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, i.e., 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 and proved that Demuth randomness is strong enough to ensure the Denjoy alternative for Markov computable functions. In this paper, we show that the set of these points is indeed 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 results, 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 by the only very recently defined notion of difference randomness. This is the first analytical characterization of difference randomness. We also consider the concept of porous points, a special type of Lebesgue non-density points for which drops in density are witnessed by single intervals. 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, porosity, randomness

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

* The second author is supported by a Feodor Lynen postdoctoral research fellowship by the Alexander von Humboldt Foundation.
† The third author is supported by the National Science Foundation under grant DMS-1001847.

Conference title on which this volume is based on.
Editors: Billy Editor, Bill Editors; pp. 1–12
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

infinite binary sequences (or reals, as any real can be represented by an infinite binary sequence) a satisfactory definition was given by Martin-Löf [11]. 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 [8, 12] 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, that is, computable in some sense. Consider 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 [10] 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 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 had very few known characterizations other than its original definition, and in particular, no known analytical characterization.

Demuth [7] 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 exist 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 a 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 than the class of Demuth random reals. Difference randomness is a notion of randomness slightly stronger than Martin-Löf randomness and significantly weaker than Demuth randomness. We show that this notion already implies the Denjoy alternative 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 that 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. Finally we show that x is difference random if and only if x is Martin-Löf random and has positive density in every effectively closed class in which x is contained.

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 a binary sequence, finite or infinite, with length at least n, $x\restriction n$ 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 $x \in [0,1]$ and $q \in [0,1] \cap \mathbb{Q}$, we have $|x - q| < 2^{-\psi(n)} \Rightarrow |f(x) - \widehat{f}(q,n)| < 2^{-n}$. 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 $\varphi : \mathbb{N} \to \mathbb{N}$ such that for all $x \in \mathbb{R}_c$, if i is a name for x, then $\varphi(i)$ 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 sometimes use the notation $f(x)_s$ to denote the approximation of $f(x)$ at precision level 2^{-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 [13] 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 $\mathrm{dom}(f)$ is a dense subset of $[0,1]$, one can consider the lower and upper *pseudo*-derivatives defined by:

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

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

Note that when the function f is continuous on its (dense) domain, which is the case for computable and Markov computable functions, one can replace $\mathrm{dom}(f)$ by any dense subset of $\mathrm{dom}(f)$ in the definition of $\underset{\sim}{D}f$ and $\widetilde{D}f$. For Markov computable functions, for example, one could use $\mathbb{Q}$ instead of $\mathbb{R}_c$ to define the pseudo-derivatives.

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. We will need the following technical lemma. We omit the proof due to space considerations.

▶ **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. If $(\mathcal{U}_n)$ is a sequence of open sets, it is said 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 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 *weakly 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 [9]. 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 [8]) 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 with $\lambda(\mathcal{U}_n) < 2^{-n}$ for all n, 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 the 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, weakly 2-random, difference random, Demuth random and computably random sequences respectively.

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

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

hold and no other implication holds in general (other than those which can be derived by transitivity from the above diagram). See for example [12] 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 (lower Lebesgue) density ρ of a set $\mathcal{C} \subseteq \mathbb{R}$ at a point x to be the quantity

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

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 originates in the work of Denjoy. See for instance [2, 5.8.124] (but note the typo in the definition there); also [4, Ex. 7:9.12].

▶ **Definition 11.** We say that $\mathcal{C}$ is *porous at x* via $\varepsilon > 0$ if for each $\alpha > 0$ there exists β with $0 < \beta \le \alpha$ 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 a *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 for almost every point x of a measurable class $\mathcal{C}$, we have that $\mathcal{C}$ is not porous at x.

