Computable functions of bounded variation and the complexity of Jordan decomposition

Marcus Anthony Triplett

The University of Auckland

A dissertation submitted in partial fulfillment of the requirements for the degree of *Bachelor of Science (Honours)* in *Computer Science*

November 2015

COMPUTABLE FUNCTIONS OF BOUNDED VARIATION
AND THE COMPLEXITY OF JORDAN DECOMPOSITION

Marcus Anthony Triplett
Supervisor: André Nies

Contents

1 Introduction

The *total variation* of a real valued function is a measure of its oscillation. A function is said to be of *bounded variation* if its total variation is finite. While functions of bounded variation are well-behaved in the sense that they are both a.e. differentiable and form a Banach algebra under the so-called *variation norm*, they can be difficult to study computationally since the space of bounded variation functions is not separable. A classical result due to Camille Jordan in 1881 states that any function of bounded variation is the difference of a pair of non-decreasing functions, called a *Jordan decomposition* [14]. With the development of measure theory and functional analysis in the early 20th century, the concepts of variation and decomposition were extended. Hans Hahn showed that any signed and bounded Borel measure can be decomposed as the difference of two non-negative Borel measures, and similarly it was shown that a functional from the dual space of $C[0, 1]$ could be decomposed as the difference of two non-negative linear functionals.

This dissertation is an exposition on the *meta*mathematics of variation and Jordan decomposition. We provide a detailed logical analysis of the class of functions of bounded variation using insights from active programmes in mathematical logic such as computable analysis, reverse mathematics, and algorithmic randomness. Our goal is to uncover effective aspects of classical results from analysis and measure theory concerning functions of bounded variation by addressing a number of questions discussed below.

(1) If f is computable, can one always find an *effective* Jordan decomposition?

The computational content of classical analysis has been studied since Turing's semantics for computation. Computable analysis is an approach to mathematics which enforces effectivity requirements on its objects of concern and determines whether classical theorems still hold. That is, computable analysis is the theory of *computable* real numbers, real functions, metric spaces, etc. For instance, the monotone convergence theorem is not *computably true*, in the sense that there is a bounded computable monotone sequence of real numbers whose limit is not a computable real. Using the methods of Zheng and Rettinger [30], we show that the Jordan decomposition theorem is not computably true: there is a computable function of bounded variation which is not the difference of two non-decreasing computable functions by Theorem 3.4. The result can even be strengthened to the case of polynomial time computability.

(2) What is the complexity of the operator which maps a function of bounded variation to its (canonical) Jordan decomposition?

With Weihrauch's type two theory of effectivity, we shift to a more general analytic setting. We describe elements of dual spaces and spaces of Borel measures by infinite binary sequences that are processed by generalised Turing machines. Using this more general framework we study the computational difficulty of decomposing these more complex objects. Theorems 7.1, 7.9, and 7.10 show that, under suitable representations, the Jordan decomposition operator is computable on the three Banach spaces consisting of functions of bounded variation, continuous linear functionals, and signed Borel measures.

(3) How difficult is Jordan's theorem to prove?

In light of the impossibility theorems which obstruct foundational programmes in mathematics, much research into the foundations of mathematics has shifted towards revealing complexity in the existing foundation. Sufficiently powerful systems of arithmetic are capable of capturing fundamental theorems in analysis, combinatorics, and algebra, among other areas. Varying the strength of such systems influences what theorems can be proven within them, and thereby provide calculi for fragments of mathematical deduction that enable us to study and measure the complexity of different theorems. For instance, the Bolzano-Weierstraß theorem (hereafter BW) states that any bounded sequence of real numbers has a convergent subsequence. The system of second-order arithmetic ACA_0 is capable of proving BW, whereas the weaker system RCA_0 is not. On the other hand, the intermediate value theorem (IV, which states that the continuous image of an interval is again an interval) *can* be proven in RCA_0. This indicates a difference in complexity between the two theorems: the statement BW is of a *higher complexity* than IV because it requires a stronger calculus for its proof.

Reverse mathematics is a programme in the foundations of mathematics which explicates the complexity of theorems in the way so described. Given a statement formulated in the appropriate formal language, which axioms are needed to prove that statement? For BW one needs *arithmetical comprehension*, while for IV *recursive comprehension* is enough. A proof-theoretic analysis undertaken by mathematical logicians over past decades has revealed that just five particular subsystems of second-order arithmetic are enough to capture almost all of the theorems in non-set-theoretic mathematics which are considered important. These systems actually define complexity *classes* for theorems. In many cases a theorem is not only provable within a system, but is in fact equivalent to that system over RCA_0. For example, BW and the principle of arithmetical comprehension are equivalent over RCA_0. In this way we demonstrate that arithmetical comprehension is not just sufficient to prove BW, but is indeed *essential* to proving it.

The large scale characterisation of theorems into just five systems has been a major triumph of the reverse mathematics programme. Traditionally these theorems have been mined from pre-20th century mathematics, but in recent years the scope of reverse mathematical investigation has broadened. For example, the complexity of a number of recurrence theorems in topological dynamics [7], and of convergence theorems in measure theory [1] have been characterised. The negative answer to Question (1) actually shows that the Jordan decomposition theorem cannot be proven in RCA_0. This motivates the investigation into a characterisation of its complexity in Section 5.

(4) How strong is the principle asserting the differentiability of functions of bounded variation?

The work of the Czech constructivist Osvald Demuth [8] in the 1970s has recently led to new interactions between computable analysis and the theory of algorithmic randomness

[11, 22, 20, 5]. Effective analogues of theorems which assert the differentiability of a class of functions almost everywhere allow one to characterise the points of non-differentiability for those functions in terms of effective null-sets. Certain varieties of these effective null-sets turn out to correspond precisely to the null-sets used to define randomness notions. This realization has led to a number of principles that fit within the following form: for a class of computable functions $\mathcal{C}$ and a real $z \in [0,1]$,

z satisfies a certain randomness notion $\Leftrightarrow$ every function in $\mathcal{C}$ is differentiable at z.

For instance, by Theorem 4.9 below, a real z is Martin-Löf random if and only if every computable function of bounded variation is differentiable at z. We will prove and then use results of this kind in Section 6 to study the reverse mathematical complexity of statements which assert that functions of bounded variation have a point of differentiability. This allows us to formulate a precise theorem in second-order arithmetic which captures the idea that randomness is essential to differentiability.

2 Preliminaries on classical theory and computable analysis

We develop some of the classical theory of bounded variation functions. These notions will be appropriately modified for reverse mathematical analysis in Section 5. Unless otherwise mentioned all real-valued functions will have domain $[0,1]$ and codomain $\mathbb{R}$.

Definition 2.1. Let $f : [0,1] \to \mathbb{R}$. The *variation function for f* is the function $\boldsymbol{v}_f : [0,1] \to \mathbb{R} \cup \{\infty\}$ defined by

$$\boldsymbol{v}_f(x) = \sup \sum_{i=0}^{n-1} |f(t_{i+1}) - f(t_i)|$$

where the supremum is taken over all collections $t_0 < \cdots < t_n$ in $[0,x]$. Let $\mathbf{V}(f) = \boldsymbol{v}_f(1)$. The function f is said to be of *bounded variation* if $\mathbf{V}(f) < \infty$. The set of all functions of bounded variation is denoted by $\mathbf{BV}$.

Theorem 2.2 (The Jordan decomposition theorem)**.** *For any function $f : [0,1] \to \mathbb{R}$ of bounded variation, there are non-decreasing functions $f^+, f^- : [0,1] \to \mathbb{R}$ such that $f = f^+ - f^-$.*

For any function f of bounded variation, $\boldsymbol{v}_f$ and $(\boldsymbol{v}_f - f)$ are non-decreasing functions whose difference is f. Note that the expression $f^+ - f^-$ is not unique, but the pair $((\boldsymbol{v}_f + f)/2, (\boldsymbol{v}_f - f)/2)$ is the unique *minimal* decomposition for such an f. That is, for any other pair (g^+, g^-) of non-decreasing functions with $f = g^+ - g^-$ one has $(\boldsymbol{v}_f + f)/2 \leq g^+$ and $(\boldsymbol{v}_f - f)/2 \leq g^-$.

If $I \subseteq [0,1]$ is an interval we write $\mathbf{V}(f, I)$ for the variation of f on the interval I,

$$\mathbf{V}(f, I) = \sup \sum_{i=0}^{n-1} |f(t_{i+1}) - f(t_i)|,$$

where the supremum is taken over all collections $t_0 < \cdots < t_n$ in I. There is an associated norm $||\cdot||_{\mathbf{BV}} : \mathbf{BV} \to [0,\infty)$ on the set $\mathbf{BV}$ given by

$$||f||_{\mathbf{BV}} = |f(0)| + \mathbf{V}(f).$$

We call $||\cdot||_{\mathbf{BV}}$ the *variation norm.* Note that $||f||_{\mathbf{BV}} \geq ||f||_\infty$, and therefore convergence under the variation norm implies uniform convergence. In fact, $\mathbf{BV}$ is a Banach space under the variation norm.

Recall that a function $f : [0,1] \to \mathbb{R}$ is *absolutely continuous* if for any $\varepsilon > 0$ there is $\delta > 0$ such that

$$\sum_{i=0}^{n-1} |f(t_{i+1}) - f(t_i)| < \varepsilon,$$

for any partition $t_0 < \cdots < t_n$ of $[0,1]$ with $\sum_{i=0}^{n-1} |t_{i+1} - t_i| < \delta$. It is not difficult to show that any absolutely continuous function is of bounded variation, and thus the vector space of absolutely continuous functions $(\mathbf{AC}, ||\cdot||_{\mathbf{BV}})$ is a subspace of $(\mathbf{BV}, ||\cdot||_{\mathbf{BV}})$. Let L_1 be the space of all integrable functions on $[0,1]$ with respect to the Lebesgue measure. Let $\mathcal{L}_1$ be the usual quotient structure of L_1 modulo a.e. equality. Then the map $f \mapsto f'$ is an isometry from $(\mathbf{AC}, ||\cdot||_{\mathbf{BV}})$ to $(\mathcal{L}_1, ||\cdot||_1)$ whose inverse is the integration map. We will use this fact in Section 4.2 to construct a computable function of bounded variation whose points of differentiability are precisely the Martin-Löf random reals.

For an interval $A \subseteq [0,1]$, $h \in \mathbb{Q}$, and $r \in \mathbb{N}$, let

$$\mathrm{M}_A(h,r)$$

denote a *sawtooth function on the interval A with r-many teeth of height h*; so

$$\mathrm{M}_A(h,r)(x) = 0$$

whenever $x \notin A$. Note that $\mathbf{V}(\mathrm{M}_A(h,r), A) = 2hr$. We will often use such sawtooth functions to construct functions which encode computationally useful information in their variation. For instance, in Section 3 we show how to build a computable function whose total variation computes the halting problem.

There are results similar to Theorem 2.2 for continuous linear functionals and signed Borel measures. Let $(C^*[0,1], ||\cdot||)$ be the space of continuous linear functionals $F : C[0,1] \to \mathbb{R}$ with the norm $||F|| = \sup\{F(h) : h \in C[0,1], ||h||_\infty \leq 1\}$.

