The Denjoy alternative for computable functions

Laurent Bienvenu[1], Rupert Hölzl*[2], Joseph S. Miller[3], and André Nies[4]

1,2 LIAFA, CNRS & Université de Paris 7, Case 7014, 75205 Paris Cedex 13, France
`laurent.bienvenu@liafa.jussieu.fr, r@hoelzl.fr`

3 Department of Mathematics, University of Wisconsin, Madison, WI 53706-1388, USA
`jmiller@math.wisc.edu`

4 Department of Computer Science, University of Auckland, Private bag 92019, Auckland, New zealand
`andre@cs.auckland.ac.nz`

Abstract

The Denjoy-Young-Saks Theorem from classical analysis states that for an arbitrary function $f\colon \mathbb{R} \to \mathbb{R}$, the Denjoy alternative holds outside a null set. This means that for almost every real x, either the derivative of f exists at x, or the derivative fails to exist in the worst possible way: the limit superior of the slopes around x equals $+\infty$, and the limit inferior $-\infty$. Algorithmic randomness allows us to define randomness notions giving rise to different concepts of *almost everywhere*. It is then natural to wonder which of these concepts corresponds to the *almost everywhere* notion appearing in the Denjoy-Young-Saks theorem. To answer this question Demuth investigated effective versions of the theorem. In his first variation, the function f is stipulated to be computable, and in the second one the function f is only Markov computable. For this second version, Demuth introduced a strong notion of randomness (stronger for example than Martin-Löf randomness) now known as Demuth randomness, which he proved to be sufficient to satisfy the Denjoy alternative for all Markov computable functions. In this paper, we in turn investigate these two effective theorems. We first show that the set of points that fulfill the Denjoy alternative for computable functions coincides with the set of computably random reals. We then show that the set of points that fulfill the Denjoy alternative for Markov computable functions is strictly bigger than the set of Demuth random reals — showing that Demuth's sufficient condition was too strong — and moreover is incomparable with Martin-Löf randomness (meaning in particular that it does not correspond to any known set of random reals).

To prove these two theorems, we study density-type theorems, such as the Lebesgue density theorem and obtain results of independent interest. We show for example that the classical notion of Lebesgue density can be characterized in an interesting way by the only very recently defined notion of difference randomness: x being difference random is equivalent to it being Martin-Löf random and having positive density in every effectively closed class in which x is contained. This is to our knowledge the first analytical characterization of difference randomness. We also consider the concept of porous points, a special type of Lebesgue non-density points that are well-behaved in the sense that the "density holes" around the point are continuous intervals whose length follows a certain systematic rule. An essential part of our proof will be to argue that porous points of effectively closed classes can never be difference random.

1998 ACM Subject Classification F.1.1 Models of Computation

Keywords and phrases Differentiability, Denjoy alternative, density, porous, randomness

* The second author is supported by a Feodor Lynen postdoctoral research fellowship by the Alexander von Humboldt Foundation.

Conference title on which this volume is based on.
Editors: Billy Editor, Bill Editors; pp. 1–18
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Digital Object Identifier 10.4230/LIPIcs.xxx.yyy.p

1 Introduction

The aim of the theory of algorithmic randomness is to give a precise definition of what it means for a single object (usually a finite or infinite binary sequence) to be random. For infinite binary sequences (or reals, as any real can be represented by an infinite binary sequence) a satisfactory definition was given by Martin-Löf [10]. Informally an infinite sequence x is Martin-Löf random if it does not belong to any set which can computably be shown to have measure 0. Even though Martin-Löf's definition is still believed to be the best one (at least the most well-behaved), many alternative notions of randomness have appeared in the literature over the years, some weaker than Martin-Löf randomness, some stronger. We refer the reader to the two recent books [7, 11] for an extensive survey of these notions.

An interesting line of research is to study the connections between algorithmic randomness and computable analysis. The latter is concerned with effective versions of classical theorems in analysis, i.e. analytical theorems where the objects involved (functions, sets, points, etc.) are effective, i.e., computable in some sense. Now take a classical theorem of type "for any function f, for almost every x, ...". Its effective version will look like "for any effective function f, for almost every x, ...". Now, since there are only countably many effective functions (no matter what meaning is given to effective), one can reverse the quantifiers, and get a statement of type: "for almost every x, for every effective function f, ...". Therefore, a sufficiently random x will satisfy the conclusion of the theorem. For each such theorem, we can thus look at the following question: *How much* randomness is needed for x to satisfy the conclusion of the theorem? A recent example is a result proven in [1] and [9] showing that Martin-Löf randomness is precisely the level of randomness needed to satisfy the most natural effective version of Birkhoff's ergodic theorem. Another is a result of Brattka, Miller and Nies [3], which is closely connected to study conducted in this paper. They considered the effective version of the following theorem. If f is a non-decreasing function from $\mathbb{R}$ to $\mathbb{R}$, then f is differentiable almost everywhere. Following the above scheme, they studied the class of reals x such that every *computable* non-decreasing function f is differentiable at x and were able to show that this class precisely coincides with the class of computably random reals. This was surprising as computable randomness — a weakening of Martin-Löf randomness that in many ways behaves quite pathologically — had very few known characterizations other than its original definition, and in particular no known analytical characterization.

Demuth [6] studied an effective version of a related theorem, the so-called Denjoy-Young-Saks theorem, which asserts that any function $f : \mathbb{R} \to \mathbb{R}$ satisfies the Denjoy alternative at almost all points. The Denjoy alternative at a point x states that either the function is differentiable at x or the derivative fails to exists in the most dramatic way, i.e. the function f has around x arbitrarily large positive slopes and negative slopes. Demuth mainly studied the Denjoy alternative for Markov computable functions (which we will define in the moment) and studied the set DA of points x such that any Markov computable function satisfies the Denjoy alternative at x. Demuth introduced a randomness notion, now called Demuth randomness, which he proved to be sufficient to be in DA. The main result of this paper is that Demuth randomness is in fact too strong a condition, and that the class DA is strictly larger that the class of Demuth random reals.

We show that difference randomness (a notion of randomness strictly stronger than Martin-Löf randomness and strictly weaker than Demuth randomness) constitutes a sufficient condition for the Denjoy alternative to hold for Markov computable functions.

▶ **Theorem 1.** *Every difference random real belongs to* DA.

We then show that this result cannot be strengthened to Martin-Löf randomness: in fact, Martin-Löf randomness is neither sufficient nor necessary to ensure the Denjoy alternative for Markov computable functions.

▶ **Theorem 2.** *The set* DA *of reals which satisfy the Denjoy alternative for all Markov computable functions is incomparable under inclusion with the set of Martin-Löf random reals.*

These results will be proven in Section 3.

1.1 Preliminaries

We provide notation, recall the definitions of computable and Markov computable functions on the real numbers, and recall the definitions of Martin-Löf randomness, difference randomness, and computable randomness.

Basic notation. The set of finite binary sequences (we also say strings) is denoted by $2^{<\omega}$, and the set of infinite binary sequences, called Cantor space, is denoted by 2^ω. For a string σ, $|\sigma|$ is the length of σ. If σ is a string, and x is either a string or an infinite binary sequence, we say that σ is a prefix of x, which we write $\sigma \preceq x$, if the first $|\sigma|$ bits of x are exactly the string σ. Given an binary sequence, infinite or finite with length at least n, $x{\upharpoonright}n$ denotes the string made of the first n bits of x.

The Cantor space is classically endowed with the product topology. A basis of this topology is the set of cylinders: given a string $\sigma \in 2^{<\omega}$, the cylinder $[\![\sigma]\!]$ is the set of elements of 2^ω having σ as a prefix. If A is a set of strings, $[\![A]\!]$ is the union of the cylinders $[\![\sigma]\!]$ with $\sigma \in A$. The Lebesgue measure λ (or uniform measure) on the Cantor space is the probability measure assigning to each bit the value 0 with probability 1/2 and the value 1 with probability 1/2, independently of all other bits. Equivalently it is the measure λ such that $\lambda([\![\sigma]\!]) = 2^{-|\sigma|}$ for all σ. We abbreviate $\lambda([\![\sigma]\!])$ by $\lambda(\sigma)$. Given two subsets $\mathcal{X}$ and $\mathcal{Y}$, the second one being of positive measure, the conditional measure $\lambda(\mathcal{X}|\mathcal{Y})$ of $\mathcal{X}$ knowing $\mathcal{Y}$ is the quantity $\lambda(\mathcal{X} \cap \mathcal{Y})/\lambda(\mathcal{Y})$. As before, if $\mathcal{X}$ or $\mathcal{Y}$ is a cylinder $[\![\sigma]\!]$, we will simply write it as σ.

