(x)$ for all $x \in \{0,1\}^*$ and $i \in \{0,1\}$. The same applies to K and F. Using Lemma 4.3 below, there is a constant c and a sequence A which satisfies for all n the following three properties:

- if $F(A\restriction_{cn}) \geq 0$ then $F(A\restriction_{cn+c}) \leq F(A\restriction_{cn}) - 1$;
- if $F(A\restriction_{cn}) < 0$ then $F(A\restriction_{cn+c}) \geq F(A\restriction_{c_n}) + 1$.

It follows that the there is a constant c' with $-c' \leq F(A\restriction_m) \leq c'$ for all m. So A is weakly but not strongly Chaitin ρ-random. It follows from the results in the previous section that A is Solovay ρ-null and Martin-Löf ρ-random. □

Lemma 4.3: *Let ρ be an unbounded premeasure. There exists a constant c with the following property: For all strings x there exist strings y, z of length c such that*

$$K(xy) + \log \rho(xy) \geq K(x) + \log \rho(x) + 1;$$
$$K(xz) + \log \rho(xz) \leq K(x) + \log \rho(x) - 1.$$

Proof: Again assume ρ is an unbounded (p,q)-premeasure. We first outline an algorithm which treats c as an input. The result is obtained by fixing c to a sufficiently large value.

Given x, c, the strings $y, z \in \{0,1\}^c$ are determined as follows. Order the 2^c strings in $\{0,1\}^c$ according to the size of $\rho(xu)$, $u \in \{0,1\}^c$, in descending order and let, for $b \in \{1,2,\ldots,2^c\}$, $g(x,c,b)$ be the b-th string in this ordering prefixed by x, so $\rho(g(x,c,1)) \geq \rho(g(x,c,2)) \geq \ldots \geq \rho(g(x,c,2^c))$. As ρ is recursive, one can take g to be a recursive function as well. Note that

$$\begin{aligned} q^c\rho(x) &\leq \sum_{u\in\{0,1\}^c} \rho(xu) \\ &\leq \rho(g(x,c,1)) + \sum_{a\in\{0,1,\ldots,c-1\}} 2^a \cdot \rho(g(x,c,2^a+1)) \end{aligned}$$

and $\rho(g(x,c,1)) \leq p^c\rho(x)$. Now choose $a \in \{0,1,\ldots,c-1\}$ such that the $2^a \cdot \rho(g(x,c,2^a+1))$ is maximal and choose y such that $xy = g(x,c,b)$ for the $b \in \{1,2,\ldots,2^a\}$ for which $K(g(x,c,b))$ is maximal. Furthermore, let $z = 0^c$, that is, z consists of c zeroes. Now, the following statements hold for all sufficiently large c, all x and the a, y, z chosen for them as above.

- $K(xy) \geq K(x) + a - K(c) - K(a) - \log(c) \geq K(x) + a - 4\log(c)$ where c has to be sufficiently large so that $\log(c)$ absorbs the constants involved;
- $\rho(xy) \geq q^c \rho(x) 2^{-a} \cdot c^{-2}$ where c is sufficiently large that $q^c - p^c \geq q^c/c$;
- $K(xy) + \log\rho(xy) \geq K(x) + \log\rho(x) + \log(q)c - 6\log(c)$ where $\log(q) > 0$ as ρ is unbounded;
- $K(xz) \leq K(x) + K(c) + \log(c) \leq K(x) + 3\log(c)$ as z can be computed from x and c and $\log(c)$ absorbs the involved constants if c is sufficiently large;
- $\log\rho(xz) \leq \log\rho(x) + c\log(p)$ where $\log(p) < 0$;
- $K(xz) + \log\rho(xz) \leq K(x) + \log\rho(x) + c\log(p) + 3\log(c)$.

Now one can choose a constant c sufficiently large so that all of the above hold, that $\log(q)c - 6\log(c) > 1$, and that $c\log(p) + 3\log(c) < -1$. With c thus chosen, one can find for every x some $y, z \in \{0,1\}^c$ with the desired properties. □

Though not equivalent, Solovay ρ-tests and Martin-Löf ρ-tests induce the same notion of effective dimension. They distinguish between sequences on a rather "fine" level of complexity oscillations. For details on effective dimension see e.g. [9].

Proposition 4.4: *Let ρ be a premeasure. If $X \in \{0,1\}^\omega$ is covered by an effective Solovay ρ-test, then X is Martin-Löf ρ'-null for any premeasure ρ' such that* $\lim_n \rho'(X\restriction_n)/\rho(X\restriction_n) = 0$.