Theorem 2.3 (Lemma 8.13 [2]). *For every functional $F \in C^*[0,1]$ there are non-negative functionals $F^+, F^- : C[0,1] \to \mathbb{R}$ such that $F = F^+ - F^-$.*

Let $\mathcal{B}$ be the Borel σ-algebra on $[0,1]$, and let $\mathbf{BM}$ be the set of all signed Borel measures $\mu : \mathcal{B} \to \mathbb{R}$. We define the *variation norm* $||\cdot||_m : \mathbf{BM} \to [0,\infty)$ by $||\mu||_m = \sup \sum_{I\in\pi} |\mu(I)|$, where the supremum is taken over all partitions π of $[0,1]$ into finitely many intervals. Note that if μ is non-negative, then $||\mu||_m = \mu([0,1])$ by σ-additivity.

Theorem 2.4 (Corollary 3.1.2 [3])**.** *For any $\mu \in \mathbf{BM}$ there are non-negative Borel measures $\mu^+, \mu^- : \mathcal{B}[0,1] \to \mathbb{R}$ such that $\mu = \mu^+ - \mu^-$.*

The spaces $(C^*[0,1], ||\cdot||)$ and $(\mathbf{BM}, ||\cdot||_m)$ are Banach spaces, and as in Theorem 2.2, for Theorem 2.3 and Theorem 2.4 there are minimal decompositions. Hence we make the following specifications.

1. For each functional $F \in C^*[0,1]$, the *minimal Jordan decomposition* of F is the pair (F^+, F^-) such that $F^+, F^- \in C^*[0,1]$ are non-negative and $F = F^+ - F^-$, and for any other pair (G^+, G^-) with $G^+, G^- \in C^*[0,1]$ non-negative and $F = G^+ - G^-$ one has $F^+ \le G^+$ and $F^- \le G^-$.

2. For each $\mu \in \mathbf{BM}$, the *minimal Jordan decomposition* of μ is the pair (μ^+, μ^-) such that $\mu^+, \mu^- \in \mathbf{BM}$ are non-negative and $\mu = \mu^+ - \mu^-$, and for any other pair (ν^+, ν^-) with $\nu^+, \nu^- \in \mathbf{BM}$ non-negative and $\mu = \nu^+ - \nu^-$ one has $\mu^+ \le \nu^+$ and $\mu^- \le \nu^-$.

By the Riesz representation theorem, every functional $F \in C^*[0,1]$ can be written as the Riemann-Stieltjes integral with respect to a unique function $g \in \mathbf{BV}$. The effective aspects of Riesz's theorem have been explored by Jafarikhah and Weihrauch in [17, 12]. This work was necessary to establish their results concerning the type two computability of the Jordan decomposition operator for linear functionals and signed Borel measures. We present these results in Section 7.

There are further analogues of the Jordan decomposition theorem, such as the Doob decomposition theorem, which states that any submartingale can be decomposed into the sum of a martingale and an increasing stochastic process. The effective content of this result will not be considered in this dissertation, though we do study computable martingales in Section 4.1.

Definition 2.5. A real α is *left-c.e.* (also called *left computable* and *left semicomputable*) if there is a computable strictly increasing sequence of rationals $(q_s)_{s\in\mathbb{N}}$ such that $\lim_s q_s = \alpha$. One says that α is *computable* if there is a computable sequence of rationals $(q_s)_{s\in\mathbb{N}}$, called a *Cauchy name*, such that $|q_n - \alpha| < 2^{-n}$ for all $n \in \mathbb{N}$. A function $f : \mathbb{R} \to \mathbb{R}$ is *computable* if

1. $f(q)$ is a computable real uniformly in a rational q, and
2. f is effectively uniformly continuous; i.e., there is a computable function $h : \mathbb{N} \to \mathbb{N}$ such that $|f(x) - f(y)| < 2^{-n}$ whenever $|x - y| < 2^{-h(n)}$.

A classic result in computable analysis asserts that computability in the sense of Definition 2.5 is equivalent to the *effective Weierstraß condition*, which states that there is a computable sequence $(p_s)_{s\in\mathbb{N}}$ of rational polynomials or polygonal functions such that $||f - p_s||_\infty < 2^{-s}$ for all s.

Finally, we define the *slope* of f at a pair a, b of distinct rationals as the quotient

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

In Section 7 we proceed to the type two theory of effectivity, where the previous notions are situated in a more general framework.

3 Effective Jordan decomposability

The principal concern of this section is the effective status of the Jordan decomposition theorem. Given a computable real-valued function f, is it always possible to decompose it into two non-decreasing computable functions? This question is important because an affirmative answer immediately characterises the complexity of the theorem in terms of reverse mathematics. On the other hand, a negative answer may prompt the rejection of Jordan's theorem by various constructively inclined schools of mathematics, such as the Markov school or followers of Bishop's programme.

Indeed, Bridges [6] has shown that the statement *the difference of two non-decreasing functions on* $[0, 1]$ *has a variation* entails the *limited principle of omniscience*, which does not hold when restricted to an intuitionistic logic. Thus, intuitionistically at least, one may construct a function of bounded variation with no (constructible) variation function. We study the computable analogue of this phenomenon after characterising the left-c.e. reals as the variation functions of computable functions.

Theorem 3.1 (Zheng and Rettinger [30])**.** *For any left-c.e. real* $\alpha \in [0, 1]$*, there is a computable function* f *such that* $\mathbf{V}(f) = \alpha$*.*

Proof. We define f as the sum of sawtooth functions. Let $(q_n)_{n\in\mathbb{N}}$ be a computable sequence of rationals converging to α such that $q_0 = 0$ and $q_{n+1} > q_n$. On each interval $[q_n, q_{n+1}]$ we will specify f to be a sawtooth function such that $\mathbf{V}(f, [q_n, q_{n+1}]) = q_{n+1} - q_n$.

Define a computable function $r : \mathbb{N} \to \mathbb{N}$ by

$$r(n) = \mu m > 0 \left[\frac{q_{n+1} - q_n}{2m} < 2^{-n} \right].$$

Let $(h_n)_{n\in\mathbb{N}}$ be the sequence in $\mathbb{R}$ given by $h_n = [(q_{n+1} - q_n)/2r(n)]$. For all $n \in \mathbb{N}$ put $I_n = [q_n, q_{n+1}]$. Recall that a sawtooth function $\mathrm{M}_A(h, r)$ has r-many teeth of height h. Define $f : [0, 1] \to \mathbb{R}$ by

$$f = \sum_{n=0}^{\infty} \mathrm{M}_{I_n}(h_n, 2r(n)).$$

Then $f(x) = 0$ for all $x \in [\alpha, 1]$. Each interval I_n contributes precisely $q_{n+1} - q_n$ to the variation of f. Hence

$$\mathbf{V}(f) = \sum_{n=0}^{\infty} \mathbf{V}(f, I_n) + \mathbf{V}(f, [\alpha, 1]) = \sum_{n=0}^{\infty} (q_{n+1} - q_n) = \alpha.$$

So f has total variation α. It remains to show that f is computable. To see this, consider the sequence of polygonal functions $p_n : [0,1] \to \mathbb{R}$ defined by

$$p_n(x) = \begin{cases} f(x) & \text{if } x \in [0, q_{n+1}], \\ 0 & \text{otherwise.} \end{cases}$$

Then for any $n \in \mathbb{N}$ one has $||f - p_n||_\infty \leq h_n < 2^{-n}$, as required by the effective Weierstraß condition. □

The construction works by coding the real α into the variation of f. Thus we may put, for instance, $\alpha = \sum_{n \in \emptyset'} 2^{-n}$ and obtain via Theorem 3.1 a computable function f such that $\mathbf{V}(f) \equiv_T \emptyset'$.

Definition 3.2. Let $f : [0,1] \to \mathbb{R}$ be a computable function. An *effective Jordan decomposition* of f is a pair of non-decreasing computable functions (f^+, f^-) such that $f = f^+ - f^-$. The function f is *effectively Jordan decomposable* if it has an effective Jordan decomposition.

The notion of effective Jordan decomposability is the effectivisation of regular decomposability: we now require that any function which contributes to the decomposition must be computable. The result of Theorem 3.4 below shows that not every computable function of bounded variation has an *effective* Jordan decomposition. Thus Jordan's theorem cannot hold in the minimal ω-model of RCA_0, which consists entirely of computable sets and functions. This leads us to the investigation of the reverse mathematical complexity of Jordan's theorem in Section 5.

Lemma 3.3. *Let f be an effectively Jordan decomposable function. Then $\boldsymbol{v}_f$ has a computable modulus of uniform continuity.*

Proof. Let (f^+, f^-) be an effective Jordan decomposition of f, and let h_1 and h_2 be their respective computable moduli of uniform continuity. Then the computable function $h : \mathbb{N} \to \mathbb{N}$ defined by $h(n) = \max(h_1(n+1), h_2(n+1))$ is a modulus of uniform continuity for f. To see this, observe that for any $x < y$ with $|x - y| < 2^{-h(n)}$ one has

$$\begin{aligned} |\boldsymbol{v}_f(y) - \boldsymbol{v}_f(x)| &= \mathbf{V}(f, [x,y]) \\ &= \mathbf{V}(f^+ - f^-, [x,y]) \\ &\leq \mathbf{V}(f^+, [x,y]) + \mathbf{V}(f^-, [x,y]) \\ &= (f^+(y) - f^+(x)) + (f^-(y) - f^-(x)) \\ &\leq 2^{-n}. \end{aligned}$$

□

Theorem 3.4 (Zheng and Rettinger [30])**.** *There is a computable function f of bounded variation which is not effectively Jordan decomposable.*

Proof. Let A be an incomputable c.e. set. Fix an effective enumeration $(a_s)_{s\in\mathbb{N}}$ of A without repetitions, and let $(m_s)_{s\in\mathbb{N}}$ and $(t_s)_{s\in\mathbb{N}}$ be defined by $m_s = \max\{a_t+s+1 : t \leq s\}$ and $t_s = m_s - (a_s+1)$, respectively. Then for all s one has $t_s \geq 0$ and $m_s < m_{s+1}$. Now take f to be the sawtooth function of type $[0,1] \to \mathbb{R}$ defined by

$$f = \sum_{s=0}^{\infty} \mathrm{M}_{[2^{-(s+1)},2^{-s}]}(2^{-m_s}, 2^{t_s}). \tag{1}$$

We will show that f is a computable function of bounded variation, but that $\boldsymbol{v}_f$ does not have a computable modulus of uniform continuity, and hence by Lemma 3.3 f cannot be effectively Jordan decomposable.

Define a sequence $(p_s)_{s\in\mathbb{N}}$ of polygonal functions by putting $p_s(x) = f(x)$ if $x \in [2^{-(s+1)}, 1]$, and $p_s(x) = 0$ otherwise. Then, since $m_s > s$,

$$\begin{aligned} \sup\{|f(x)-p_s(x)| : x \in [0,1]\} &= \sup\{|f(x)-p_s(x)| : x \in [0,2^{-(s+1)}]\} \\ &\leq 2^{-m_s} \\ &\leq 2^{-s}. \end{aligned}$$

Hence $||f-p_s||_\infty \leq 2^{-s}$, so f is computable by the effective Weierstraß condition. Now we verify that f is of bounded variation. Note that on the interval $[2^{-s}, 2^{-(s+1)}]$ the function f is equal to the sawtooth function with 2^{t_s}-many teeth of height 2^{-m_s}. Thus

$$\mathbf{V}(f) = \sum_{s=0}^{\infty} \mathbf{V}(f, [2^{-s}, 2^{-(s+1)}]) = \sum_{s=0}^{\infty} 2^{t_s} 2^{-m_s} 2 = \sum_{s=0}^{\infty} 2^{-a_s} < \infty.$$

Finally, assume for contradiction that $h : \mathbb{N} \to \mathbb{N}$ is a computable modulus of uniform continuity for $\boldsymbol{v}_f$. Then

$$\sum_{t \geq h(s)} 2^{-a_t} = \mathbf{V}(f, [0, 2^{-h(s)}]) = |\boldsymbol{v}_f(0) - \boldsymbol{v}_f(2^{-h(s)})| \leq 2^{-s}.$$

Hence if $t \geq h(s)$ then $a_t < s$. In particular, $n \in A \Leftrightarrow n \in A_{h(n)}$, which contradicts the assumption that A is not computable. □

Our choice of notation in equation (1) greatly simplifies the construction of Zheng and Rettinger, and enables the subsequent verification to proceed easily.