Computable real-valued functions. Most of the paper will focus on functions from $[0,1]$ to $\mathbb{R}$. The set $[0,1]$ is typically identified with 2^ω, where a real $x \in [0,1]$ is identified with its binary expansion. This extension is unique, except for dyadic rationals (of the form $a2^{-b}$ with a, b positive integers) which have two. A cylinder $[\![\sigma]\!]$ will be commonly identified with the open interval $(0.\sigma, 0.\sigma + 2^{-n})$, where $0.\sigma$ is the dyadic rational whose binary extension is $0.\sigma000\ldots$.

We say that a function $f : [0,1] \to \mathbb{R}$ is computable (over the reals) if it can be effectively approximated with arbitrary precision. More precisely, f is computable (over the reals) if there exists a computable function $\widehat{f} : \mathbb{Q} \times \mathbb{N} \to \mathbb{Q}$ and a computable $\psi : \mathbb{N} \to \mathbb{N}$ such that for all $x \in [0,1]$, $|x - q| < 2^{-\psi(n)} \Rightarrow |f(x) - \widehat{f}(q,n)| < 2^{-n}$. Note that a computable function over the reals is by this definition necessarily continuous. A real x is computable if the constant function x is computable. Equivalently, a real is computable if its binary expansion, seen as a function from $\mathbb{N}$ to $\{0,1\}$ is computable.

We denote the set of computable reals by $\mathbb{R}_c$. The image of a computable real by a computable function is itself a computable real. Since the computable reals form a dense subset of the reals, a computable function is uniquely determined by its restriction $\mathbb{R}_c \to \mathbb{R}_c$.

The class of Markov computable functions is a larger class of functions $\mathbb{R}_c \to \mathbb{R}_c$. As we just said, a real x is computable if there is a computable function β which computes its binary expansion. Any index i of β in an uniform enumeration $(\phi_i)_i$ of partial computable functions is called a *name* for x. A function $f : \mathbb{R}_c \to \mathbb{R}_c$ is said to be Markov computable if from a name of $x \in \mathbb{R}_c$, one can effectively compute a name for $f(x)$. More precisely, f is Markov computable if there exists a partial computable function $\check{f} : \mathbb{N} \to \mathbb{N}$ such that for all $x \in \mathbb{R}_c$, if i is a name for x, then $\check{f}(i)$ is defined and is a name for $f(x)$. Given a Markov computable function f, $x \in [0,1]$ and $s \in \mathbb{N}$, we write sometimes use the notation $f(x)_s$ to denote the approximation of $f(x)$ at stage s. Unless specified otherwise, a Markov computable function is always assumed to be total on $[0,1] \cap \mathbb{R}_c$. An important theorem of Tseitin [12] states that a total Markov computable function is always continuous on its domain.

We define the following analytical notations: for a function f, the *slope* at a pair a, b of distinct reals in its domain is

$$S_f(a,b) = \frac{f(a) - f(b)}{a - b}.$$

Recall the following definitions for the case that z is in the domain of f.

$$\overline{D}f(z) = \limsup_{h \to 0} S_f(z, z+h) \quad \text{and} \quad \underline{D}f(z) = \liminf_{h \to 0} S_f(z, z+h)$$

The derivative $f'(z)$ exists if and only if these values are equal and finite.

In this article we will work with functions that are not necessarily defined on all reals, e.g. Markov computable functions. When working with these functions $\overline{D}f(z)$ and $\underline{D}f(z)$ are not defined for all z. Nonetheless, in case the set $dom(f)$ is a dense subset of $[0,1]$, one can consider the lower and upper *pseudo*-derivatives defined by:

$$\underset{\sim}{D}f(z) = \liminf_{h \to 0^+} \{S_g(a,b) \colon a, b \in \mathrm{dom}(f) \ \wedge \ a \le x \le b \ \wedge \ 0 < b - a \le h\}.$$

$$\widetilde{D}f(z) = \limsup_{h \to 0^+} \{S_g(a,b) \colon a, b \in \mathrm{dom}(f) \ \wedge \ a \le x \le b \ \wedge \ 0 < b - a \le h\}.$$

It is not true in general that if a function f is defined at a point z, then $\underset{\sim}{D}f(z) = \underline{D}f(z)$ and $\widetilde{D}f(z) = \overline{D}f(z)$. This is true however when the function f is continuous on its (dense) domain, which is the case for computable and Markov computable functions. Note also that in this special case, one can replace, in the definition of $\underset{\sim}{D}f$ and $\widetilde{D}f$, dom(h) by any dense subset of dom(h) (for Markov computable functions, one could for example take $\mathbb{Q}$ instead).

We also later need the following technical lemma, which we prove in the appendix.

▶ **Lemma 3.** *Let $h\colon \ \subseteq [0,1] \to \mathbb{R}_0^+$ be a computable function that is defined and non-decreasing on an effectively closed class $\mathcal{C}$. Then $h\restriction_{\mathcal{C}}$ can be extended to a function $g\colon [0,1] \to \mathbb{R}_0^+$ that is computable and non-decreasing on $[0,1]$.*

Randomness notions. As we have seen above, an open set $\mathcal{U} \subseteq 2^\omega$ is a union of cylinders. If it is a union of a *computably enumerable* (c.e.) family of cylinders, it is said to be effectively open (or c.e. open). A set is called effectively closed set if its complement is effectively open. If $(\mathcal{U}_n)$ is a sequence of open sets, it is said a to be a uniformly c.e. sequence of open sets if there is a sequence (W_n) of uniformly c.e. sets of strings such that each $\mathcal{U}_n$ is the union of the cylinders generated by the strings in W_n.

A *Martin-Löf test* is a uniformly c.e. sequence $(\mathcal{U}_n)_n$ of open sets such that for all n, $\lambda(\mathcal{U}_n) < 2^{-n}$. A *difference test* is a pair $((\mathcal{U}_n)_n, \mathcal{C})$ of a uniformly c.e. sequence $(\mathcal{U}_n)_n$ of open classes and a single effectively closed class $\mathcal{C}$ such that for all n, $\lambda(\mathcal{U}_n \cap \mathcal{C}) < 2^{-n}$. A strong test is a sequence of uniformly c.e. sequence $(\mathcal{U}_n)_n$ of open sets with the weaker condition that $\lim_n \lambda(\mathcal{U}_n) = 0$.

▶ **Definition 4.** A sequence $x \in 2^\omega$ is called *Martin-Löf random* if there is no Martin-Löf test *covering* it, i.e., for any Martin-Löf test $(\mathcal{U}_n)_n$ we have $x \notin \bigcap_n \mathcal{U}_n$. A sequence $x \in 2^\omega$ is called *weak-2-random* if there is no strong test covering it, i.e., for any strong test $(\mathcal{U}_n)_n$ we have $x \notin \bigcap_n \mathcal{U}_n$. A sequence $x \in 2^\omega$ is called *difference random* if there is no difference test *covering* it, i.e., if for any difference test $((\mathcal{U}_n)_n, \mathcal{C})$ we have $x \notin \bigcap_n (\mathcal{U}_n \cap \mathcal{C})$.

The notion of difference randomness was introduced by Franklin and Ng [8]. They proved that the set of difference random reals in fact coincides with the set of Martin-Löf random reals that are Turing incomplete.

▶ **Proposition 5.** *For both Martin-Löf randomness and difference randomness, it is equivalent (see for example [7]) to require "almost avoidance": a sequence $x \in 2^\omega$ is Martin-Löf random (resp. difference random) if and only if for every Martin-Löf test $(\mathcal{U}_n)$ (resp. difference test $((\mathcal{U}_n), \mathcal{C})$), x only belongs to finitely many $\mathcal{U}_n$ (resp. finitely many $\mathcal{U}_n \cap \mathcal{C}$).*

Note that this type of "almost avoidance" variation of definitions is not admissible for weak 2-randomness.

Another strengthening of Martin-Löf randomness is Demuth randomness. A Demuth test is a sequence $(\mathcal{U}_n)$ of effectively open sets, which is not necessarily uniformly c.e. but instead enjoys the following weak form of uniformity: there exists an ω-c.e. function $f : \mathbb{N} \to \mathbb{N}$ which for each n gives a c.e. index for a set of strings generating $\mathcal{U}_n$.

▶ **Definition 6.** A sequence $x \in 2^\omega$ is said to be Demuth random if for every Demuth test $(\mathcal{U}_n)$, x belongs to only finitely many $\mathcal{U}_n$.

The last notion of randomness we will discuss in the paper is computable randomness. Its definition involves the notion of martingale.

▶ **Definition 7.** A martingale is a function $d : 2^{<\omega} \to [0, \infty)$ such that for all $\sigma \in 2^{<\omega}$

$$d(\sigma) = \frac{d(\sigma 0) + d(\sigma 1)}{2}.$$

Intuitively, a martingale represents a betting strategy where a player successively bets money on the values of the bits of an infinite binary sequence (doubling its stake when guess is correct); $d(\sigma)$ then represents the capital of the player after betting on initial segment σ. With this intuition, a martingale succeeds against a sequence x if $\limsup_n d(x{\upharpoonright}n) = +\infty$. A computably random sequence is a sequence against which no computable betting strategy succeeds. In other words:

▶ **Definition 8.** A sequence $x \in 2^\omega$ is computably random if and only if for every computable martingale d, $\limsup_n d(x{\upharpoonright}n) < +\infty$

We denote by MLR, W2R, DiffR, DemR, CR the classes of Martin-Löf random, weak-2-random, difference random, Demuth random and computably random sequences respectively.