▶ **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 $c \in \mathbb{N}$ such that $\mathcal{C}$ is porous at z via 2^{-c+2}. For each string σ consider the set of minimal "porous" extensions at stage t,

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

We claim that

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

To see this, let R be the set of strings ρ in (1). Let V be the set of prefix-minimal strings that occur as witnesses τ in (1). Then the open sets generated by R and by V are disjoint. Thus, if r and v denote their measures, respectively, we have $r + v \le 2^{-|\sigma|}$. By definition of $N_t(\sigma)$, for each $\rho \in R$ there is $\tau \in V$ such that $|0.\tau - 0.\rho| \le 2^{-|\tau|+c}$. This implies $r \le 2^{c+1} v$. The two inequalities together imply (1).

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. 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 $N_{t+1}(\sigma')$. $\Diamond$

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

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

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, $U_n = \bigcup_t [\![B_{n,t}]\!]$ is a nested union, so the second claim implies that $\lambda(\mathcal{C} \cap U_n) \leq (1-2^{-c-2})^n$. We show $z \in \bigcap U_n$ by induction on n. Clearly $z \in U_0$. If $n > 0$ suppose inductively $\sigma \prec z$ where $\sigma \in \bigcup_t B_{n-1,t}$. Since z is random there is η such that $\sigma \prec \eta \prec z$ and η ends in 0^c1^c. 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 $(b-a)/4$. Thus, since $\mathcal{C}$ is porous at z via 2^{-c+2}, there is t, $\rho \succeq \eta$ and τ satisfying the condition in the definition of $N_t(\sigma)$. By the choice of η one verifies that $\tau \succeq \sigma$. Thus $z \in U_n$.

Now take a computable subsequence $(U_{g(n)})_{n\in\mathbb{N}}$ such that $\lambda(\mathcal{C} \cap U_{g(n)}) \leq 2^{-n}$ to obtain a difference test that z 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 [5]) states that the Denjoy alternative holds at almost all points for *any* function f.

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

▶ **Definition 14** (Demuth [6]). 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 [6, 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 need for (2) ⇒ (1) is that $f(q)$ 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 f by this weaker hypothesis.

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

We now turn to the proof of Theorem 1. 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 $x \in$ 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 $\underline{D}f(z) > -n$ 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 $f(x) + (n+1)x$, we may assume that for $x \in \mathcal{C}$ and some r and s, we have

$$\forall a, b\ [r < a \le x \le b < s \to S_f(a,b)_0 > 0].$$

We restrict our attention to the interval $[r,s]$; for notational simplicity we assume that $[r,s] = [0,1]$. Let $f_*(x) = \sup_{a \le x} f(a)$. Then f_* is nondecreasing on $\mathcal{C}$.

Claim. The function $f_* \upharpoonright_{\mathcal{C}}$ 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 for $p, q \in (r,s)$ we have $p < x \leftrightarrow f(p) < d$, 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

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

furthermore, since $\mathcal{C}$ is not porous at x, we may assume that 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

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

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:

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

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 that is not Martin-Löf random but nonetheless satisfies the Denjoy alternative for all Markov computable 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 x would be a density point of L. Since $x \in \mathsf{CR}$ 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 $L := \{x \succeq \sigma \mid M(x{\upharpoonright}n) < \varepsilon$ 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 $m := \inf_{\tau\succeq\sigma} M(\tau)$. 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 $M(\sigma)$ is arbitrarily close to m and thereby raise the density to the desired level $1-q$. Although this method only controls the density in L of the dyadic intervals containing x, we are able to show that this is sufficient. ◀

▶ **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 $\mathcal{U}_1$. It is not hard to see that since $\mathcal{U}_1$ is a c.e. open set, it is an effective union $\bigcup_t I_t$ of *closed* rational intervals I_t that overlap only on their endpoints. Let $\mathcal{U}_1[s] = \bigcup_{t<s} I_t$ and let Ω_s be the leftmost point of $\mathcal{U}_1[s]$. Then Ω is approximated from below by the computable sequence of rationals $(\Omega_s)_s$.