The subrecursive case of Jordan decomposition was explored in Ko [15]. A function is said to be *polynomial time Jordan decomposable* if it is effectively Jordan decomposable and the functions comprising the decomposition are polynomial time computable. In line with the results of this section, Ko has shown that there is a polynomial time computable function of bounded variation which is not polynomial time Jordan decomposable [15]. Zheng and Rettinger were able to strengthen this by constructing a polynomial time computable function which is not the difference of two *computable* functions, and even has a polynomial time computable modulus of absolute continuity.

4 Characterising points of differentiability with randomness

In this section our efforts are devoted to characterising the reals which are simultaneously points of differentiability for all computable functions of bounded variation. We begin with the subclass consisting of Lipschitz functions before proceeding to the general case. The table in Section 4.3 below summarises the known correspondences between points of differentiability and the randomness notions they characterise, which extend the presented results. On our way to the result for Lipschitz functions, in Theorem 4.3 we provide a characterisation of their variation functions in terms of *interval-c.e.* functions.

4.1 Differentiability of computable Lipschitz functions

Computable martingales. We note some preliminaries on martingales and computable randomness. See [19, Chapter 7] or [10, Section 6.3] for a comprehensive exposition.

Definition 4.1. A function $\mathbf{M} : 2^{<\mathbb{N}} \to [0, \infty)$ is a *martingale* if it satisfies the fairness condition $\mathbf{M}(\sigma 0) + \mathbf{M}(\sigma 1) = 2\mathbf{M}(\sigma)$ for all $\sigma \in 2^{<\mathbb{N}}$. $\mathbf{M}$ is *computable* if $\mathbf{M}(\sigma)$ is a computable real uniformly in σ. We say that $\mathbf{M}$ *succeeds* on a sequence $Z \in 2^{<\mathbb{N}}$ if $\sup_n \mathbf{M}(Z|_n) = \infty$.

Classically, a martingale is defined as a stochastic process with respect to a decreasing sequence of σ-algebras, and is intended to model the fortune of a gambler playing a fair game. By the optional stopping theorem, discrete-time martingales have the property that, on average, the expected fortune of a gambler is no more than when the gambler began playing. In the context of algorithmic randomness we use martingales to formalize the notion of a betting strategy, see the introduction to computable randomness below. With a martingale $\mathbf{M}$ we may construct a Borel measure $\mu_{\mathbf{M}}$ by specifying

$$\mu_{\mathbf{M}}[0.\sigma, 0.\sigma + 2^{-|\sigma|}) = \mathbf{M}(\sigma)2^{-|\sigma|}.$$

Since the set of all intervals of the form $[0.\sigma, 0.\sigma + 2^{-|\sigma|})$ is a semiring on [0,1], by an application of Carathéodory's extension theorem we can assume that $\mu_{\mathbf{M}}$ is unique and complete.

Given a martingale $\mathbf{M}$, we define the associated *cumulative distribution function* $\mathsf{cdf}_{\mathbf{M}} : [0, 1] \to \mathbb{R}$ by

$$\mathsf{cdf}_{\mathbf{M}}(x) = \mu_{\mathbf{M}}[0, x).$$

Definition 4.2. A function $\mathbf{L} : 2^{<\mathbb{N}} \to \mathbb{R}$ is a *signed martingale* if it satisfies the fairness condition of Definition 4.1.

From a signed martingale $\mathbf{L}$ we define the *variation martingale* $\mathbf{V_L} : 2^{<\mathbb{N}} \to [0, \infty]$ by

$$\mathbf{V_L}(\sigma) = \sup_k 2^{-k} \sum_{|\eta|=k} |\mathbf{L}(\sigma\eta)|.$$

This is an analogue of the variation of a real valued function, and itself defines a martingale.

The variation of a computable Lipschitz function. We say that a non decreasing function $f : [0,1] \to \mathbb{R}$ is *interval-c.e.* if $f(0) = 0$ and $f(y) - f(x)$ is a left-c.e. real uniformly in rationals $x < y$. Theorem 4.3 below shows that interval-c.e. functions are essentially the variation functions of computable Lipschitz functions.

Theorem 4.3 (Freer et al. [11]). *Let f be an interval-c.e. function with Lipschitz constant κ. Then there is a computable κ-Lipschitz function g such that $f = \boldsymbol{v}_g$.*

Proof. Define the martingale $\mathbf{M}$ by $\mathbf{M}(\sigma) = S_f(0.\sigma, 0.\sigma + 2^{-|\sigma|})$. There is a computable signed martingale $\mathbf{L}$ such that $\mathbf{V_L} = \mathbf{M}$ and $|\mathbf{L}| \leq \mathbf{M}$ (see [11], Lemma 3.3). Thus $|\mathbf{L}| \leq S_f(0.\sigma, 0.\sigma + 2^{-|\sigma|}) \leq \kappa$.

We define a Lipschitz function $g : [0,1] \to \mathbb{R}$ by specifying its values at dyadic rationals, which is sufficient since the set of dyadic rationals is dense in $[0,1]$. For $\sigma \in 2^{<\mathbb{N}}$ define

$$g(0.\sigma) = 2^{-|\sigma|} \sum \{\mathbf{L}(\tau) : 0.\tau < 0.\sigma \wedge |\tau| = |\sigma|\}.$$

Let $\sigma, \eta \in 2^{<\mathbb{N}}$ be of length n with $0.\eta < 0.\sigma$. Then

$$\begin{aligned} |g(0.\sigma) - g(0.\eta)| &= 2^{-n} \Big| \sum_{|\tau|=n} \{\mathbf{L}(\tau) : 0.\tau < 0.\sigma\} - \sum_{|\tau|=n} \{\mathbf{L}(\tau) : 0.\tau < 0.\eta\} \Big| \\ &\leq 2^{-n} \sum_{|\tau|=n} \{|\mathbf{L}(\tau)| : 0.\eta \leq 0.\tau < 0.\sigma\} \\ &\leq \kappa |0.\sigma - 0.\eta|, \end{aligned}$$

the last inequality holding since there are as many as 2^n binary strings of length n. Thus g is κ-Lipschitz on the set of dyadic rationals, and so defines a total computable κ-Lipschitz function $\widehat{g}(x) = \sup\{g(0.\sigma) : \sigma \in 2^{<\mathbb{N}} \wedge 0.\sigma \leq x\}$. For ease of notation we will suppress the hat overscript and simply write g.

Next we show that $f = \boldsymbol{v}_g$. Since f and g are continuous it suffices to show this for dyadic rationals. For $\sigma \in 2^{<\mathbb{N}}$ of length n we have

$$\begin{aligned} \boldsymbol{v}_g(0.\sigma) &= 2^{-n} \sum \{\mathbf{V_L}(\tau) : 0.\tau < 0.\sigma \wedge |\tau| = n\} \\ &= 2^{-n} \sum \{\mathbf{M}(\tau) : 0.\tau < 0.\sigma \wedge |\tau| = n\} \\ &= 2^{-n} \sum \{S_f(0.\tau, 0.\tau + 2^{-|\tau|}) : 0.\tau < 0.\sigma \wedge |\tau| = n\} \\ &= f(0.\sigma). \end{aligned}$$

□

With Rute, Freer et al. [11] extended Theorem 4.3 to continuous interval-c.e. functions. In particular, they showed that every continuous interval-c.e. function is the variation function of a computable function.

Lipschitz functions and computable randomness. Recall that a sequence $Z \in 2^{\mathbb{N}}$ is *computably random* if no computable martingale succeeds on Z, and a real $z \in [0,1]$ is computably random if its binary expansion is computably random. Schnorr introduced this randomness notion to capture the *unpredictability* paradigm in randomness; i.e., that a random sequence should have unpredictable bits, so that betting on which bits occur in such a sequence cannot yield unbounded profit.

There is an effective analogue to Lebesgue's differentiation theorem: a real z is computably random iff every computable non-decreasing function is differentiable at z [5]. A crucial component of the proof, the following theorem asserts that for functions computable on the rational unit interval, just being non-decreasing on the rationals is sufficient to be differentiable at computably randoms.

Theorem 4.4 (Brattka et al. [5]). *Let $f : [0,1] \to \mathbb{R}$ be a function computable on $[0,1]_{\mathbb{Q}}$. If f is non-decreasing on $[0,1]_{\mathbb{Q}}$, then f is differentiable at each computably random real z.*

Theorem 4.4 allows us to extend the effective Lebesgue differentiation theorem to Lipschitz functions. The characterisation of reals which are simultaneously points of differentiability for all computable Lipschitz functions aligns with the characterisation for non-decreasing functions.

Theorem 4.5 (Freer et al. [11]). *Let $z \in [0,1]$. Then z is computably random iff every computable Lipschitz function is differentiable at z.*

Proof. $\Rightarrow$: Let f be a computable function with Lipschitz constant κ. Then the function $g : [0,1] \to \mathbb{R}$ defined by $g(x) = f(x) + \kappa x$ is computable and non-decreasing. Hence, by Theorem 4.4, g is differentiable at z. Thus f is differentiable at z.

$\Leftarrow$: By contrapositive. Let $\mathbf{M}$ be a computable martingale which succeeds on (the binary expansion of) z. We will construct a bounded computable Lipschitz martingale $\mathbf{B}$ such that $\mathsf{cdf}_{\mathbf{B}}$ is not differentiable at z.

By a result of Schnorr [24] we may assume that $\mathbf{M}$ takes only rational values, and by [5, Proposition 3.4] we may also assume that $\mathbf{M}$ has the savings property; i.e., that $\mathbf{M}(\sigma\tau) \geq \mathbf{M}(\sigma) - 1$ for all strings σ, τ. We will construct $\mathbf{B}$ inductively by considering its value at each string σ, and defining $\mathbf{B}(\sigma 0)$ and $\mathbf{B}(\sigma 1)$ accordingly. For every σ, the martingale $\mathbf{B}$ will be in either the *up phase* or the *down phase*. In the up phase we will ensure that $\mathbf{B}(\sigma) < 3$, and in down phase we will ensure that $\mathbf{B}(\sigma) > 2$. This property will be the inductive condition.

Construction 4.6. Put $\mathbf{B}(\varnothing) = 2$ and declare $\mathbf{B}(\varnothing)$ to be in the up phase. Then $\mathbf{B}(\varnothing)$ satisfies the inductive condition. Now assume that $\mathbf{B}(\sigma)$ has been defined.

Case 1. $\mathbf{B}(\sigma)$ is in the up phase. For $k \in \{0,1\}$ let $r_k = \mathbf{B}(\sigma) + \mathbf{M}(\sigma k) - \mathbf{M}(\sigma)$. If $r_0, r_1 < 3$ let $\mathbf{B}(\sigma k) = r_k$; declare $\mathbf{B}(\sigma k)$ to be in the up phase.