Given a sequence $x \in 2^\omega$, the following implications

$$\begin{array}{ccccc} x \in \mathsf{W2R} & & & & \\ & \searrow & & & \\ & & x \in \mathsf{DiffR} \longrightarrow & x \in \mathsf{MLR} \longrightarrow & x \in \mathsf{CR} \\ & \nearrow & & & \\ x \in \mathsf{DemR} & & & & \end{array}$$

hold and no other implication holds in general (other than those which can be derived by transitivity from the above diagram). See for example [11] for a detailed exposition.

2 Density and porosity

In this section, we initiate the study of effective aspects of Lebesgue density, which will be crucial in the proofs of Theorems 1 and 2. In this section, we mostly focus on what is needed for the proofs of these theorems. In Section 4 we will provide further results on density.

Let us first recall the concept of Lebesgue density:

▶ **Definition 9.** We define the (Lebesgue) density ρ of a set $\mathcal{C} \subseteq \mathbb{R}$ at a point x to be the quantity

$$\rho(x|\mathcal{C}) := \liminf_{\delta \to 0^+} \frac{\lambda([x-\delta, x+\delta] \cap \mathcal{C})}{\lambda([x-\delta, x+\delta])},$$

where λ is the Lebesgue measure.

Intuitively, this measures what fraction of the space is filled by $\mathcal{C}$ around x if we "zoom in" arbitrarily close. Note that the density of a set at a point is between 0 and 1.

Again, in the rest of the paper, we will freely identify 2^ω and $[0,1]$ and will therefore be able to talk about density of a set $\mathcal{C} \subseteq 2^\omega$ at a point.

▶ **Theorem 10** (Lebesgue density theorem)**.** *Let $\mathcal{C} \subseteq \mathbb{R}$ be a measurable set. Then $\rho(x|\mathcal{C}) = 1$ for all points $x \in \mathcal{C}$ outside a set of measure 0.*

The concept of porosity of a set at a point forms a cornerstone of the proofs of Theorems 1 and 2. The following definition is due to E.P. Dolzhenko, 1967 (see for instance [2, 5.8.124], but note the typo in the definition there).

▶ **Definition 11.** We say that $\mathcal{C}$ is *porous at x* via $\varepsilon > 0$ if for each $\alpha > 0$ there exists β with $0 < \beta \leq \alpha$ such that $(x-\beta, x+\beta)$ contains an open interval of length $\varepsilon\beta$ that is disjoint from $\mathcal{C}$. We say that $\mathcal{C}$ is porous at x if it is porous at x via some ε. We call *non-porosity point* a real x such that every effectively closed class to which it belongs is non-porous at x.

Clearly porosity of $\mathcal{C}$ at x implies $\rho(x|\mathcal{C}) < 1$. Therefore, given an effectively closed class $\mathcal{C}$, for almost every point x of this class, $\mathcal{C}$ is not porous at x. Since there are only countably many effectively closed sets, it follows that the set of non-porosity points has measure 1. Our next proposition makes this more precise.

▶ **Lemma 12.** *Let $\mathcal{C} \subseteq [0,1]$ be an effectively closed class. If $z \in \mathcal{C}$ is difference random, then $\mathcal{C}$ is not porous at z.*

Proof. In this proof, we say that a string σ meets $\mathcal{C}$ if $[\![\sigma]\!] \cap \mathcal{C} \neq \emptyset$.

Fix $c \in \mathbb{N}$. We build a difference test covering the points x at which $\mathcal{C}$ is porous via 2^{-c+2}. For each string σ consider the set of minimal "porous" extensions at stage t,

$$N_t(\sigma) = \left\{ \rho \succeq \sigma \,\middle|\, \exists \tau \left[\begin{array}{c} |\tau| = |\rho| \wedge |0.\tau - 0.\rho| \leq 2^{-|\rho|+c} \wedge \\ [\![\tau]\!] \cap \mathcal{C}_t = \emptyset \wedge \rho \text{ is minimal with this property} \end{array} \right] \right\}.$$

Note that by the formal details of this definition even the "holes" τ are ρ's, and therefore contained in the sets $N_t(\sigma)$. This will be essential for the proof of the first of the following two claims. In contrast to this, note that if ρ meets $\mathcal{C}$ then $\rho \neq \tau$, which implies

$$\sum_{\substack{\rho \in N_t(\sigma) \\ \rho \text{ meets } \mathcal{C}}} 2^{-|\rho|} \leq (1 - 2^{-c}) 2^{-|\sigma|}. \tag{1}$$

At each stage t of the construction we define recursively a sequence of anti-chains as follows.

$$B_{0,t} = \{\emptyset\}, \text{ and for } n > 0\colon B_{n,t} = \bigcup\{N_t(\sigma)\colon\ \sigma \in B_{n-1,t}\}$$

Claim. If a string ρ is in $B_{n,t}$ then it has a prefix ρ' in $B_{n,t+1}$.
This is clear for $n = 0$. Suppose inductively that it holds for $n-1$. Suppose further that ρ is in $B_{n,t}$ via a string $\sigma \in B_{n-1,t}$. By the inductive hypothesis there is $\sigma' \in B_{n-1,t+1}$ such that $\sigma' \preceq \sigma$. Since $\rho \in N_t(\sigma)$, ρ is a viable extension of σ' at stage $t+1$ in the definition of $N_{t+1}(\sigma')$, except maybe for the minimality. Thus there is $\rho' \preceq \rho$ in $N_{t+1}(\sigma')$. This establishes the claim. $\Diamond$

Claim. For each n, t, we have $\sum\{2^{-|\rho|}\colon\ \rho \in B_{n,t} \wedge \rho \text{ meets } \mathcal{C}\} \le (1-2^{-c})^n$.
This is again clear for $n = 0$. Suppose inductively it holds for $n-1$. Then, by (1),

$$\begin{aligned}\sum_{\substack{\rho\in B_{n,t}\\ \rho \text{ meets } \mathcal{C}}} 2^{-|\rho|} \quad &= \quad \sum_{\substack{\sigma\in B_{n-1,t}\\ \sigma \text{ meets } \mathcal{C}}} \ \sum_{\substack{\rho\in N_t(\sigma)\\ \rho \text{ meets } \mathcal{C}}} 2^{-|\rho|} \\ &\le \quad \sum_{\substack{\sigma\in B_{n-1,t}\\ \sigma \text{ meets } \mathcal{C}}} 2^{-|\sigma|}(1-2^{-c}) \quad \le \quad (1-2^{-c})^n.\end{aligned}$$

This establishes the claim. $\Diamond$

Now let $U_n = [\![\bigcup_t B_{n,t}]\!]$. Clearly the sequence $(U_n)_{n\in\mathbb{N}}$ is uniformly effectively open. By the first claim we have $U_n = \bigcup_t [\![B_{n,t}]\!]$, so the second claim implies that $\lambda(\mathcal{C} \cap U_n) \le (1-2^{-c})^n$.

Every interval $(a, b) \subseteq [0, 1]$ contains an interval of the form $[\![\rho]\!]$ for a dyadic string ρ such that the length of $[\![\rho]\!]$ is no less than $(b-a)/4$. Therefore, if $\mathcal{C}$ is porous at x via 2^{-c+2} then $x \in \bigcap_n U_n$ for each n. Take a computable subsequence $(U_{g(n)})_{n\in\mathbb{N}}$ such that $\lambda(\mathcal{C} \cap U_{g(n)}) \le 2^{-n}$ to obtain a difference test that x fails. ◀

3 Effective forms of the Denjoy-Young-Saks Theorem

We begin with the formal definition of the Denjoy alternative.

▶ **Definition 13.** Let $f :\subseteq [0, 1] \to \mathbb{R}$ be a partial function whose domain is dense. We say that f satisfies the Denjoy alternative at x if

- either the pseudo-derivative of f at x exists (meaning that $\widetilde{D}f(x) = \underset{\sim}{D}f(x)$)
- or $\widetilde{D}f(x) = +\infty$ and $\underset{\sim}{D}f(x) = -\infty$

Intuitively this means that there are two ways for the alternative to hold: either the function behaves well on x by having a derivative at this place, or, if it behaves badly, it does so in the worst possible way, that being the fact that the limit superior and the limit inferior diverge as much as possible. The Denjoy-Young-Saks theorem (see, e.g., Bruckner [4]) states that the Denjoy alternative holds at almost all points for *any* function f.