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 $s+1$, consider I_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 I_s.
2. Or, this interval does change the value of Ω: $\Omega_{s+1} > \Omega_s$. In this case, define F on I_s to be the triangular function taking value 0 on the endpoints of I_s 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 $a < x < b$ such that the interval $[a, b]$ is contained either in one or in the union of two of the intervals 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 for each interval, and both the zero function and the triangular function are computable on $\mathbb{R}_c$. (In the triangular case, note 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 computably real points arbitrarily close to Ω (at least the endpoints of intervals I_s 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 $2^{-n/2}$.Thus, letting x_k be middle point of $[\Omega_s, \Omega_{s+1}]$ and $q > \Omega$ be a rational such that $q - a_k < 2^{-k}$, we have

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

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

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 at most $2^{-n/2}$. Therefore $\sum_n ||h_n|| < \infty$, so by the Weierstrass M-test we can conclude that the convergence is uniform and 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 Martin-Löf random real x.*

1. *x is difference random.*
2. *x has incomplete Turing degree.*
3. *x is a point of positive lower Lebesgue 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 [9].

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

$$U_k = \{z \mid \exists \text{ interval } I\colon z \in I \text{ and } \lambda(\mathcal{C} \mid I) < 2^{-k}\}.$$

Since $\mathcal{C}$ is effectively closed, these classes are uniformly effectively open; and clearly $x \in \mathcal{C}$. The measure bound on this difference test follows from the following lemma, the proof of which we omit due to space considerations.

▶ **Lemma 21.** *Let $\mathcal{C} \subseteq [0, 1]$ be closed. Let $U_k = \{z \mid \exists$ interval $I\colon z \in I$ and $\lambda(\mathcal{C} \mid I) < 2^{-k}\}$. Then $\lambda(\mathcal{C} \cap U_k) \leq 2^{-k+1}$.*

(3) ⇒ (2): We only sketch the proof due to space considerations.

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 $\lambda(\mathcal{C}_\varepsilon \mid x{\upharpoonright}n) < \varepsilon$ 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 do not 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 in a neighborhood and wait until we see a certain type of behavior (the reduction outputs 0 on a certain value) 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 $x \in \mathsf{MLR}$. ◀

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

▶ **Corollary 22.** *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.*

▶ Remark. If x is weakly 2-random and $\mathcal{C}$ is an effectively closed class containing x, then $\rho(x|\mathcal{C}) = 1$. This is because for x to have $\rho(x|\mathcal{C}) < 1$ in some $\mathcal{C}$ can be written as $\forall \delta_0\ \exists \delta < \frac{\delta_0}{2} : \frac{\lambda([x-\delta,x+\delta]\cap\mathcal{C})}{\lambda([x-\delta,x+\delta])} < 1-\varepsilon$ for some fixed ε, which is a Π^0_2 condition. By the Lebesgue density theorem, the set of these x (for each $\mathcal{C}$) is also null, so they are covered by a strong test.

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 Bruckner, Judith Bruckner, and Brian Thomson. *Real Analysis.* Prentice Hall (Pearson), 2007.

5 Andrew M. Bruckner. *Differentiation of real functions*, volume 659 of *Lecture Notes in Mathematics*. Springer, Berlin, 1978.

6 Osvald Demuth. *Preprint for* Remarks on Denjoy sets. Technical report, 1988.

7 Osvald Demuth. Reducibilities of sets based on constructive functions of a real variable. *Commentationes Mathematicae Universitatis Carolinae*, 29(1):143–156, 1988.

8 Rodney Downey and Denis Hirschfeldt. *Algorithmic randomness and complexity.* Theory and Applications of Computability. Springer, 2010.

9 Johanna Franklin and Keng Meng Ng. Difference randomness. *Proceedings of the AMS*, 139:345–360, 2011.

10 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.

11 Per Martin-Löf. The definition of random sequences. *Information and Control*, 9:602–619.

12 André Nies. *Computability and randomness.* Oxford Logic Guides. Oxford University Press, 2009.

13 G. S. Tseitin. Algorithmic operators in constructive metric spaces. *Transactions of the American Mathematical Society*, 64, 1967.