If not, then there is a unique k such that $r_k \geq 3$. To see this, notice that otherwise one has

$$\mathbf{B}(\sigma) + \mathbf{M}(\sigma 0) - \mathbf{M}(\sigma) \geq 3, \text{ and}$$
$$\mathbf{B}(\sigma) + \mathbf{M}(\sigma 1) - \mathbf{M}(\sigma) \geq 3,$$

so that $\mathbf{M}(\sigma) < \mathbf{M}(\sigma 0)$ and $\mathbf{M}(\sigma) < \mathbf{M}(\sigma 1)$. Hence $2\mathbf{M}(\sigma) < \mathbf{M}(\sigma 0) + \mathbf{M}(\sigma 1)$, which contradicts the fairness condition for martingales.

For this k let $\mathbf{B}(\sigma k) = 3$ and declare $\mathbf{B}(\sigma k)$ to be in the down phase. Let $\mathbf{B}(\sigma(1 - k)) = 2\mathbf{B}(\sigma) - 3$ and declare $\mathbf{B}(\sigma(1 - k))$ to be in the up phase. Clearly the inductive condition holds for both $\mathbf{B}(\sigma 0)$ and $\mathbf{B}(\sigma 1)$.

Case 2. $\mathbf{B}(\sigma)$ is in the down phase. For $k \in \{0, 1\}$ let $r_k = \mathbf{B}(\sigma) - (\mathbf{M}(\sigma k) - \mathbf{M}(\sigma))$. If $r_0, r_1 > 2$ let $\mathbf{B}(\sigma k) = r_k$ and declare $\mathbf{B}(\sigma k)$ to stay in the down phase. Otherwise, by an argument similar to the first case, there is a unique k such that $r_k \leq 2$. Let $\mathbf{B}(\sigma k) = 2$ and declare $\mathbf{B}(\sigma k)$ to be in the up phase. Let $\mathbf{B}(\sigma(1 - k)) = 2\mathbf{B}(\sigma) - 2$ and stay in the down phase. The construction preserves the inductive condition for $\sigma 0$ and $\sigma 1$. *This completes the construction.*

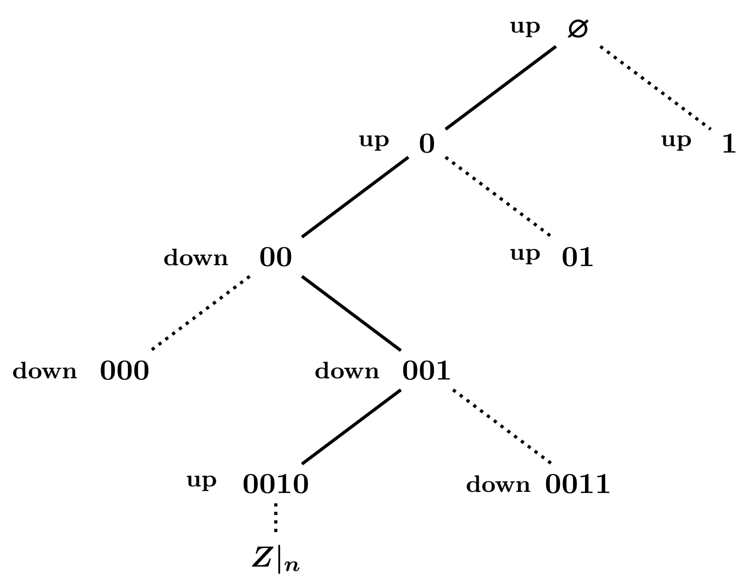


Figure 1: An illustration of the inductive construction of the martingale $\mathbf{B}$. There is unbounded capital to use on the construction of $\mathbf{B}(Z|_n)$ since $\mathbf{M}$ succeeds on the binary expansion of z.

Claim 4.7. $\mathsf{cdf}_{\mathbf{B}}$ *is Lipschitz.*

Proof. We first show that for each string η we have $1 \leq \mathbf{B}(\eta) \leq 4$. Suppose that $\mathbf{B}$ is in the up phase at η. Then our inductive condition ensures that $\mathbf{B}(\eta) < 3$. Moreover, if $\mathbf{B}$ entered the up phase from $\sigma \prec \eta$ with $|\sigma|$ greatest, then $\mathbf{B}(\sigma) = 2$. By the savings property $\mathbf{M}(\eta) - \mathbf{M}(\sigma) \geq -1$. Thus $\mathbf{B}(\eta) = \mathbf{B}(\sigma) + \mathbf{M}(\eta) - \mathbf{M}(\sigma) \geq 1$.

Now suppose that $\mathbf{B}$ is in the down phase at η. Again, the inductive condition ensures that $\mathbf{B}(\eta) > 2$. If $\mathbf{B}$ entered the up phase from $\sigma \prec \eta$ with $|\sigma|$ greatest, then $\mathbf{B}(\sigma) = 3$. An application of the savings property shows that $\mathbf{B}(\eta) = \mathbf{B}(\sigma) - (\mathbf{M}(\eta) - \mathbf{M}(\sigma)) \leq 4$.

Next we show that $\mathsf{cdf}_{\mathbf{B}}$ has Lipschitz constant 4. Let $x, y \in [0,1]$ with $x < y$. Let $n \in \mathbb{N}$. Put $i = \lfloor x2^n \rfloor$ and $j = \lceil y2^n \rceil$. Then $i2^{-n} \leq x < (i+1)2^{-n}$ and $(j-1)2^{-n} < y \leq j2^{-n}$. Hence,

$$\begin{aligned}\mathsf{cdf}_{\mathbf{B}}(y) - \mathsf{cdf}_{\mathbf{B}}(x) = \mu_{\mathbf{B}}[x,y) &\leq \sum_{r=i}^{j-1} \mu_{\mathbf{B}}[r2^{-n}, (r+1)2^{-n}) \\ &\leq 4((j-1)-i)2^{-n} \\ &= 4((j-1)2^{-n} - (i+1)2^{-n} + 2^{-n}) \\ &< 4(y - x + 2^{-n}).\end{aligned}$$

A similar proof shows that $\mathsf{cdf}_{\mathbf{B}}(y) - \mathsf{cdf}_{\mathbf{B}}(x) \geq 1$. This proves the claim. ◊

Claim 4.8. $\mathsf{cdf}_{\mathbf{B}}$ *is not differentiable at* z.

Proof. Let Z be the binary expansion of z. For any $\sigma \in 2^{<\mathbb{N}}$, one has $\mathbf{B}(\sigma) = S_{\mathsf{cdf}_{\mathbf{B}}}(0.\sigma, 0.\sigma + 2^{-|\sigma|})$. Since $\mathbf{M}$ succeeds on Z, there is unbounded capital to spend in the construction of $\mathbf{B}$ for prefixes of Z. Thus as $n \to \infty$, $\mathbf{B}(Z|_n)$ oscillates between up phases and down phases; i.e., between values ≤ 2 and ≥ 3. Hence $\mathsf{cdf}_{\mathbf{B}}$ cannot be differentiable at z. ◊

□

4.2 Differentiability of computable functions of bounded variation

For an interval $A \subseteq [0,1]$ with endpoints a and b, we denote by $|A|$ the length $b - a$ of A. If $p \in \mathbb{N}$, let $\Lambda_{A,p}$ denote the sawtooth function which is 0 outside of A, reaches $p|A|/2$ at the middle point of A, and is linearly interpolated elsewhere.

We work with the usual topology on Cantor space $2^{\mathbb{N}}$ generated by the collection of basic open cylinders

$$[\sigma] = \{Z \in 2^{\mathbb{N}} : \sigma \prec Z\}$$

for each $\sigma \in 2^{<\mathbb{N}}$. Let λ denote the product measure on $2^{\mathbb{N}}$; i.e., for basic cylinders $\lambda[\sigma] = 2^{-|\sigma|}$, and if $U = \bigcup_{i=0}^{\infty}[\sigma_i]$ where $\{\sigma_i : i \in \mathbb{N}\}$ is prefix-free, then

$$\lambda U = \sum_{i=0}^{\infty} \lambda[\sigma_i].$$

If $U \subseteq 2^{<\mathbb{N}}$ we also define $[\![U]\!] = \{Z \in 2^{\mathbb{N}} : \sigma \prec Z \text{ for some } \sigma \in U\}$.

A *Martin-Löf test* is a uniformly computably enumerable sequence $(\mathcal{G}_m)_{m\in\mathbb{N}}$ of open sets such that $\lambda\mathcal{G}_m \leq 2^{-m}$ for every m. A sequence $Z \in 2^{\mathbb{N}}$ is *Martin-Löf random* if $Z \notin \bigcap_m \mathcal{G}_m$ for every Martin-Löf test $(\mathcal{G}_m)_{m\in\mathbb{N}}$. There is a *universal* Martin-Löf test $(\mathcal{U}_m)_{m\in\mathbb{N}}$, whereby Z is Martin-Löf random if and only if $Z \notin \bigcap_m \mathcal{U}_m$.

In the proof of Theorem 4.4 we turned a martingale which succeeded on a sequence into a Lipschitz function which could not be differentiated at the corresponding computably random real; this showed that the differentiability of all computable Lipschitz functions at some real z implies that z is computably random. Now, the existence of a universal test allows the construction of a *single* function f whose points of differentiability are exactly the Martin-Löf random reals.

The construction was first completed by Demuth [8] according to the constructivist paradigm. We use the construction due to Brattka, Miller, and Nies [5], which recasts Demuth's result in the modern framework of computable analysis and algorithmic randomness. Given a universal Martin-Löf test $(\mathcal{G}_m)_{m\in\mathbb{N}}$ such that $\mathcal{G}_{m+1} \subseteq \mathcal{G}_m$ and $\lambda\mathcal{G}_m \leq 8^{-m}$ for all m, one divides each $\mathcal{G}_m$ into a computable double sequence $(C_{m,i})_{m,i\in\mathbb{N}}$ of disjoint open intervals. On each $\mathcal{G}_m$, one defines the function f_m as the infinite sum of sawtooth functions, one tooth per $C_{m,i}$:

$$f_m = \sum_{i=0}^{\infty} \Lambda_{C_{m,i},4^m}.$$

The construction is illustrated in Figure 2. We take $f = \sum_{m=0}^{\infty} f_m$.

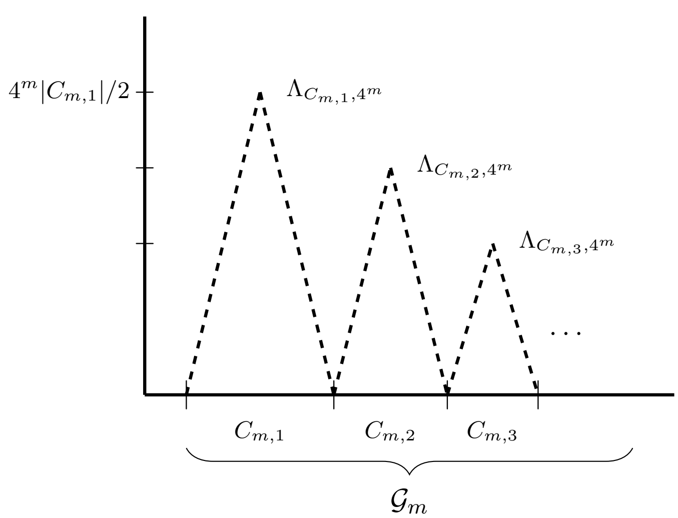


Figure 2: Construction of f_m on the open set $\mathcal{G}_m$.