3.1 Computable randomness means that all computable function satisfy the Denjoy alternative

▶ **Definition 14** (Demuth [5]). A real $z \in [0, 1]$ is called Denjoy random (or a Denjoy set) if for no Markov computable function g we have $\underset{\sim}{D}g(z) = +\infty$.

In a preprint by Demuth [5, p. 6] it is shown that if $z \in [0,1]$ is Denjoy random, then for every computable $f\colon [0,1] \to \mathbb{R}$ the Denjoy alternative holds at z. By combining this result with the results in [3] we can achieve the following result that provides a pleasing characterization of computable randomness through differentiability of computable functions.

▶ **Theorem 15** (Demuth, Miller, Nies, Kučera). *The following are equivalent for a real $z \in [0,1]$.*
1. *z is Denjoy random.*
2. *z is computably random.*
3. *For every computable $f\colon [0,1] \to \mathbb{R}$ the Denjoy alternative holds at z.*
Note that all we needed for the last implication was that $g(q)$ is a computable real uniformly in a rational $q \in [0,1] \cap \mathbb{Q}$. Thus, in Definition 14 we can replace Markov computability of g by this weaker hypothesis.

3.2 Difference randomness implies that all Markov computable function satisfy the Denjoy alternative

We now turn to the proof of Theorem 1. Using the results of the previous section, it will be enough to prove the following.

▶ **Proposition 16.** *Let $x \in 2^\omega$ be a computably random real that is also a non-porosity point. Then $x \in \mathsf{DA}$, i.e., all Markov computable functions satisfy the Denjoy alternative at x.*

To get Theorem 1 from this proposition, remember that a difference random point is always computably random, and, by Lemma 12 a difference random real is also always a non-porosity point.

Proof. We first prove a lemma, which takes advantage of the special way in which a set is arranged around its non-porosity points.

▶ **Lemma 17.** *Suppose that $f\colon \subseteq [0,1] \to \mathbb{R}$ is Markov computable. Let $\mathcal{C} \subseteq [0,1]$ be an effectively closed class such that there is an n with $\underset{\sim}{D}f(z) > -n$ for all $z \in \mathcal{C}$. Let the computably random real $x \in \mathcal{C}$ be a non-porosity point of $\mathcal{C}$. Then f is differentiable at x.*

Proof. We effectivize an argument of Bogachev [2, p. 371]. Replacing f by $f(x) + nx$, we may assume that for $x \in \mathcal{C}$, we have

$$\forall a, b\ [r \le a \le x \le b \le s \to S_f(a,b)_0 > 1].$$

Let $f_*(x) = \sup_{a \le x} f(a)$. Then f_* is nondecreasing on $\mathcal{C}$.

Claim. The function $f_* \restriction_E$ is computable.
To see this, recall that p, q range over $[0,1] \cap \mathbb{Q}$, and let $f^*(x) = \inf_{q \ge x} f(q)$. If $x \in \mathcal{C}$ and $f_*(x) < f^*(x)$ then x is computable: fix a rational d in between these two values. Then $p < x \leftrightarrow f(p) < d$, and $q > x \leftrightarrow f(q) > d$. Hence x is both left-c.e. and right-c.e., and therefore computable. Now a Markov computable function is continuous at every computable x. Thus $f_*(x) = f^*(x)$ for each x in $\mathcal{C}$.

To compute $f(x)$ for $x \in \mathcal{C}$ up to precision 2^{-n}, we can now simply search for rationals $p < x < q$ such that $0 < f(q)_{n+2} - f(p)_{n+2} < 2^{-n-1}$, and output $f(p)_{n+2}$. If during this search we detect that $x \notin \mathcal{C}$, we stop. This shows the claim. ♢

By Lemma 3 there is a computable nondecreasing function g defined on $[0,1]$ that extends f_*. Then by a classic theorem of Lebesgue, $g'(x)$ exists for a.e. $x \in [0,1]$. A result by Brattka, Miller and Nies [3, Thm. 4.1] states that in fact computable randomness of x is enough to guarantee the existence of $g'(x)$.

It therefore suffices to show that for each $x \in \mathcal{C}$ such that $g'(x)$ is defined and $\mathcal{C}$ is not porous at x, we have $\widetilde{D}f(x) \leq g'(x) \leq \underset{\sim}{D}f(x)$. Since $\underset{\sim}{D}f(x) \leq \widetilde{D}f(x)$, this would establish the theorem.

To see this, we show $\widetilde{D}f(x) \leq g'(x)$, the other inequality being symmetric. Fix $\varepsilon > 0$. Choose $\alpha > 0$ such that

$$\forall u, v \in \mathcal{C}\,[(u \leq x \leq v \wedge 0 < v - u \leq \alpha) \to S_{f_*}(u, v) \leq g'(x)(1 + \varepsilon)] \tag{2}$$

furthermore, since $\mathcal{C}$ is not porous at x, for each $\beta \leq \alpha$, the interval $(x - \beta, x + \beta)$ contains no open subinterval of length $\varepsilon\beta$ that is disjoint from $\mathcal{C}$. Now suppose that $a, b \in [0, 1] \cap \mathbb{Q}$, $a < x < b$ and $\beta = 2(b - a) \leq \alpha$. There are $u, v \in \mathcal{C}$ such that $0 \leq a - u \leq \varepsilon\beta$ and $0 \leq v - b \leq \varepsilon\beta$. Since $u, v \in \mathcal{C}$ we have $f_*(u) \leq f(a)$ and $f(b) \leq f_*(v)$. Therefore $v - u \leq b - a + 2\varepsilon\beta = (b - a)(1 + 4\varepsilon)$. It follows that

$$S_f(a, b) \leq \frac{f_*(v) - f_*(u)}{b - a} \leq S_{f_*}(u, v)(1 + 4\varepsilon) \leq g'(x)(1 + 4\varepsilon)(1 + \varepsilon).$$ ◀

We are now ready to prove Proposition 16. Suppose x is computably random and is a non-porosity point. Let f be a Markov computable function. Suppose that f does not satisfy the Denjoy alternative at x by strong failure of the existence of the pseudo-derivative at x. We therefore are $\underset{\sim}{D}f(x) > -\infty$ or $\widetilde{D}f(x) < +\infty$. Suppose the first one holds (the proof for the other case is similar), and take an n such that $\underset{\sim}{D}f(x) > -n$. By definition of $\underset{\sim}{D}$, this means that for some fixed positive rational ε and some fixed t, x belongs to the effectively closed class:

$$\mathcal{C} = \{x \in [0, 1] \mid \forall q_1, q_2 \in \mathbb{Q}\ s.t.\ x \in [q_1, q_2] \wedge |q_2 - q_1| < \varepsilon,\ S_f(q_1, q_2)_t > -n\}$$

We can therefore apply Proposition 16 to this class $\mathcal{C}$ (every point $z \in \mathcal{C}$ is such that $\underset{\sim}{D}f(z) > -n$, x belongs to $\mathcal{C}$, x is computably random and is a non-porosity point). Therefore, f is differentiable at x, and thus the Denjoy alternative holds indeed. ◀

3.3 The class DA is incomparable with the Martin-Löf random reals

▶ **Theorem 18.** *There exists a real x which is not Martin-Löf random but nonetheless satisfies the Denjoy alternative for all constructive functions.*

Proof sketch. The Denjoy alternative at x can be met in two ways. We will say "the DA for f is fulfilled by existence" if the (pseudo-)derivative of f at x exists and say that "the DA for f is fulfilled by failure" in the other case. To prove the statement we construct a set x by forcing that is CR, not MLR and for every Markov computable function either fulfills the DA by failure for this function or is a density point (and therefore certainly not a porosity point) of a certain effectively closed class L such that L and f fulfill the requirements of Lemma 17. The argument to prove the statement then goes like this: if we fulfill the DA by failure we are done. Otherwise we will be in the effectively closed class L. Since $x \in$ CR we can invoke Lemma 17 to show that f is differentiable and therefore fulfills the DA by existence.

Assume we have constructed the initial segment σ of x so far, and are given a computable martingale M. The most interesting part of the argument is how to ensure that we are density points of certain effectively closed classes of the form $L := \{x \succeq \sigma \mid M(x{\upharpoonright}n) < \varepsilon \text{ for all } n > |\sigma|\}$. To do this we need to make sure that the density of x in L will be 1 in the limit. At every stage of the construction we will make sure that the density of x is at least

$1-q$ for some q by choosing the right extension σ' of σ. When we later extend σ' further we will make q smaller and smaller while forever staying inside L. This way in the limit we reach density 1 in L.