Proof: Assume that C is an effective Solovay ρ-cover for X and let ρ' be a premeasure with $\lim_n \rho'(X\restriction_n)/\rho(X\restriction_n) = 0$. Deleting a finite number of strings from C does not change the covering properties of a Solovay test, so we may assume that $\sum_{w \in C} \rho(w) \leq 1$. Given $n \geq 0$, we define a r.e. set C_n by enumerating only those elements of C for which

$$\frac{\rho'(w)}{\rho(w)} \leq 2^{-n}.$$

Then X is covered by C_n and it holds that

$$\sum_{w \in C_n} \rho(w) = \sum_{w \in C_n} \frac{\rho'(w)}{\rho(w)} \rho(w) \leq 2^{-n} \sum_{w \in C_n} \rho(w) \leq 2^{-n}.$$

Hence, (C_n) is a Martin-Löf ρ'-test for X. □

On the other hand, Solovay tests can always be covered by strong Martin-Löf tests.

Theorem 4.5: *For every premeasure ρ, every Solovay ρ-nullset is covered by a strong Martin-Löf ρ-test.*

Proof: Let W be a Solovay ρ-test. Again, we may assume that $\rho(W) < 1$. Assume further that for $A \in \{0,1\}^\omega$ there are infinitely many $w \in W$ such that $w \sqsubset A$. We distinguish two cases:

Case 1: It holds that

$$(\forall n)(\exists x \in W)[x \sqsubset A \ \wedge\ \rho(W_x^+) \geq \rho(x)2^n]. \tag{4.1}$$

Define

$$V_n = \{x \in W :\ \rho(W_x^+) \geq \rho(x)2^n\}.$$

We claim that $(V_n)_{n\in\mathbb{N}}$ is a strong Martin-Löf ρ-test that covers A. Obviously, A is covered by each V_n, due to the assumption above. Furthermore, if $V \subseteq V_n$ is prefix-free, then

$$2^n\rho(V) = 2^n \sum_{x\in V} \rho(x) \leq \sum_{x\in V} \rho(W_x^+) \leq \sum_{w\in W} \rho(w) < 1,$$

so $\rho(V) < 2^{-n}$.

Case 2: We have

$$(\exists n)(\forall x \in W)[x \sqsubset A \ \Rightarrow\ \rho(W_x^+) < \rho(x)2^n]. \tag{4.2}$$

We can strengthen this to

$$\begin{aligned} &(\exists^\infty x \in W)[x \sqsubset A \ \wedge\ \rho(W_x^+) > r\rho(x)] \\ \text{and}\ \ &(\forall^\infty x)[x \sqsubset A \ \Rightarrow\ \rho(W_x^+) < (r + \tfrac{1}{2})\rho(x)], \end{aligned}$$

where r is some rational number. By removing finitely many elements from W we can even assume

$$(\forall z \sqsubset A)[\rho(W_x^+) < (r + \tfrac{1}{2})\rho(x)].$$

Now let $x_0, x_1, \ldots$ be an enumeration of W and construct inductively sets $T_n \subseteq \{0,1\}^\star$ starting with $T_0 = \{x_0\}$. If $Q = T_n \cup \{x_{n+1}\}$ satisfies

$$(\forall y \in Q)[\rho(Q_y^+) < (r + \tfrac{1}{2})\rho(y)]$$

then let $T_{n+1} = T_n \cup \{x_n\}$ else let $T_{n+1} = T_n$. It is easy to see that the resulting union $T = \bigcup T_n$ satisfies

$$(\forall y \in T)[\rho(T_y^+) < (r + \tfrac{1}{2})\rho(y)].$$

Furthermore, every prefix of A in W is also in T. To see this, assume that x_n is a prefix of A. If $n = 0$ then $x_n \in T$ anyway. If $n > 0$ then consider

$Q = T_{n-1} \cup \{x_n\}$ and any $y \in Q$. If $y \not\sqsubseteq x_n$ then $\rho(Q_y^+) = \rho((Q - \{x_n\})_y^+)$ and y does prevent x_n from being added to T. If $y \sqsubseteq x_n$ then $y \sqsubset A$ and $\rho(Q_y^+) \leq \rho(W_y^+) \leq (r + \frac{1}{2})\rho(y)$ and again y does prevent x_n from being enumerated into T. Thus $x_n \in T_n$ and $x_n \in T$. So all prefixes of A in W are also in T and T covers A. The set T is obviously enumerable.