The measure $\lambda\mathcal{G}_m$ approaches 0 as $m \to \infty$, and the intervals $C_{m,i}$ are constructed so that this fact ensures their length decreases to 0 at a rate which makes the sum of the heights of the sawteeth $\Lambda_{C_{m,i},4^m}$ converge. This guarantees that $f(x)$ is defined for each $x \in [0,1]$. A number of careful estimates shows also that one may compute $f(q)$ uniformly in a rational q.

Our constructed function f is absolutely continuous: for an interval $A \subseteq [0,1]$ and $p \in \mathbb{N}$, let $\Theta_{A,p}$ be the function that is undefined at the endpoints and middle points of A, takes the value p on the left half of A, and takes the value $-p$ on the right half of A. Then

$$\int_0^x \Theta_{A,p} d\lambda = \Lambda_{A,p}(x).$$

For each $m \in \mathbb{N}$ define $g_m : [0,1] \to \mathbb{R}$ by $g_m = \sum_{i=0}^{\infty} \Theta_{C_{m,i},4^m}$. The construction of g_m is illustrated in Figure 3.

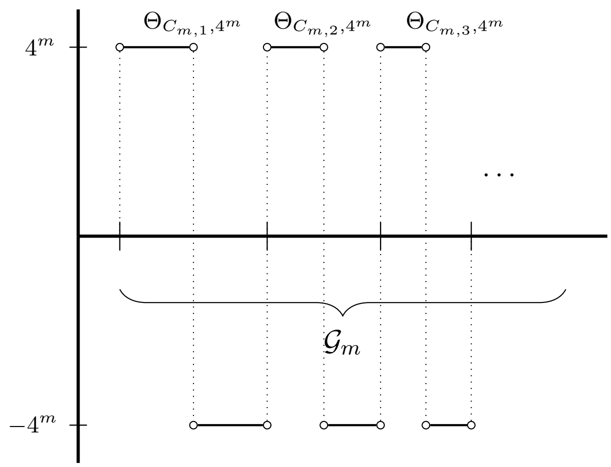


Figure 3: Construction of g_m on the open set $\mathcal{G}_m$.

We show that g_m is integrable. Note that $|g_m| = 4^m \mathbb{1}_{\mathcal{G}_m}$ a.e. since $\bigsqcup_i C_{m,i} = \mathcal{G}_m$, so

$$\int |g_m| d\lambda = 4^m \lambda \mathcal{G}_m \leq 4^m 8^{-m} \leq 2^{-m}.$$

Hence $\sum_{m=0}^{\infty} \int |g_m| d\lambda \leq 2$. By the Lebesgue dominated convergence theorem $\sum_{m=0}^{\infty} g_m$ is finite almost everywhere [3, Theorem 2.8.1]. Define $g : [0,1] \to \mathbb{R} \cup \{\infty\}$ by $g(x) = \sum_{m=0}^{\infty} g_m(x)$. Then g is integrable and $\int_0^x g d\lambda = \sum_m \int_0^x g_m d\lambda$. Thus since $f_m(x) = \int_0^x g_m d\lambda$, one has $f(x) = \int_0^x g d\lambda$. Hence, by [3, Theorem 5.3.6], f is absolutely continuous. In particular, f is of bounded variation.

Theorem 4.9 (Brattka et al. [5])**.** *Let $z \in [0,1]$. Then z is Martin-Löf random iff every computable function of bounded variation is differentiable at z.*

Proof. $\Leftarrow$: Any real which is not Martin-Löf random is not a point of differentiability for the function f that we constructed above.

$\Rightarrow$: Let $f : [0,1] \to \mathbb{R}$ be a computable function of bounded variation. By the uniform computability condition for computable real functions, there is a computable function $\varphi : \mathbb{Q} \times \mathbb{N} \to \mathbb{Q}$ such that for all n, $|f(q) - \varphi(q,n)| \leq 2^{-n}$. Without loss of generality we may assume that both f and $\mathbf{V}(f|_{\mathbb{Q}})$ are bounded by 1.

Let $\Sigma = \{-1, 0, 1\}$, and let $(x_n)_{n\in\mathbb{N}}$ be an effective listing of $[0,1]_{\mathbb{Q}}$. There is a surjective function δ from Σ^ω to the set of all functions of type $[0,1]_{\mathbb{Q}} \to [0,1]$, called a *representation*, such that $\delta(\zeta) = h$ iff $\zeta\langle x_k, n\rangle$ is the nth term in a Cauchy name of $h(x_k)$. In this case we call ζ a *name* of h. This system of representations underlies the type two theory of effectivity, which we develop in detail in Section 7; for now this remark will suffice. Let $\hat{\zeta} = \{\langle\zeta_0, \zeta_1\rangle : \zeta_i$ names a non-decreasing function $f_i : [0,1]_{\mathbb{Q}} \to [0,1]$ for $i = 0, 1$, and $f(q) = f_0(q) - f_1(q)$ for all $q \in [0,1]_{\mathbb{Q}}\}$. Then to determine whether $\langle\zeta_0, \zeta_1\rangle \in \hat{\zeta}$, one must verify that

$$\forall k, l, n\big[x_k < x_l \to (\zeta_0\langle x_k, n\rangle \leq \zeta_0\langle x_l, n\rangle + 2^{-n} \ \wedge\ \zeta_1\langle x_k, n\rangle \leq \zeta_1\langle x_l, n\rangle + 2^{-n})\big],$$

and

$$\forall k, n\big|\varphi(x_k, n) - (\zeta_0\langle x_k, n\rangle - \zeta_1\langle x_k, n\rangle)\big| \leq 2^{-n}.$$

Hence $\hat{\zeta}$ is a Π^0_1 class.

By the low for z basis theorem [9, Proposition 7.4], z is Martin-Löf random (and therefore computably random) relative to some z-c.e. pair of names $\langle\zeta_0, \zeta_1\rangle \in \hat{\zeta}$. Relativizing Theorem 4.4 to $\delta(\zeta_0) \oplus \delta(\zeta_1)$ we have that $\delta(\zeta_0)$ and $\delta(\zeta_1)$ are differentiable at z. Thus since $f = \delta(\zeta_0) - \delta(\zeta_1)$, f is also differentiable at z.

□

The proof of Theorem 4.9 used the language of representations. In Section 5.3 we prove once more that the set of Jordan decompositions for a function is a Π^0_1 class, though in the framework of second-order arithmetic. As we will see, the differences between the proofs are conspicuous, and quite a lot more work must be done to achieve a similar outcome. For instance, we must explicitly construct an infinite computable binary tree corresponding to the desired Π^0_1 class, whose paths encode the components of a Jordan decomposition.

4.3 Other classes of a.e. differentiable functions

Table 1 lists known correspondences between randomness notions and classes of effective functions. Two such correspondences rely on a generalisation of variation. For a real $p > 1$, the *p-variation* $\mathbf{V}_p(f, I)$ of f in the interval I is

$$\mathbf{V}_p(f, I) = \sup \sum_{i=0}^{n-1} \frac{|f(t_{i+1}) - f(t_i)|^p}{|t_{i+1} - t_i|^{p-1}},$$

where the supremum is taken over all collections $t_0 < \cdots < t_n$ in I. Similar to the variation norm from Section 2, one defines the *p-variation norm* $||\cdot||_{\mathbf{BV}_p}$ by

$$||f||_{\mathbf{BV}_p} = |f(0)| + (\mathbf{V}_p(f, [0,1]))^{\frac{1}{p}}.$$

Class of effective functions	Randomness notion	
polynomial time computable	polynomial time randomness	[20]
computable and non-decreasing computable and Lipschitz	computable randomness	[11]
interval c.e.	density randomness	[18]
Lipschitz and p-variation computable p-variation computable with finite p-variation norm	Schnorr randomness	[11]
computable and bounded variation computable and absolutely continuous	Martin-Löf randomness	[5]
computable and a.e. differentiable in $[0,1]^n$	weak 2-randomness in $[0,1]^n$	[5]

Table 1: Correspondences between effective functions and randomness notions. A real satisfies a given randomness notion iff it is a point of differentiability for every function in the corresponding class.

Using the p-variation norm one can obtain a characterisation of Schnorr randoms in terms of differentiability. By [11, Corollary 5.2], a real z is Schnorr random iff every Lipschitz function computable in the p-variation norm is differentiable at z.

It was also shown recently by Nies in [20] that a real z is a point of differentiability for every polynomial-time computable function iff no polynomial-time computable martingale succeeds on the binary expansion of z. This is a refinement of the effective Lebesgue differentiation theorem for the case of feasible analysis. Rute in [23] has given several other pertinent characterisations. Among other things, he showed that if g is a computable function of bounded variation, $\mathbf{V}(g)$ is a computable real, and g' is computable in the $\mathcal{L}_1$ norm, then g is differentiable at all Schnorr randoms [23, Corollary 9.20].

A set X is said to be *high* if $X' \geq_T \emptyset''$. By [19, Theorem 7.5.9], for any high set C there is a computably random real z such that $z \equiv_T C$. Thus any high set computes a real which is simultaneously a point of differentiability for every non-decreasing computable function. Similarly, for any high set C there is a Schnorr random real $z \equiv_T C$ such that z is not computably random [19, Theorem 7.5.10]. Thus any high set also computes a point of differentiability for every Lipschitz function which is computable in the p-variation norm.

The results suggest a certain kind of equivalence in computational complexity between differentiating computable functions and asserting that random reals exist. We use this idea in Section 6 to establish a principle in reverse mathematics stating just this equivalence, relating functions of bounded variation and Martin-Löf randomness.

5 Reverse mathematics of the Jordan decomposition theorem

The content of this section, together with the content of Section 6, was initiated by the preliminary results in the 2013 Logic Blog [21]. This dissertation completes and elaborates on some of these preliminary results, which we present momentarily. We first proceed quickly through the necessary background on second-order arithmetic.

5.1 Subsystems of second-order arithmetic

We shall introduce the reverse mathematics framework, using mostly the conventions due to Simpson [25]. The *language $\mathcal{L}_2$ of second-order arithmetic* is a two-sorted first-order language. The variables are divided into (i) *number variables* $i, j, k, m, n, p, \ldots$ which are intended to range over $\mathbb{N}$; and (ii) *set variables* $X, Y, Z, \ldots$ which are intended to range over subsets of $\mathbb{N}$. *Numerical terms* are number variables, constant variables from the set $\{0, 1\}$, and $t_1 + t_2$ and $t_1 \cdot t_2$ whenever t_1 and t_2 are numerical terms. The *atomic formulas* are all expressions $t_1 = t_2$, $t_1 < t_2$, $t_1 \in X$, where t_1 and t_2 are numerical terms and X is any set variable. $\mathcal{L}_2$-*formulas* (often simply called *formulas*) are generated from atomic formulas in the usual way with the connectives $\neg, \wedge, \vee, \rightarrow, \leftrightarrow$, the number quantifiers $\forall n$, $\exists n$, and set quantifiers $\forall X, \exists X$.

The language $\mathcal{L}_2$ is interpreted by $\mathcal{L}_2$-*models*, which are 7-tuples

$$\mathcal{M} = (|\mathcal{M}|, \mathcal{S}, +, \cdot, 0, 1, <),$$

where $|\mathcal{M}|$ is a set acting as the number variables, and $\mathcal{S} \subseteq \mathcal{P}(|\mathcal{M}|)$ acts as the set variables. We assume that $|\mathcal{M}|$ and $\mathcal{S}$ are always disjoint, and we will often denote such a model by the pair $(\mathcal{M}, \mathcal{S})$. Formulas in $\mathcal{L}_2$-models are interpreted in the standard way.