To achieve density $1-q$ as required we look at the quantity $m := \inf_{\tau \succeq \sigma} d(\tau)$. We interpret m as an amount of capital that the martingale M has put on a savings account and is not touching anymore. Of course this implies that M also has less capital available for betting and can therefore reach capital ε only on a smaller fraction of the extensions of σ. By applying the Ville-Kolmogorov inequality for martingales to $M-m$ it is clear that M can reach capital ε only on a set of extensions of σ of relative measure $1-\frac{M(\sigma)-m}{\varepsilon-m}$. Or, in other words, σ has density $1-(M(\sigma)-m)/(\varepsilon-m)$ in L. By replacing σ with a long enough extension we can make sure that $d(\sigma)$ is arbitrarily close to m and thereby raise the density to the desired level $1-q$. ◀

▶ **Theorem 19.** *There exists a Markov computable function f for which the Denjoy alternative does not hold at Chaitin's Ω. Moreover, f can be taken to be uniformly continuous, i.e., it can be built in such a way that it has a (unique) continuous extension to* $[0,1]$.

Proof. Let $(\mathcal{U}_n)$ be a universal Martin-Löf test, i.e. a test such that all reals that are not in MLR are covered by it (the existence of such a Martin-Löf test is well-known). No computable real can be Martin-Löf random, every $x \in \mathbb{R}_c$ belongs to $\mathcal{U}_1$. Let Ω be the leftmost point of the complement of $\mathcal{U}_1$. Since $\mathcal{U}_1$ is a c.e. open set Ω can be approximated from below by the computable sequence of rationals $(\Omega_s)_s$, where Ω_s is the leftmost point of $\mathcal{U}_1[s]$ (the approximation of $\mathcal{U}_1$ after s stages). Our function f is defined as the restriction to $\mathbb{R}_c$ of the following function F. Outside $\mathcal{U}_1$, F is equal to 0. On $\mathcal{U}_1$, it is constructed sequentially as follows. At stage $s+1$, enumerate a new interval (σ) inside $\mathcal{U}_1$ (we can assume that it is disjoint from $\mathcal{U}_1[s]$). There are two cases.

1. Either adding this interval does not change the value of Ω (i.e., $\Omega_{s+1}=\Omega_s$). In that case, define the function F to be equal to zero on the whole interval (σ).
2. Or, this interval does change the value of Ω: $\Omega_{s+1} > \Omega_s$. In this case, define F on (a,b) to be the triangular function taking value 0 on the endpoints of (σ) and reaching the value v at the middle point, where v is defined as follows. Let t be the last stage at which the *previous* increase of Ω occurred (i.e., t is maximal such that $t<s$ and $\Omega_{t+1} > \Omega_t$). Let n be the smallest integer such that the real interval $[\Omega_t, \Omega_{t+1}]$ contains a multiple of 2^{-n}. For that n, set $v = 2^{-n/2}$.

First, we see that the restriction f of F to $\mathbb{R}_c$ is Markov computable: given a code i, we try to compute the real x coded by i (remember that such an x might not exist) until we find a sufficiently good estimate $a < x < b$, a, b dyadic, such that the interval $[a,b]$ is contained in one of the intervals (σ) appearing in the enumeration of $\mathcal{U}_1$. It is then easy to compute F at x as one can decide which of the above cases hold, and both the zero function and the triangular function are computable on $\mathbb{R}_c$ (for the triangular function, remark that the value n of the construction can be found computably).

We claim that the function f does not satisfy the Denjoy alternative at Ω. More precisely, we have $\widetilde{D}f(\Omega)=0$ and $\underset{\sim}{D}f(\Omega)=-\infty$. Notice that f is equal to 0 on $(\Omega,1]\cap\mathbb{R}_c$ and non-negative on $[0,\Omega)\cap\mathbb{R}_c$, taking the value 0 at points arbitrarily close to Ω (at least the endpoints of intervals (σ) enumerated on the left of Ω), therefore $\widetilde{D}f(\Omega)=0$ is clear. To see that $\underset{\sim}{D}f(\Omega)=-\infty$, consider for all k the dyadic real a_k which is a multiple of 2^{-k}, is smaller than Ω and such that $\Omega - a_k < 2^{-k}$. Since $a_k < \Omega$, there exists a stage t such that $a_n \in [\Omega_t, \Omega_{t+1}]$. Let $s > t$ be the next stage at which Ω increases. By definition, F is then defined to be a triangular function on $[\Omega_s, \Omega_{s+1}]$ of height $2^{-n/2}$, where n has to be greater than k as $[\Omega_s, \Omega_{s+1}]$ does not contain any multiple of 2^{-k} (a_k is the largest such real that is smaller than Ω, and we have $a_k < \Omega$. Thus, calling x_k the middle point of $[\Omega_s, \Omega_{s+1}]$, we have

$$S_f(x_k, \Omega) = \frac{f(x_k) - f(\Omega)}{x - \Omega} \leq \frac{2^{-k} - 0}{-2^{-k/2}} \leq -2^{-k/2}$$

and this happens for all k, hence $\underset{\sim}{D}f(\Omega) = -\infty$.

It remains to show that the function F is continuous on $[0, 1]$. But this is almost immediate as one can write $F = \sum_n h_n$ where h_n is the function equal to 0 except on the intervals on which F is a triangular function of height $2^{-n/2}$, and on that interval $h_n = F$. It is obvious that the h_n are continuous and of magnitude $2^{-n/2}$. Therefore $\sum_n ||h_n|| < \infty$, so by the Weierstrass M-test we can conclude that the convergence is uniform hence the function $\sum_n h_n$ is continuous. ◀

4 Positive density as a randomness notion

We return to the notion of positive density and give an interesting characterization of incomplete Martin-Löf random sets.

▶ **Theorem 20.** *The following statements are equivalent for a real x.*

1. *x is difference random.*
2. *x is Martin-Löf random and x has incomplete Turing degree.*
3. *x is Martin-Löf random and a point of positive lower density in every effectively closed class $\mathcal{C}$ with $x \in \mathcal{C}$, that is, $\rho(x|\mathcal{C}) \neq 0$.*

Proof sketch. The equivalence between (1) and (2) has been shown by Franklin and Ng [8].

(1) ⇒ (3): Proof by contraposition. Assume that $x \in \mathsf{MLR}$ and that for all ε there is an n such that $\lambda(\mathcal{C} \mid x \restriction n) < \varepsilon$. Then for all k let

$$\mathcal{U}_k := \{z \mid \exists n \ \lambda(\mathcal{C} \mid z \restriction n) < 2^{-k}\}.$$

Since $\mathcal{C}$ is effectively closed, these classes are uniformly effectively open. Let G_k be a minimal prefix-free set of strings generating $\mathcal{U}_k$. Then the following computation shows that $(\mathcal{U}_k)_k$ is a difference test.

$$\begin{array}{rcll} \lambda(\mathcal{U}_k \cap \mathcal{C}) & = & \sum_{\sigma \in G_k} \lambda(\mathcal{C} \cap [\sigma]) & \\ & = & \sum_{\sigma \in G_k} 2^{-|\sigma|} \cdot \lambda(\mathcal{C} \mid [\sigma]) & \\ & \leq & \sum_{\sigma \in G_k} 2^{-|\sigma|} \cdot 2^{-k} & \leq \ 2^{-k}. \end{array}$$

Since $x \in \mathcal{C}$, by the original assumption, this difference test clearly covers x.

(3) ⇒ (2): Suppose now that x is Martin-Löf random and Turing complete. We are going to show that x has lower density 0 inside some effectively closed class $\mathcal{C}$. We show that, given a rational ε, we can effectively construct an effectively closed class $\mathcal{C}_\varepsilon$ such that $x \in \mathcal{C}_\varepsilon$ and $\lambda(\mathcal{C}_\varepsilon \mid x \restriction n) < \varepsilon$ for some n. It will then suffice to let $\mathcal{C} := \bigcap_\varepsilon \mathcal{C}_\varepsilon$ for an effective list of ε's that converge to 0.

Fix $\varepsilon > 0$. In this construction, we will build an auxiliary c.e. set W. By the recursion theorem, since x is complete, we can assume to know in advance a Turing reduction Γ such that $\Gamma^x = W$.

In order to lower the density of $\mathcal{C}_\varepsilon$ around x we need to remove many reals from $\mathcal{C}_\varepsilon$. Since we don't know x this comes at the risk of inadvertently removing x as well. The approach of the proof is then to make use of the fact that we control W. We keep observing the results of the reduction Γ relative to all possible oracles and wait until we see a certain type of behavior (reduction outputs 0) on all oracles except fraction ε. As soon as this happens we change W in such a way that it does exactly *not* show this behavior. Since x computes W it certainly cannot be among the $1-\varepsilon$ fraction of oracles showing the special behavior, so we can safely remove them from $\mathcal{C}_\varepsilon$.

Of course it must be avoided that we wait forever, since in that case the measure of $\mathcal{C}_\varepsilon$ would forever remain equal to 1. It will therefore be necessary to argue why we can be sure that we will eventually observe the special behavior. To see this, we will argue that if we never observe that behavior, x is in a descending chain of sets U_k such that U_k always has measure $1-\varepsilon$ relative to U_{k-1}, and that this chain actually is a Martin-Löf test covering x. This of course contradicts $x \in \mathsf{MLR}$.