From T one enumerates $S = \{x \in T : \rho(W_x^+) > r\rho(x)\}$. The set S contains infinitely many prefixes of A. Furthermore, for every $x \in S$, every prefix-free subset Q of S_x^+ and every $y \in Q$, the following inequalities hold.

$$\begin{aligned}
\rho(W_y^+) &> r\rho(y); \\
\rho(W_x^+) &\geq \rho(Q) + \sum_{y \in Q} \rho(W_y^+) \ > \ (1+r)\rho(Q); \\
\rho(W_x^+) &\leq (r + \tfrac{1}{2})\rho(x); \\
\rho(Q) &\leq \tfrac{1+2r}{2+2r}\rho(x).
\end{aligned}$$

Now let inductively

$$Q_m = \{x \in S : \forall y \in S\, (y \sqsubset x \Rightarrow y \in \cup_{k<m} Q_k)\,\}.$$

So Q_m is the set of all $x \in S$ such that the cardinality of $\{y \in S : y \sqsubset x\}$ is exactly m. Now $\rho(Q_{m+1}) \leq \frac{2r+1}{2r+2}\rho(Q_m)$ for all m. Now one can compute numbers $m_0, m_1, m_2, \ldots$ such that for all n, $(\frac{2q+1}{2q+2})^{m_n} < 2^{-n}$ and thus $\rho(Q_{m_n}) < 2^{-n}$. Now let

$$V_n = \{x \in S : \exists y \in Q_{m_n}\, (y \sqsubset x)\} = S - Q_0 - Q_1 - Q_2 - \ldots - Q_{m_n}.$$

The sets V_n are uniformly enumerable. Furthermore, if Q is a prefix-free subset of V_n, then there is for every $x \in Q$ some $y \in Q_{m_n}$ with $y \sqsubset x$. Thus Q is the union of prefix free sets Q_y^+ with $y \in Q_{m_n}$. By choice of S and V_n, $\rho(Q_y^+) < \rho(y)$ and $\rho(Q) < \rho(Q_{m_n}) < 2^{-n}$. This completes the proof. □

Corollary 4.6: *For $s \geq 0$ and ρ_s given as $\rho_s(x) = 2^{-|x|s}$, every Solovay ρ_s-nullset is covered by a strong Martin-Löf ρ_s-test.*

For unbounded, length-invariant premeasures, strong Martin-Löf tests are strictly more powerful then Solovay tests.

Theorem 4.7: *For any unbounded, length-invariant premeasure ρ there is a set A which is covered by a strong Martin-Löf ρ-test but not by a Solovay ρ-test.*

Proof: Let $I_0, I_1, \ldots$ be a recursive sequence of disjoint intervals such that

$$(\forall i)\,(\exists j \in I_i)\,(\forall k)\,[|\{x \in \{0,1\}^j : K(x) \leq k\}| < 2^{k-2i}].$$

Using Lemma 4.3, we can construct a sequence A such that, up to a constant c,

$$(\forall i)\,(\forall j \in I_i)\,[K(A\restriction_j) = -\log \rho(A\restriction_j) + i].$$

By construction A is Solovay ρ-random. Let

$$W_i = \{x \in \{0,1\}^{\max I_{i+c}} : \forall y \sqsubseteq x\,(K(y) \leq -\log \rho(y) + i + 2c)\}.$$

Obviously, every W_i is finite and covers A. Furthermore, the W_i are uniformly enumerable. We show how to modify (W_i) to obtain a strong Martin-Löf ρ-test that covers A.

We initialize with $V_i = \emptyset$. Every time some w is enumerated into W_i, we check whether there exists a $v \in V_i$ such that $v \sqsubseteq w$. If so, we let V_i unchanged. Otherwise we pick the longest $v \sqsubseteq w$ such that for all prefix-free subsets $Q \subseteq V_i \cup \{v\}$ and every $u \sqsubseteq w$ it holds that

$$\rho(Q_u^+) < \rho(u).$$

Enumerate v into V_i. It is clear that the V_i still cover A. It remains to show that for every prefix-free subset Q of V_i, $\rho(Q) \leq 2^{-i}$.