For each $k \in \mathbb{N}$, the scheme of Σ^0_k *induction* consists of all axioms of the form

$$[\varphi(0) \wedge \forall n[\varphi(n) \rightarrow \varphi(n+1)]] \rightarrow \forall n \varphi(n)$$

where $\varphi(n)$ is any Σ^0_k $\mathcal{L}_2$-formula. Similarly one defines the scheme of Π^0_k induction.

We also define, for each $k \in \mathbb{N}$, the scheme of Δ^0_k *comprehension*. The scheme consists of all axioms of the form

$$\forall n[\varphi(n) \leftrightarrow \psi(n)] \rightarrow \exists X \forall n[n \in X \leftrightarrow \varphi(n)]$$

where $\varphi(n)$ is any Σ^0_k formula and $\psi(n)$ is any Π^0_k formula.

Definition 5.1. The *recursive comprehension axiom system*, denoted RCA_0, is the formal system in the language $\mathcal{L}_2$ whose axioms consist of the schemes of Σ^0_1 induction and Δ^0_1 comprehension, together with the universal closures of the following basic axioms:

1. $n + 1 \neq 0$
2. $m + 1 = n + 1 \rightarrow m = n$

3. $m + 0 = m$
4. $m + (n + 1) = (m + n) + 1$
5. $m \cdot 0 = 0$
6. $m \cdot (n + 1) = (m \cdot n) + m$
7. $\neg m < 0$
8. $m < n + 1 \leftrightarrow (m < n \vee m = n)$

A *binary tree* T is a subset of $2^{<\mathbb{N}}$ such that for every $\tau \in T$, if $\tau' \preceq \tau$ then $\tau \in T$. A *path* through T is a binary sequence Z such that for each $n \in \mathbb{N}$, one has $Z|_n \in T$.

Definition 5.2. *Weak König's lemma* (WKL) is the statement that every infinite binary tree has a path. The formal system WKL_0 in the language $\mathcal{L}_2$ consists of RCA_0 plus WKL.

An $\mathcal{L}_2$-formula φ is *arithmetical* if it contains no set quantifiers.

Definition 5.3. The formal system ACA_0 in the language $\mathcal{L}_2$ consists of RCA_0 together with the induction axiom

$$[0 \in X \wedge \forall n[n \in X \rightarrow n + 1 \in X]] \rightarrow \forall n[n \in X]$$

and the comprehension axioms

$$\exists X \forall n[n \in X \leftrightarrow \varphi(n)]$$

for every arithmetical formula $\varphi(n)$ in which X does not occur free.

Definition 5.4. *Weak weak König's lemma* (WWKL) is the statement that for any binary tree T, if

$$\lim_{n \to \infty} \frac{|\{\tau \in T : |\tau| = n\}|}{2^n} > 0$$

then T has a path. WWKL_0 is the formal system in $\mathcal{L}_2$ consisting of RCA_0 and WWKL.

As with the case of the Church-Turing thesis in computability theory, in reverse mathematics one tends to move past the rigid architecture of axiomatic systems. Instead, definitions and proofs are presented in an informal way, but one takes great care to avoid using principles of comprehension and induction that are not accessible in the system one is working with. Whenever one does use such a principle, one must make sure to explicitly communicate that one is doing so.

Note that we take for granted the effective coding of various mathematical objects, such as rational and real numbers, sequences, functions, etc. The explicit encodings can be found in [25].

We make the following definitions within RCA_0. A *uniformly continuous function* $f : [0, 1] \to \mathbb{R}$ is presented by a *Cauchy name*; i.e., a sequence $(f_s)_{s \in \mathbb{N}}$ of rational polynomials or polygonal functions with rational breakpoints such that $||p_s - p_r||_\infty \leq 2^{-s}$ for all $r > s$. The sequence $(f_s)_{s \in \mathbb{N}}$ is intended to describe $f = \lim_{s \to \infty} f_s$.

If $\Pi = \{t_0, \ldots, t_n\}$ is a partition of a set A (i.e., $t_i < t_{i+1}$ for each i), let

$$S(f, \Pi) = \sum_{i=0}^{n-1} |f(t_{i+1}) - f(t_i)|.$$

We say that f is of *bounded variation* if there is $k \in \mathbb{N}$ such that $S(f, \Pi) \leq k$ for every partition Π of $[0, 1]$. Note that our definition of bounded variation is chosen to avoid having to *declare* that the supremum from Definition 2.1 exists.

In this section we work with the broader notion of Jordan decomposition, where the components of the decomposition are not required to be minimal. For functions $f, g : [0, 1] \to \mathbb{R}$, write

$$f \leq_{\mathsf{slope}} g \text{ iff } \forall x \forall y [x < y \to f(y) - f(x) \leq g(y) - g(x)];$$

i.e., the slopes of g are at least as big as the slopes of f. It is not difficult to see that finding a Jordan decomposition of f is equivalent to finding a non-decreasing function g such that $f \leq_{\mathsf{slope}} g$ and putting $f = g - (g - f)$.

Note that we only work with uniformly continuous functions. In computable analysis functions with domain $[0, 1]$ are uniformly continuous by definition, so we can't use results from the previous sections to prove general theorems about functions which may only be continuous.

5.2 Jordan decomposition for uniformly continuous functions

The Jordan decomposition theorem 2.2 can be strengthened: if f is uniformly continuous then the functions f^+ and f^- comprising its decomposition can be chosen to be uniformly continuous too. We first study the complexity of Jordan's theorem for uniformly continuous functions whose decomposition constituents are *required* to be uniformly continuous.

The principle $\mathsf{Jordan}_{\mathtt{cont}}$ states that for every uniformly continuous function f of bounded variation, there is a non-decreasing uniformly continuous function $g : [0, 1] \to \mathbb{R}$ such that $f \leq_{\mathsf{slope}} g$.

Theorem 5.5. $\mathsf{RCA}_0 \vdash \mathsf{Jordan}_{\mathtt{cont}} \leftrightarrow \mathsf{ACA}_0$.

Proof. $\Leftarrow$: Given a Cauchy name $(f_s)_{s \in \mathbb{N}}$ for f, we construct a Cauchy name for $\boldsymbol{v}_f$. By [25, Theorem VIII.1.12], we may use ACA_0 to declare that the jump $\{f_s : s \in \mathbb{N}\}'$ of the representation of f exists. For each $n \in \mathbb{N}$ define $g_n : [0, 1] \to \mathbb{R}$ as follows. Using $\{f_s : s \in \mathbb{N}\}'$, compute t such that $\forall u > t [||\boldsymbol{v}_{f_u} - \boldsymbol{v}_{f_t}||_\infty \leq 2^{-n}]$, and set $g_n = \boldsymbol{v}_{f_t}$. Then $(g_s)_{s \in \mathbb{N}}$ defines the desired function.

$\Rightarrow$: Fix a model of $\mathsf{RCA}_0 + \mathsf{Jordan}_{\mathtt{cont}}$. Let

$$q_n = 1 - 2^{-n-1}, \text{ and } q_{n,s} = q_n - 2^{-n-s-1}.$$

Let $h : \mathbb{N} \to \mathbb{N}$ be an injective function. We will show that the range of h exists, which is equivalent to ACA_0 ([25, Lemma III.1.3]). Recall that we let $\mathrm{M}_A(h, r)$ denote a sawtooth function on the interval A with r many teeth of height h. For each $s \in \mathbb{N}$, define a continuous function f_s as follows. On each interval of the form $I_k = [q_{h(k),k}, q_{h(k),k+1}]$ put

$$f_s = \begin{cases} \mathrm{M}_{I_k}(2^{-s}, 2^{s-h(k)}) & \text{if } s \geq h(k), \\ 0 & \text{otherwise.} \end{cases}$$

Let $f_s = 0$ elsewhere. This ensures that the range of h is encoded into the variation of f_s.

The sequence $(f_s)_{s\in\mathbb{N}}$ defines a continuous function $f = \lim_{s\to\infty} f_s$, as $||f_t - f_s||_\infty \leq 2^{-s}$ for all $t > s$. We show that f is of bounded variation with bound 1. Note that we only need to examine the variation of f on the disjoint intervals $[q_{h(k),k}, q_{h(k),k+1}]$ since $f = 0$ elsewhere.

Let $m \in \mathbb{N}$. For $k \in \{0, \dots, m\}$, let Π_k partition I_k. We estimate the variation of f on the interval $\bigcup_{k\leq m} I_k$. Without loss of generality we may assume that each partition contains the midpoints and endpoints of the sawteeth defined on I_k.[1] This allows us to easily compute the variation of f as the piece-wise combination of non-decreasing functions. For all $s \geq \max\{h(k) : 0 \leq k \leq m\}$ one has

$$\sum_{k=0}^{m} S(f, \Pi_k) = \sum_{k=0}^{m} S(f_s, \Pi_k) = \sum_{k=0}^{m} 2^{-h(k)+1} < 1,$$

which establishes the desired bound.

By $\mathsf{Jordan}_{\mathtt{cont}}$, take $g : [0,1] \to \mathbb{R}$ non-decreasing and continuous such that $f \leq_{\mathsf{slope}} g$. Note that g is presented by a Cauchy name $(g_s)_{s\in\mathbb{N}}$. Given that the range of h is encoded in the variation of f, we will use the (easily computable) variation of g on the interval $[q_{n,k}, q_{n,k+1}]$ to bound to possible pre-images under of n under h.

Define a computable function $\gamma : \mathbb{N} \to \mathbb{N}$ such that $g(q_n) - g(q_{n,\gamma(n)}) < 2^{-n}$ as follows. There is a Σ^0_0 formula $\theta(n, m, k)$ such that

$$\exists m\theta(n, m, k) \leftrightarrow g(q_n) - g(q_{n,k}) < 2^{-n}.$$

Since g is continuous and $\lim_{s\to\infty} q_{n,s} = q_n$ one has $\forall n\exists k\exists m\theta(n, m, k)$. As this sentence is Π^0_2, given any instance of the variable n one can effectively obtain a witness for the Σ^0_1 formula $\exists k\exists m\theta(n, m, k)$ (see, e.g. [25, Theorem II.3.5]). Thus by minimization we may put $\gamma(n) = k$, where $\langle m, k\rangle$ is least such that $\theta(n, m, k)$ holds.

Now if $h(k) = n$ then by the monotonicity of g,

$$g(q_n) - g(q_{n,k}) \geq g(q_{n,k+1}) - g(q_{n,k}).$$

Let Π be a partition of $[q_{n,k}, q_{n,k+1}]$ containing the endpoints and midpoints of each sawtooth defined on that interval. Then since $g - f \leq g$ and the variation of an increasing function is the difference of its values at its endpoints one has

$$2^{-n+1} = S(f, \Pi) = S(g - (g - f), \Pi) \leq S(g, \Pi) + S(g - f, \Pi) \leq 2(g(q_{n,k+1}) - g(q_{n,k})).$$

[1] Indeed, this only refines the partition and provides an improved estimate.

Thus $g(q_n) - g(q_{n,k}) \geq 2^{-n}$, and then $k < \gamma(n)$. Hence

$$n \in \operatorname{rng}(h) \leftrightarrow \exists k < \gamma(n)[h(k) = n],$$

so the range of h exists by Δ^0_1 comprehension.