The formal details of the construction can be found in the appendix.

Together with Lemma 12 we get the following corollary. To the best of our knowledge there exists no direct proof of this fact.

▶ **Corollary 21.** *For any $x \in \mathsf{MLR}$ the following implication holds: If for all effectively closed classes $\mathcal{C}$ with $x \in \mathcal{C}$ it holds that $\rho(x|\mathcal{C}) > 0$ then for all effectively closed classes $\mathcal{C}$ with $x \in \mathcal{C}$ we have that $\mathcal{C}$ is not porous at x.*

References

1. Laurent Bienvenu, Adam Day, Mathieu Hoyrup, Ilya Mezhirov, and Alexander Shen. Ergodic-type characterizations of algorithmic randomness. To appear in Information and Computation.
2. Vladimir I. Bogachev. *Measure theory. Vol. I, II.* Springer-Verlag, Berlin, 2007.
3. Vasco Brattka, Joseph S. Miller, and André Nies. Randomness and differentiability. Submitted.
4. Andrew M. Bruckner. *Differentiation of real functions*, volume 659 of *Lecture Notes in Mathematics*. Springer, Berlin, 1978.
5. Osvald Demuth. *Preprint for* Remarks on Denjoy sets. Technical report, 1988.
6. Osvald Demuth. Reducibilities of sets based on constructive functions of a real variable. *Commentationes Mathematicae Universitatis Carolinae*, 29(1):143–156, 1988.
7. Rodney Downey and Denis Hirschfeldt. *Algorithmic randomness and complexity.* Theory and Applications of Computability. Springer, 2010.
8. Johanna Franklin and Keng Meng Ng. Difference randomness. *Proceedings of the AMS*, 139:345–360, 2011.
9. Johanna N.Y. Franklin, Noam Greenberg, Joseph S. Miller, and Keng Meng Ng. Martin-Löf random points satisfy Birkhoff's ergodic theorem for effectively closed sets. To appear in Proceedings of the American Mathematical Society.
10. Per Martin-Löf. The definition of random sequences. *Information and Control*, 9:602–619.
11. André Nies. *Computability and randomness.* Oxford Logic Guides. Oxford University Press, 2009.
12. G. Tseitin. Algorithmic operators in constructive metric spaces. *Transactions of the American Mathematical Society*, 64, 1967.

A Appendix: Formal proofs

A.1 Proof of Lemma 3

Before we can prove Lemma 3 we need the following easy lemma.

▶ **Lemma 22.** *Let $\mathcal{C}$ be an effectively closed class. Let $h\colon \subseteq [0,1] \to \mathbb{R}_0^+$ be a computable function with domain containing $\mathcal{C}$. Then $\sup_{x\in\mathcal{C}}\{h(x)\}$ is right-c.e. and $\inf_{x\in\mathcal{C}}\{h(x)\}$ is left-c.e. uniformly in an index for $\mathcal{C}$.*

Proof. We proof the statement for the supremum; the proof for the infimum is analogous. We use the signed digit representation of reals, that is every real is represented by an infinite sequence in $\{-1,0,1\}^\infty$.

We run in parallel an enumeration of $\mathcal{C}^c$ and for all $x \in [0,1]$ (given by a Cauchy name $(x_n)_{n\in\mathbb{N}}$) the computations of $h(x)$ up to precision 2^{-n}. That is, we want to compute $(h(x))_n$, the n-th entry of a Cauchy name for $h(x)$.

Due to uniform continuity, for each n there is a number n' such that for all x in order to accomplish the computation it suffices to have access to the initial segment $(x_0,\dots,x_{n'})$ of the Cauchy name of x. When the computation of $(h(x))_n$ halts for some x it also halts for all other x' which have a Cauchy name that begins with $(x_0,\dots,x_{n'})$, since the computation is clearly the same. We do not know n', so we build a tree of computations that branches into three directions (-1, 0 and 1) whenever we access a new entry of the Cauchy name of the input. We remove a branch of the tree when it gets covered by $\mathcal{C}^c$. Due to the existence of n' the tree will remain finite.

Write $\sup_{x\in\mathcal{C}}\{h(x)\}[n]$ for the approximation to the value $\sup_{x\in\mathcal{C}}\{h(x)\}$ that we achieve when we proceed as described with precision level 2^{-n}. If we increase the precision, more of $\mathcal{C}^c$ may get enumerated before halting has occurred everywhere on the tree; so we see that the sequence $(\sup_{x\in\mathcal{C}}\{h(x)\}[n]+2^{-n})_{n\to\infty}$ is a right-c.e. approximation to $\sup_{x\in\mathcal{C}}\{h(x)\}$. ◀

Proof of Lemma 3. Since $\mathcal{C}$ is compact and closed, h is uniformly continuous on $\mathcal{C}$, that is, for every $\varepsilon > 0$ there exists a *single* $\delta(\varepsilon) > 0$ such that for any point $x \in E$ the continuity condition

$$|y-x| < \delta(\varepsilon) \Rightarrow |h(y)-h(x)| < \varepsilon$$

is satisfied.

Idea. We do not know $\delta(\varepsilon)$ for a given ε, but we can search for it using in parallel the following construction for different candidate δ's:

We split the unit interval into intervals $(I_k)_k$ of length δ and write l_k for the left border point of I_k. Write i_k for $\inf\{h(x) \mid x \in I_k \cap E\}$ and s_k for $\sup\{h(x) \mid x \in I_k \cap E\}$ if these values exist. We use Lemma 22 to approximate i_k and s_k for all intervals, and at the same time we enumerate $\mathcal{C}^c$, the complement of $\mathcal{C}$.

We do this until we have found a δ (called ε-*fit*) such that every interval has been dealt with; by this we mean that for every interval I_k we have either covered I_k with $\mathcal{C}^c$, or we have found that $s_k - i_k < \varepsilon$, that is we already know h up to precision ε. In the latter case we set our approximation to h to be the line from point (l_k, i_k) to point (l_{k+1}, s_k); on the remaining intervals (the former case) we interpolate linearly. Call the new function g_0. We can then output $g_0(x)$ up to precision ε at any point $x \in [0,1]$.

A problem with this construction. The following problem can occur with g_0: Assume we have for some ε found a δ that is ε-fit. We construct g_0 as described and interpolate linearly on, say, the maximal connected sequence $I_k,\dots,I_{k+n}$, all contained in $\mathcal{C}^c$. But if we look at the same construction for g_0 at a better precision $\varepsilon' < \varepsilon$, we might actually enumerate more of $\mathcal{C}^c$ until we find an ε'-fit δ', and this might extend the sequence $I_k, \ldots, I_{k+n}$ to, say, $I_{k-i}, \ldots, I_{k+n+j}$, all contained in $\mathcal{C}^c$. The linear interpolation on this sequence of intervals would then be *significantly flatter* than at level ε. So for some $x \in I_{k-i} \cup \cdots \cup I_{k+n+j}$ we might have that the approximation to g_0 with precision ε differs by more than ε from the approximation to g_0 with precision ε', which is not allowed.

To fix this problem we need to define g inductively over all precision levels, and "commit" to all linear interpolations that have happened at earlier precision levels, as will be described now.

Formal construction. Assume we want to compute the n-th entry of a Cauchy name for $g(x)$, that is we want to compute $g(x)$ up to precision 2^{-n}. We say that *we are at precision level n*. We do not know $\delta(2^{-n})$ so we do the following with $\delta = 2^{-p}$ in parallel for all p until we find a δ that is *n-fit*, defined as follows:

Split the interval $[0, 1]$ into intervals of length δ and write

$$I_k = [(k-1) \cdot 2^{-p}, k \cdot 2^{-p})$$

for the k-th interval and $l_k = (k-1) \cdot 2^{-p}$ for the left border point of I_k. For mathematical precision set $I_p := [1 - 2^{-p}, 1]$. Call an interval I_k n-treated if there exists a smaller precision level $n' < n$ where I_k has been covered by $\mathcal{C}^c$ and therefore a linear interpolation on I_k has been defined. We say that δ is *n-fit* if

- for every interval I_k we have that
  1. I_k is n-treated or
  2. $I_k \cap E = \emptyset$ or
  3. $s_k - i_k < 2^{-n}$
- and if for all k, where both I_k and I_{k+1} fulfill condition 3, we have $i_{k+1} < s_k$; that is, we have that intervals that directly follow each other have a "vertical overlap" in their approximations.

The following linear interpolation is an 2^{-n}-close approximation to g:

First, inductively replay the construction for all precision levels $n' < n$ to find all n-treated intervals. For the remaining intervals, run in parallel the right-c.e. and left-c.e. approximations to s_k and i_k, respectively, and the enumeration of $\mathcal{C}^c$, until for every interval either condition 2 or 3 are satisfied.