Let $j \in I_{i+c}$ be such that for all k,

$$|\{x \in \{0,1\}^j : K(x) \leq k\}| < 2^{k-2i}.$$

Consider the cover $U = \{u \in \{0,1\}^j : \exists w \in W_i (u \sqsubseteq w)\}$. ρ is length-invariant, so let r be the unique valle of all $\rho(u)$, $u \in \{0,1\}^j$. It follows by the choice of j that

$$\rho(U) \leq r2^{-\log(r)+i+2c-2(i+c)} = 2^{-i}.$$

We claim that for any prefix-free set $Q \subseteq V_i$, $\rho(Q) \leq \rho(U)$. Assume this is not the case for some prefix-free $Q \subseteq V_i$. Let $g : Q \to \{1,\ldots,n\}$ and $h : U \to \{1,\ldots,n\}$ such that $g(x) = h(u)$ if and only if x and u are compatible, i.e. $x \sqsubseteq u$ or $u \sqsubset x$. Since $\rho(Q) > \rho(U)$, there must be an $m \leq n$ such that, for $Q_m = g^{-1}(\{m\})$ and $U_m = h^{-1}(\{m\})$

$$\rho(Q_m) > \rho(U_m).$$

Note that at least one of Q_m, U_m must contain at most one element. If $U_m = \{u\}$ and for all $x \in Q_m$, $x \sqsupset u$, then $\rho(u) < \rho(Q_m) \leq \rho(Q_u^+)$, so one of the elements of Q_m could not have been enumerated in the construction of V_i. If, on the other hand, $Q_m = \{x\}$ and for all $u \in U_m$, $u \sqsupset x$, then this contradicts the choice of x as the longest string possible. ◻

References

1. Jin-Yi Cai and Juris Hartmanis. On Hausdorff and topological dimensions of the Kolmogorov complexity of the real line. *Journal of Computer and System Sciences*, 49(3):605–619, 1994.
2. Cristian Calude, Ludwig Staiger and Sebastiaan A. Terwijn. On partial randomness. *Annals of Pure and Applied Logic*, to appear.
3. Gregory J. Chaitin. A theory of program size formally identical to information theory. *Journal of the ACM*, 22:329–340, 1975.
4. Gregory J. Chaitin. Information-theoretic characterizations of recursive infinite strings. *Theoretical Computer Science*, 2(1):45–48, 1976.
5. Rodney G. Downey and Denis R. Hirschfeldt. *Algorithmic randomness and complexity*. Springer, to appear.
6. Ming Li and Paul Vitányi. *An Introduction to Kolmogorov Complexity and its Applications.* Second Edition. Springer, 1997.
7. Jack H. Lutz. Gales and the constructive dimension of individual sequences. In *Automata, languages and programming* (Geneva, 2000), volume 1853 of Lecture Notes in Computer Science, pages 902–913. Springer, Berlin, 2000.
8. Per Martin-Löf. The definition of random sequences. *Information and Control*, 9:602–619, 1966.
9. Jan Reimann. *Computability and fractal dimension*. Doctoral Dissertation, Universität Heidelberg, 2004.
10. Jan Reimann and Frank Stephan. Effective Hausdorff dimension. In *Logic Colloquium '01*, volume 20 of *Lecture Notes in Logic*, pages 369–385. Association of Symbolic Logic, Urbana, IL, 2005.
11. Piergiorgio Odifreddi. *Classical Recursion Theory.* North-Holland, 1989.
12. Claude Ambrose Rogers. *Hausdorff Measures.* Cambridge University Press, 1970.
13. Boris Y. Ryabko. Coding of combinatorial sources and Hausdorff dimension. *Doklady Akademii Nauk SSSR*, 277(5):1066–1070, 1984.
14. Boris Y. Ryabko. Noise-free coding of combinatorial sources, Hausdorff dimension and Kolmogorov complexity. *Problemy Peredachi Informatsii*, 22(3):16–26, 1986.
15. Claus-Peter Schnorr. *Zufälligkeit und Wahrscheinlichkeit*. Lecture Notes in Mathematics 218, Springer, 1971.
16. Robert M. Solovay. *Draft of a paper on Chaitin's work*. Manuscript, IBM Thomas J. Watson Research Center, New York, May 1975, 215 pages.
17. Robert I. Soare. *Recursively Enumerable Sets and Degrees.* Springer, 1987.
18. Ludwig Staiger. Kolmogorov complexity and Hausdorff dimension. *Information and Computation*, 103(2):159–194, 1993.
19. Kohtaro Tadaki. A generalization of Chaitin's halting probability Ω and halting self-similar sets. *Hokkaido Mathematical Journal*, 31(1):219–253, 2002.
20. Alexander K. Zvonkin and Leonid A. Levin. The complexity of finite objects and the basing of the concepts of information and randomness on the theory of algorithms. *Uspekhi Matematicheskikh Nauk*, 25(6):85–127, 1970.