□

5.3 Jordan decomposition for functions of rational domain

We have shown that requiring a Jordan decomposition to consist of uniformly continuous functions causes the complexity of the Jordan decomposition theorem to reach ACA_0. Working with complicated conditions like uniform continuity (a Π^0_3 statement) gives one more freedom to perform encodings of various objects. Above, we encoded the range of an injective function h into the variation of a function of bounded variation f. Decomposing f into uniformly continuous functions allowed us to recover enough information to decide whether some number was the image of another under h. Working with more primitive objects prevents the encoding of high complexity sets, however. We now relax the requirements of the Jordan decomposition by only stipulating that the decomposition is given by functions which are defined on the rationals. Such functions can be represented by finite strings that cumulatively describe the behaviour of the function at each rational.

Let $[0,1]_{\mathbb{Q}} := [0,1] \cap \mathbb{Q}$. We present a function $g : [0,1]_{\mathbb{Q}} \to \mathbb{R}$ as a binary sequence Z_g in the following way. Fix an effective listing $(p_n, q_n)_{n\in\mathbb{N}}$ of all rationals $p_n, q_n \in \mathbb{Q}$ such that $0 \leq p_n \leq 1$. We let $Z_g(2n) = 1$ iff $g(p_n) < q_n$, and $Z_g(2n+1) = 1$ iff $g(p_n) > q_n$. Then Z defines the (total) function $g : [0,1] \to \mathbb{R}$ specified by $g(p) = \inf\{q \in \mathbb{Q} : g(p) > q\}$.

We modify the $\leq_{\mathsf{slope}}$ notation for functions of rational domain. For $f, g :\subseteq [0,1] \to \mathbb{R}$ we let

$$f \leq^*_{\mathsf{slope}} g \text{ iff } \forall x, y \in [0,1]_{\mathbb{Q}}[x < y \to (f(y) - f(x) \leq g(y) - g(x))].$$

The principle $\mathsf{Jordan}_{\mathbb{Q}}$ is the statement that for every continuous function f of bounded variation, there is a non-decreasing function $g : [0,1]_{\mathbb{Q}} \to \mathbb{R}$ such that $f \leq^*_{\mathsf{slope}} g$.

Theorem 5.6. $\mathsf{RCA}_0 \vdash \mathsf{Jordan}_{\mathbb{Q}} \leftrightarrow \mathsf{WKL}_0$.

Proof. $\Leftarrow$: Let f be a function of bounded variation, given by a Cauchy name $(f_s)_{s\in\mathbb{N}}$. We construct a binary tree T such that any path Z through T encodes a non-decreasing function $g : [0,1]_{\mathbb{Q}} \to \mathbb{R}$ with $f \leq^*_{\mathsf{slope}} g$. To do this, we must ensure the following two requirements at each stage k for every potential function g:

$\mathcal{R}_1$: if $p_s < p_r$, $q_s = q_r$, and $g(p_s) > q_s$ with $r, s < k$ then $g(p_r) > q_r$;

$\mathcal{R}_2$: if $p_s < p_r$ with $r, s < k$ then for any q_j with $j < k$ and $g(p_r) - g(p_s) < q_j$ one has $f_k(p_r) - f_k(p_s) < q_j + 2^{-k}$.

The first requirement guarantees that any g encoded by a path Z_g through T is non-decreasing. The second guarantees the slope condition. Formally, let $\ell(\tau) = \left\lceil \frac{|\tau|-1}{2} \right\rceil$. By Δ^0_1 comprehension take T to be the set of all $\tau \in 2^{<\mathbb{N}}$ such that

(i) $\forall r, s < \ell(\tau) \Big[(p_s < p_r \wedge q_s = q_r \wedge \tau(2s+1) = 1) \rightarrow \tau(2r+1) = 1\Big]$, and

(ii) $\forall r, s < \ell(\tau) \Big[(p_s < p_r \wedge \tau(2s) = 1 \wedge \tau(2r+1) = 1)$

$$\rightarrow \forall j < |\tau| (q_s - q_r < q_j \rightarrow |f_{|\tau|}(p_s) - f_{|\tau|}(p_r)| < q_j + 2^{-(|\tau|+1)})\Big].$$

To see that T is infinite, notice that since f is of bounded variation, the string $Z_{v_{f_s}}|_s$ is an element of T for every $s \in \mathbb{N}$. Thus by WKL, T has a path Z.

Let g be the unique function of type $[0,1]_{\mathbb{Q}} \to \mathbb{R}$ such that $Z = Z_g$.

Claim 5.7. *The function g is non-decreasing.*

Proof. Take $x, y \in [0,1]_{\mathbb{Q}}$ with $x < y$. Let $q \in \mathbb{Q}$. It suffices to show that if $g(x) > q$ then $g(y) > q$. There is $r, s \in \mathbb{N}$ such that $p_s = y$, $q_s = q$, $p_r = x$, and $q_r = q$. If $g(p_r) > q_r$ then by clause (i), $g(p_s) > q_s$. ◊

Claim 5.8. $f \leq^*_{\mathsf{slope}} g$.

Proof. Let $x, y \in [0,1]_{\mathbb{Q}}$ such that $x < y$. There is $k \in \mathbb{N}$ and $r, s, j < k$ such that

1. $p_r \leq x < y \leq p_s$,
2. $q_r < g(p_r) < g(p_s) < q_s$, and
3. $q_j + 2^{-(k-2)} < q_s - q_r < q_j + 2^{-(k-3)}$.

Since g is non-decreasing this implies that $q_j + 2^{-(k-2)} < g(y) - g(x) < q_j + 2^{-(k-3)}$. By the continuity of f one can ensure p_r and p_s are chosen so close to x and y that

4. $|f(y) - f(x)| \leq |f(p_s) - f(p_r)| + 2^{-k}$.

Using clause (ii) one has

$$\begin{aligned} |f(y) - f(x)| &\leq |f(p_s) - f(p_r)| + 2^{-k} \\ &= |f(p_s) - f_k(p_s) + f_k(p_s) - f_k(p_r) + f_k(p_r) - f(p_r)| + 2^{-k} \\ &\leq |f(p_s) - f_k(p_s)| + |f_k(p_s) - f_k(p_r)| + |f_k(p_r) - f(p_r)| + 2^{-k} \\ &< 2^{-k} + q_j + 2^{-(k+1)} + 2^{-k} + 2^{-k} \\ &< q_j + 2^{-(k-2)} \\ &< g(y) - g(x). \end{aligned}$$

◊

⇒: We reason within RCA_0. Let $T \subseteq 2^{<\mathbb{N}}$ be an infinite binary tree. We will show that T has a path. Let $\widetilde{T} = \{\tau \in 2^{<\mathbb{N}} : \tau \notin T \wedge \tau|_{(|\tau|-1)} \in T\}$. Without loss of generality we may assume that $\widetilde{T}$ is infinite. Let $h : \mathbb{N} \to \mathbb{N}$ be defined by

$$\begin{aligned} h(0) &= \min\{\sigma \in T : \forall \tau \in T[\sigma \not\prec \tau]\}, \\ h(n+1) &= \min\{\sigma \in T \setminus \{h(k) : k \leq n\} : \forall \tau \in T \setminus \{h(k) : k \leq n\}[\sigma \not\prec \tau]\}, \end{aligned}$$

where the minimum is taken with respect to the usual integer encoding of binary strings. Then $T \setminus \operatorname{rng}(h) = \{\tau \in T : \tau \text{ has infinitely many extensions in } T\}$. Let $(\widetilde{\sigma}_k)_{k\in\mathbb{N}}$ be an enumeration of $\widetilde{T}$ such that $|\widetilde{\sigma}_i| \leq |\widetilde{\sigma}_{i+1}|$. Note that for any $k, \ell \in \mathbb{N}$,

$$|\widetilde{\sigma}_k| \leq \ell \to k \leq 2^\ell. \tag{2}$$

For all $\sigma \in 2^{<\mathbb{N}}$ put $I_\sigma = [0.\sigma, 0.\sigma + 2^{-|\sigma|}]$. For each $s \in \mathbb{N}$ define a polygonal function $f_s : [0,1] \to \mathbb{R}$ as follows. On the interval $I_{\widetilde{\sigma}_k}$ set

$$f_s = \begin{cases} \mathrm{M}_{I_{\widetilde{\sigma}_k}}(2^{-s}, 2^{s-h(k)}) & \text{if } s \geq h(k), \\ 0 & \text{otherwise.} \end{cases}$$

Let $f_s = 0$ elsewhere. Then $(f_s)_{s\in\mathbb{N}}$ defines a continuous function $f = \lim_s f_s$. We show that f is of bounded variation. As before, we need only consider the variation of f on the disjoint intervals $I_{\widetilde{\sigma}_k}$. Let $m \in \mathbb{N}$, and for each $k \leq m$ let Π_k be a partition of $I_{\widetilde{\sigma}_k}$ containing the midpoints and endpoints of each sawtooth defined on that interval. For all $s \geq \max\{h(k) : 0 \leq k \leq m\}$ one has

$$\sum_{k=0}^{m} S(f, \Pi_k) = \sum_{k=0}^{m} S(f_s, \Pi_k) = \sum_{k=0}^{m} 2^{-h(k)+1} < 1,$$

as required.

By $\mathsf{Jordan}_{\mathbb{Q}}$ there exists $g : [0,1]_{\mathbb{Q}} \to \mathbb{R}$ non-decreasing such that $f \leq^*_{\mathsf{slope}} g$. Define $\Delta : \mathbb{N} \to \mathbb{R}$ by $\Delta(k) = \max\{g(0.\sigma + 2^{-|\sigma|}) - g(0.\sigma) : \sigma \in T \wedge |\sigma| = k\}$. Then Δ is non-increasing.

There are two cases to consider. If $\lim_{n\to\infty} \Delta(n) = 0$, then g behaves as a continuous function. This provides a decomposition of f that allows us to use an argument similar to the one in Theorem 5.5 to prove the existence of $\operatorname{rng}(h)$. One can then find a path through T by avoiding this set.

Otherwise, there is a jump-type discontinuity of g. The intervals around this point correspond to strings which form an infinite subtree $\widehat{T}$ of T. One can bound the size of any prefix-free subset of $\widehat{T}$ using the size of the jump, and thus effectively find a path through $\widehat{T}$.

Case 1. $\lim_{n\to\infty} \Delta(n) = 0$. Take $\gamma : \mathbb{N} \to \mathbb{N}$ such that $\Delta(\gamma(n)) < 2^{-n}$. This can be done using the method in the left to right direction in the proof of Theorem 5.5. If $h(k) = n$ then

$$g(0.\widetilde{\sigma}_k + 2^{-|\widetilde{\sigma}_k|}) - g(0.\widetilde{\sigma}_k) \geq 2^{-n}.$$

Hence $|\widetilde{\sigma}_k| \leq \gamma(n)$, and then by (2) $k \leq 2^{\gamma(n)}$. This gives

$$n \in \text{rng}(h) \leftrightarrow \exists k \leq 2^{\gamma(n)}[h(k) = n],$$

so rng(h) exists by Δ^0_1 comprehension. Then $T \setminus \text{rng}(h)$ exists. Define $Z : \mathbb{N} \to \mathbb{N}$ by primitive recursion by setting $Z(0) = \varnothing$ (the empty string), and $Z(n+1) = \mu\sigma[Z(n)\sigma \in T \setminus \text{rng}(h)]$. Then Z is a path through T.