Build the following piecewise linear function: For all intervals that are already n-treated, keep the linear interpolations from the earlier precision level $n-1$. In all remaining intervals I_k that fulfill condition 3 we set g to be the line from point (l_k, i_k) to (l_{k+1}, s_k), that is, for $x \in I_k$ we let

$$g(x) = i_k + (x - l_k) \cdot \frac{(s_k - i_k)}{l_{k+1} - l_k} = i_k + (x - l_k)\frac{(s_k - i_k)}{\delta}.$$

Now look at the remaining intervals that have not yet been assigned a linear interpolation. In every maximal connected sequence $I_k, \ldots, I_{k+n}$ of such intervals every interval must fulfill condition 2. We interpolate linearly over $I_k, \ldots, I_{k+n}$ in the straightforward way, that is we draw a line from point $(s_k, g(l_k))$ to point $(s_{k+n+1}, g(l_{k+n+1}))$ (strictly speaking $g(l_k)$ is not yet defined, so use $\lim_{x \to l_k} g(x)$ instead).

Verification. It is clear that g is everywhere defined and non-decreasing. Whenever h was defined on a point $x \in I_k$ inside E, g gets assigned a value between i_k and s_k and since $s_k > h(x) > i_k$ and $s_k - i_k < 2^{-n}$ we have $|h(x) - g(x)| < 2^{-n}$. To see that g is computable note that s_k and i_k are defined on any interval that is not entirely contained in $\mathcal{C}^c$ and that these values can be approximated in a right-c.e. and left-c.e. way, respectively, by using lemma 22. ◀

A.2 Proof of Theorem 15

Proof. (1) ⇒ (3) is a result of Demuth [5].

(3) ⇒(2): Let f be a nondecreasing computable function. Then f satisfies the Denjoy alternative at z. Since $\underset{\sim}{D}f(z) \geq 0$, this means that $f'(z)$ exists. This implies that z is computably random by Brattka et al. [3, Thm. 4.1].

(2) ⇒ (1): Given a binary string σ, we write $S_f(\sigma)$ to mean $S_f(a,b)$ where $(a,b) = [\![\sigma]\!]$.

By hypothesis z is incomputable, and in particular not a rational. Suppose that the function g is Markov computable and $\underset{\sim}{D}g(z) = +\infty$. Choose dyadic rationals a, b such that $(a,b) = [\![\sigma]\!]$ for some string σ, $z \in (a,b) \subseteq [0,1]$ and $S_g(r,s) > 4$ for each r,s such that $z \in (r,s) \subseteq (a,b)$.

Define a computable martingale M on extensions $\tau \succeq \sigma$ that succeeds on (the binary expansion of) z. In the following τ ranges over such extensions.

First note that $S_g(\tau)$ is a computable real uniformly in τ. Furthermore, the function $\tau \mapsto S_g(\tau)$ satisfies the martingale equality, and succeeds on z in the sense that its values are unbounded (even converge to ∞) along z. However, this function may have negative values; informally, this is as if we were allowed to "bet with debt" because we can increase our capital at a string $\sigma 0$ beyond $2S_g(\sigma)$ by incurring a debt, i.e. negative value, at $S_g(\sigma 1)$. We now define a computable martingale M that succeeds on z and does not bet with debt.

Let $M(\emptyset) = S_g(\emptyset)$. Suppose now that $M(\tau)$ has been defined and is positive.

Case 1. There is $u \in \{0,1\}$ such that, where $v = 1-u$, we have $S(\tau v)_1 < 1$ (this is the second term in the Cauchy name for the computable real $S(\tau v)$, which is at most 1/2 away from that real). Then $S(\tau v) < 2$. By choice of a, b we now know that z is not an extension of τv. Thus, we let M double its capital along τu, let $M(\rho) = 0$ for all $\rho \succeq \tau v$. (The martingale M stops betting on these extensions.)

Case 2. Otherwise. Then $S(\tau v) > 0$ for $v = 0, 1$. We let M bet with the same betting factors as S_g:

$$M(\tau u) = M(\tau)\frac{S_g(\tau u)}{S_g(\tau)} \quad \text{for } u = 0, 1.$$

Note that $M(\tau u) > 0$.

If Case 1 applies to infinitely many initial segments of the binary expansion of z, then M doubles its capital along z infinitely often. Since M has only positive values along z, this means that $\lim_n M(z{\upharpoonright}_n)$ fails to exist, whence z is not computably random by the effective version of the Doob martingale convergence theorem [7, Theorem 7.1.3].

Otherwise, along z, M is eventually in Case 2. So M succeeds on z because S_g does. ◀

A.3 Proof of Theorem 18

We will use a finite extension argument which we present in the language of forcing. Our set of conditions $\mathbb{P}$ will have the form $\langle M, \sigma, \varepsilon \rangle$ where M is a computable martingale, σ a string and ε a rational, and such that $M(\sigma) < \varepsilon$. We say that $\langle M', \sigma', \varepsilon' \rangle$ extends $\langle M, \sigma, \varepsilon \rangle$ which we write $\langle M', \sigma', \varepsilon' \rangle \leq \langle M, \sigma, \varepsilon \rangle$ if

- $M' \geq M$
- $\sigma \preceq \sigma'$ and $M'(\tau) < \varepsilon$ for all $\sigma \preceq \tau \preceq \sigma'$

- $\varepsilon' \leq \varepsilon$

To each condition $p = \langle M, \sigma, \varepsilon \rangle$ we associate the effectively closed set (of positive measure)

$$L_{\langle M,\sigma,\varepsilon\rangle} = \{x \in 2^\omega \mid (\forall \tau)\ \sigma \preceq \tau \prec x \to M(\tau) < \varepsilon\}$$

(notice that $\langle M', \sigma', \varepsilon' \rangle \leq \langle M, \sigma, \varepsilon \rangle$ implies $L_{\langle M',\sigma',\varepsilon'\rangle} \subseteq L_{\langle M,\sigma,\varepsilon\rangle}$)

We claim that for a sufficiently generic filter $G \subseteq \mathbb{P}$, the closed set

$$\bigcap_{\langle M,\sigma,\varepsilon\rangle \in G} L_{\langle M,\sigma,\varepsilon\rangle}$$

is a singleton $\{x\}$ and x is Martin-Löf random but satisfies the Denjoy alternative for every Markov computable function.

Claim 1. For any filter $G \subseteq \mathbb{P}$, the set $\mathcal{G} = \bigcap_{\langle M,\sigma,\varepsilon\rangle \in G} L_{\langle M,\sigma,\varepsilon\rangle}$ is non-empty. If G is sufficiently generic, $\mathcal{G}$ is in fact a singleton x which is equal to the union of the strings appearing in the conditions of G.

Proof. By compactness, if $\mathcal{G}$ is empty, then there are finitely many conditions $\langle M_i, \sigma_i, \varepsilon_i \rangle$ such that $\bigcap_i L_{\langle M_i,\sigma_i,\varepsilon_i\rangle}$. Since G is a filter, let $\langle M^*, \sigma^*, \varepsilon^* \rangle$ be a condition extending all the $\langle M_i, \sigma_i, \varepsilon_i \rangle$. We have $\bigcap_i L_{\langle M_i,\sigma_i,\varepsilon_i\rangle} \supseteq L_{\langle M^*,\sigma^*,\varepsilon^*\rangle}$, and the latter is non-empty as it has positive measure.

Now, for all n, let D_n to be the set of conditions $\langle M, \sigma, \varepsilon \rangle$ with $|\sigma| \geq n$. One can see that D_n is dense: indeed, for a condition $\langle M, \sigma, \varepsilon \rangle$, if $|\sigma| < n$, then diagonalizing against M during $n - |\sigma|$ steps, one can find an extension τ of σ of length at least n such that $\langle M, \tau, \varepsilon \rangle$ extends $\langle M, \sigma, \varepsilon \rangle$. Therefore if G is sufficiently generic, it contains conditions $\langle M, \sigma, \varepsilon \rangle$ for arbitrarily long σ. Since G is a filter, all the strings appearing in its elements must be comparable, hence there is a unique real x that extends them all. Therefore $\mathcal{G}$ contains at most the singleton $\{x\}$. Since $\mathcal{G}$ is non-empty, it is equal to the singleton $\{x\}$. ♢

From now on we assume that "sufficiently generic" means in particular that $\mathcal{G}$ is a singleton $\{x\}$.

Claim 2. If G is sufficiently generic, and $\langle M, \sigma, \varepsilon \rangle$ is a condition in G, then $M(x{\upharpoonright}n) < \varepsilon$ for all $n \geq |\sigma|$.

Proof. Trivial by definition of x. ♢

Claim 3. If G is sufficiently generic, x is computably random.

Proof. This is the usual argument. Let N be a computable martingale. Let $\langle M, \sigma, \varepsilon \rangle$ be a condition. Let δ be a rational such that $M(\sigma) < \varepsilon - \delta$. Then, setting $M' = M + 2^{-|\sigma|}\delta N$, it is easy to see that $\langle M', \sigma, \varepsilon \rangle$ is a condition, and that N does not succeed on any element of $L_{\langle M',\sigma,\varepsilon\rangle}$. ♢

Claim 4. If G is sufficiently generic, and $\langle M, \sigma, \varepsilon \rangle$ is a condition in G, then x is a density point of $L_{\langle M,\sigma,\varepsilon\rangle}$.

Proof. Suppose $\langle M, \sigma, \varepsilon \rangle$ is a condition, i.e., $M(\sigma) < \varepsilon$. We want to show that x is a density point of $L_{\langle M,\sigma,\varepsilon\rangle}$ i.e. that the conditional measure of $L_{\langle M,\sigma,\varepsilon\rangle}$ below $x{\upharpoonright}n$ tends to 1. Fix δ arbitrarily small. Let $\varepsilon' < \varepsilon$ be the infimum of the values M can take on extensions of σ, i.e.,

$$m = \inf\{M(\tau) \mid \tau \succeq \sigma\}$$

Then take ε' a rational between m and ε that is much closer to m than to ε. More precisely, take it such that

$$\frac{\varepsilon' - m}{\varepsilon - m} < \delta$$

Now, pick $\tau \succeq \sigma$ such that $M(\tau) < \varepsilon'$ (which must exist by definition of m and ε'). It is clear that $\langle M, \tau, \varepsilon' \rangle$ is an extension of $\langle M, \sigma, \varepsilon \rangle$. We claim that if this new condition is in G, then the conditional measure of $L_{\langle M,\sigma,\varepsilon\rangle}$ below $x{\upharpoonright}n$ for all $n \geq |\tau|$ is at least $1 - \delta$, which will complete the proof.

To prove this claim, pick some $n \geq |\tau|$ and consider the initial segment $x{\upharpoonright}n$, which extends τ. Define the martingale N by $N(\nu) = M(x{\upharpoonright}n^\frown\nu) - m$. This is a martingale of initial capital $M(\tau) - m$ which is non-negative by definition of m. Therefore, the Ville-Kolmogorov inequality yields

$$\mu\big\{z \mid (\forall k)\ N(z{\upharpoonright}k) < \varepsilon - m\big\} \geq 1 - \frac{M(\tau) - m}{\varepsilon - m} > 1 - \delta$$

Now, notice that by definition of N, the measure $\mu\big\{z \mid (\forall k)\ N(z{\upharpoonright}k) < \varepsilon - m\big\}$ is exactly $\mu_{x\upharpoonright n}(L_{\langle M,\tau,\varepsilon\rangle})$, but also $\mu_{x\upharpoonright n}(L_{\langle M,\sigma,\varepsilon\rangle})$ as τ is an extension of σ. Therefore, we have

$$\mu_{x\upharpoonright n}(L_{\langle M,\sigma,\varepsilon\rangle}) > 1 - \delta$$

for all $n \geq |\tau|$. ◊

Claim 5. If G is sufficiently generic, then every constructive function satisfies the Denjoy alternative at x.

Proof. Let $\langle M, \sigma, \varepsilon \rangle$ be a condition and let f be a constructive function. We distinguish two cases.

Case 1. The elements of the set $L_{\langle M,\sigma,\varepsilon\rangle}$ have a uniform bound on their lower derivative. Formally,

$$(\exists n)\ \forall z \in L_{\langle M,\sigma,\varepsilon\rangle}\ \underset{\sim}{D}f(z) > -n$$

In this case, G being sufficiently generic ensures that x is computably random (Claim 3) and is a density point of $L_{\langle M,\tau,\varepsilon\rangle}$ (Claim 4), so we can directly apply Lemma 17, from which we get that f is differentiable at x.

Case 2. If the above case does not hold, then for any large given n, say $n = |\sigma| + 1$, one can find some $z \in L_{\langle M,\tau,\varepsilon\rangle}$ and $a, b \in \mathbb{R}_c$ with $a < b$ such that

- $z \in (a, b)$
- $|b - a| < 2^{-n}$
- $\frac{f(b)-f(a)}{b-a} < -n$

In that case, one can extend $\langle M, \sigma, \varepsilon \rangle$ to $\langle M, \tau, \varepsilon \rangle$ where $[\tau] \subseteq (a, b)$.

Now, G being sufficiently generic, we ensure that either $\underset{\sim}{D}f(x) = -\infty$ (when there are conditions $\langle M, \sigma, \varepsilon \rangle$ in G with arbitrarily large σ that put us in Case 2) or that f is differentiable at x (when for all sufficiently large σ, all $\langle M, \sigma, \varepsilon \rangle$ in G put us in Case 1). ◊

Claim 6. If G is sufficiently generic, x is not Martin-Löf random.

Proof. This part is quite standard. Let $\langle M, \sigma, \varepsilon \rangle$ be a condition and let c be a constant. Since one can computably diagonalize against a computable martingale there exists a computable sequence z extending σ such that $M(z{\upharpoonright}n) < \varepsilon$ for all $n > |\sigma|$. Since z is computable, it is possible to take n large enough so that $\tau = z{\upharpoonright}n$ satisfies $K(\tau) < |\tau| - c$. This proves that a sufficiently generic G will yield a sequence x that is not Martin-Löf random. ◊

A.4 Proof of Theorem 20

Here we give the details of the proof that were left out in the main text.

Formal construction. Look at the following effective procedure $P(\sigma, k, \varepsilon)$ that enumerates the complement of $\mathcal{C}_\varepsilon$.

1) Pick a fresh integer $n = n(\sigma, k)$.
2) Enumerate the set $\mathcal{V}$ of reals y that extend σ and satisfy $\Gamma^y(n) = 0$. Due to the use principle this will always happen on whole open cylinders of y's. Therefore we can represent $\mathcal{V}$ by a c.e. enumeration of finite strings. It is easy to see that we can force this enumeration to be prefix-free. As long as the conditional measure of $\mathcal{V}$ above σ does not exceed $1 - \varepsilon$, put every enumerated string τ into U_k and for each of them start the procedure $P(\tau, k + 1, \varepsilon)$.
3) If at some stage s some new τ is found such that $\Gamma^\tau(n) = 0$ but the conditional measure $U_k[s] \cup \{\tau\}$ above σ exceeds $1 - \varepsilon$, then enumerate n into W, enumerate all of $U_k[s] \cup \{\tau\}$ in the complement of $\mathcal{C}_\varepsilon$, and terminate the whole tree of procedures.

Verification. We need to verify that running $P(\lambda, 0, \varepsilon)$ actually produces $\mathcal{C}_\varepsilon$ as desired. First, notice that x remains in $\mathcal{C}_\varepsilon$ at all time during the construction. Indeed, when we put a string τ into the complement of $\mathcal{C}_\varepsilon$ during some procedure $P(\sigma, k, \varepsilon)$ at step 3, this τ satisfies $\Gamma^\tau(n) = 0$, and we precisely make sure at that point that $n \in W$. Since $\Gamma^x = W$, this shows that τ is not a prefix of x.

Secondly, note that if for some prefix σ of x and some k, the procedure $P(\sigma, k, \varepsilon)$ gets executed and reaches step 3, then we are done since step 3 ensures that $\lambda(\mathcal{C}_\varepsilon \mid \sigma) < \varepsilon$. So we need to argue that the conditions for step 3 will eventually be met. Assume that this is not the case. Then the following claim states that U_k will become defined for all $k \in \mathbb{N}$.

Claim. Let σ be a prefix of x and k an integer. If the procedure $P(\sigma, k, \varepsilon)$ is executed but never reaches step 3, then some longer prefix σ' of x is put into U_k and procedure $P(\sigma', k + 1, \varepsilon)$ is executed.

Proof. Indeed, if step 3 is never reached, then $n = n(\sigma, k)$ never gets enumerated in W, therefore $\Gamma^x(n) = W(n) = 0$, which by definition of step 2 shows that some prefix σ' of x is put into U_k and procedure $P(\sigma', k + 1)$ is executed. ◊

Observe that $(U_k)_k$ is a Martin-Löf test, by induction: every time a string σ is enumerated U_k, some of its extensions are enumerated into U_{k+1} during procedure $P(\sigma, k + 1, \varepsilon)$. The relative weight of these strings above σ is at most $1 - \varepsilon$, therefore we have $\lambda(U_{k+1}) \leq (1 - \varepsilon)\lambda(U_k)$ for all k. By induction this proves $\lambda(U_k) \leq (1 - \varepsilon)^k$ for all k.

The claim shows that if step 3 is never reached in any of the executed procedures $P(\sigma, k, \varepsilon)$ with σ a prefix of x, then $x \in [U_k]$ for all k. This contradicts the fact that x is Martin-Löf random. ◀