Case 2. There exists $q \in \mathbb{Q}$ such that $q > 0 \wedge \forall m \exists n[n > m \wedge \Delta(n) \geq q]$. There is a Σ^0_0 formula φ such that for all $\sigma \in 2^{<\mathbb{N}}$,

$$g(0.\sigma + 2^{-|\sigma|}) - g(0.\sigma) < q \leftrightarrow \exists s\ \varphi(\sigma, s).$$

Recall that in Theorem 4.9 we defined a Π^0_1 class $\hat{\zeta}$ containing the names of Jordan decompositions of f as the set of paths on a Π^0_1 tree, and could compute an element of this class by the low for z basis theorem. Presently we cannot argue that this class exists since Π^0_1 comprehension is not available to us. To handle this we define a Δ^0_1 tree which has the same paths as $\hat{\zeta}$, but which we can assert to exist in our model.

By Δ^0_1 comprehension, let $\widehat{T} = \{\sigma \in T : \forall \rho[\exists s < |\sigma|\ \varphi(\rho, s) \to \rho \not\preceq \sigma]\}$. Then $\widehat{T}$ is an infinite subtree of T. To see that $\widehat{T}$ is infinite, notice that by the case assumption there are infinitely many strings σ such that $\forall s\ \varphi(\sigma, s)$. For any such σ, if $\rho \preceq \sigma$ then $g(0.\rho + 2^{-|\rho|}) - g(0.\rho) \geq q$, so $\neg\exists s < |\sigma|\ \varphi(\rho, s)$, whereby $\sigma \in \widehat{T}$. This argument also shows that $\widehat{T}$ is closed under prefixes.

There is $K \in \mathbb{N}$ such that for any prefix-free $P \subseteq \widehat{T}$ one has $|P| < K$. For instance, one may take K so that $Kq > g(1) - g(0)$. Then for any prefix-free $P \subseteq \widehat{T}$,

$$|P|q < \sum_{\sigma \in P} g(0.\sigma + 2^{-|\sigma|}) - g(0.\sigma) \leq g(1) - g(0) < Kq.$$

Claim 5.9. *$\widehat{T}$ has a path.*

Proof. By Σ^0_1 induction, take

$$k = \max\{i \leq K : \text{there is a prefix-free set } P \subseteq \widehat{T} \text{ with } |P| = i\}. \tag{3}$$

Let $P_k \subseteq \widehat{T}$ witness (3). Let $\sigma = \max P_k$, where the max is taken with respect to the usual integer encoding of binary strings. Let $\ell = \max\{|\tau| : \tau \in P_k\}$. Any $\tau \in \widehat{T}$ with $|\tau| > \ell$ must extend an element of P_k, and must have at most one successor. The Π^0_1 set

$$\{\tau \in P_k : \forall v \exists \rho \in \widehat{T}\,[|\rho| = v \wedge \exists \tau' \in P_k(\tau' \leq \tau \wedge \tau' \preceq \rho)]\}$$

is non-empty because it contains σ. Thus, by Σ^0_1 induction, it has a least element τ. Then $\tau \in \text{Ext}(\widehat{T})$. Since each extension of τ whose length exceeds ℓ has exactly one successor, we can effectively find a path through $\widehat{T}$ extending τ. ◊

□

We thank Paul Shafer who provided useful comments on a previous instance of this argument.

6 Differentiability of functions of bounded variation in WWKL_0

In Section 3 we showed that proving decomposability of functions of bounded variation could not be undertaken in RCA_0, and in the previous section we showed that this operation requires either WKL_0 or ACA_0. It is interesting that proving decomposability requires systems of this strength: as we will see momentarily, functions of bounded variation can be proven to be a.e. differentiable using just WWKL_0. Despite the fact that differentiation seems to be an operation of higher complexity (given that it involves, for instance, the convergence of a sequence), the interaction between differentiability and Martin-Löf randomness that we saw in Section 4.2 means that it suffices to guarantee the existence of Martin-Löf random reals.

Definition 6.1. Let $f : [0,1] \to \mathbb{R}$. The *upper* and *lower pseudo-derivatives of f* are defined by

$$\widetilde{D}f(x) = \lim_{h\to 0^+} \sup\{S_f(a,b) : a,b \in [0,1]_{\mathbb{Q}} \ \wedge\ a \le x \le b \ \wedge\ 0 < b-a < h\}, \text{ and}$$
$$\underset{\sim}{D}f(x) = \lim_{h\to 0^+} \inf\{S_f(a,b) : a,b \in [0,1]_{\mathbb{Q}} \ \wedge a \le x \le b \ \wedge\ 0 < b-a < h\}.$$

A function $f :\subseteq [0,1] \to \mathbb{R}$ with domain containing $[0,1]_{\mathbb{Q}}$ is *pseudo-differentiable* at $z \in (0,1)$ if $\underset{\sim}{D}f(z)$ and $\widetilde{D}f(z)$ exist, are equal, and are both finite.

We say that f is pseudo-differentiable *almost surely* if for any family of open intervals $\mathcal{U} = \{(u_i,v_i)\}_{i\in\mathbb{N}}$ there exists $m \in \mathbb{N}$ such that, if $\mathcal{U}$ covers every pseudo-differentiable point of f, then $\sum_{i=0}^{m}(v_i - u_i) \ge 1$.

The property of being "pseudo-differentiable almost surely" is intended to capture the notion of a.e. differentiability in a way that is appropriate for second-order arithmetic.

Lemma 6.2 (Simpson and Yokoyama [26])**.** *For any countable model $(\mathcal{M},\mathcal{S}) \models \mathsf{WWKL}_0$ there is $\widehat{\mathcal{S}} \supseteq \mathcal{S}$ satisfying*

1. $(\mathcal{M},\widehat{\mathcal{S}}) \models \mathsf{WKL}_0$*, and*
2. *for any $A \in \widehat{\mathcal{S}}$ there is $z \in \mathcal{S}$ such that z is Martin-Löf random relative to A.*

Theorem 6.3 (WWKL_0)**.** *Every uniformly continuous function of bounded variation is pseudo-differentiable at some point.*

Proof. We show that the result holds in any countable model of WWKL_0. Let $(\mathcal{M},\mathcal{S})$ be a countable model of WWKL_0, and let $f : [0,1] \to \mathbb{R}$ be a continuous function of bounded variation in $(\mathcal{M},\mathcal{S})$. By Lemma 6.2, there is a model $(\mathcal{M},\widehat{\mathcal{S}}) \models \mathsf{WKL}_0$ such that $\mathcal{S} \subseteq \widehat{\mathcal{S}}$. By Theorem 5.6,

$$(\mathcal{M},\widehat{\mathcal{S}}) \models \mathsf{Jordan}_{\mathbb{Q}}.$$

Hence $\widehat{\mathcal{S}}$ contains a non-decreasing function $g : [0,1]_{\mathbb{Q}} \to \mathbb{R}$ such that $f \le^*_{\mathsf{slope}} g$.

Within $(\mathcal{M}, \widehat{\mathcal{S}})$, define $h : [0,1]_{\mathbb{Q}} \to \mathbb{R}$ by $h(x) = g(x) - f(x)$. By Lemma 6.2 again, there is a real $z \in (0,1)$ such that $z \in \mathcal{S}$ and $z \in \mathsf{MLR}^{g \oplus h}$. The functions g and h are pseudo-differentiable at z in $(\mathcal{M}, \widehat{\mathcal{S}})$ by Theorem 4.4, and therefore f is pseudo-differentiable at z in $(\mathcal{M}, \widehat{S})$. Thus f is pseudo-differentiable at z in $(\mathcal{M}, \mathcal{S})$. □

Theorem 6.4 (WWKL_0)**.** *Every continuous function of bounded variation is pseudo-differentiable almost surely.*

Proof. Let $(\mathcal{M}, \mathcal{S})$ be a countable model of WWKL_0. Let $f : [0,1] \to \mathbb{R}$ be a continuous function of bounded variation. As in the proof of Theorem 6.3, we obtain non-decreasing functions $g, h : [0,1]_{\mathbb{Q}} \to \mathbb{R}$ such that $f|_{\mathbb{Q}} = g - h$, and a $\mathsf{MLR}^{g \oplus h}$ real $z \in (0,1)$ such that $z \in \mathcal{S}$. Reasoning within $(\mathcal{M}, \mathcal{S})$, let $\mathcal{U} = \{(u_i, v_i)\}_{i \in \mathbb{N}}$ be a family of open intervals such that for some rational $\varepsilon > 0$ one has

$$\forall m \sum_{i=0}^{m} (v_i - u_i) < 1 - \varepsilon.$$

Let $\mathcal{V} = [0,1] \setminus \bigcup \mathcal{U}$. We will obtain a tail $u \in \mathcal{V}$ of z such that u is Martin-Löf random, and therefore a point of (pseudo-) differentiability for f. Without loss of generality we may assume that the endpoints of $\mathcal{U}$ are dyadic rationals. The family $\mathcal{U}$ is then given by a sequence $(\sigma_i)_{i \in \mathbb{N}}$ of strings; i.e., for all $i \in \mathbb{N}$

$$(0.\sigma_i, 0.\sigma_i + 2^{-|\sigma_i|}) = (u_i, v_i).$$

Thus $\lambda[\![\mathcal{U}]\!] < 1$. For each $i \in \mathbb{N}$, let

$$S_i = [\![\{\sigma_{k_1} \cdots \sigma_{k_i} : \sigma_{k_n} \in \{\sigma_j : j \in \mathbb{N}\} \text{ for each } n \in \{1, \dots, i\}\}]\!].$$

Note that $S_1 = [\![\mathcal{U}]\!]$, hence $\lambda S_1 < 1$; and by Σ^0_1 induction one has $\lambda S_i = (\lambda S_1)^i$. There is a computable function $\gamma : \mathbb{N} \to \mathbb{N}$ which thins $(S_i)_{i \in \mathbb{N}}$ to a Martin-Löf test $(S_{\gamma(i)})_{i \in \mathbb{N}}$. Since z is Martin-Löf random, there is j least such that $z \notin S_{\gamma(j)}$. Thus there exists $\sigma \in 2^{<\mathbb{N}}$ and $u \in 2^{\mathbb{N}}$ such that $z = \sigma u$ and $u \notin [\![\mathcal{U}]\!]$. Then u is the desired suffix.

Assume that $u \notin \mathsf{MLR}^{g \oplus h}$. Then there is a Martin-Löf test $(\mathcal{G}_n)_{n \in \mathbb{N}}$ relative to $g \oplus h$ such that $u \in \bigcap_n \mathcal{G}_n$. For each n, put $\widehat{\mathcal{G}}_n = \{\sigma g : g \in \mathcal{G}_n\}$. Then $\lambda(\widehat{\mathcal{G}}_n) \le \lambda(\mathcal{G}_n) \le 2^{-n}$. Hence $(\widehat{\mathcal{G}}_n)_{n \in \mathbb{N}}$ is a Martin-Löf test relative to $g \oplus h$ which captures z, a contradiction. □

The proof of Theorem 6.4 makes use of Kučera's argument [19, Proposition 3.2.24] which proves, for any given Π^0_1 class P of positive measure and Martin-Löf random sequence Z, the existence of a tail of Z which lies in P.

We are close to a reversal of Theorem 6.3 and Theorem 6.4. The reversal would assert that the following are equivalent over RCA_0.

1. WWKL_0
2. Every uniformly continuous function of bounded variation is pseudo-differentiable at some point.

3. Every uniformly continuous function of bounded variation is pseudo-differentiable almost surely.

The implications $1 \Rightarrow 2$ and $1 \Rightarrow 3$ are established by Theorem 6.3 and Theorem 6.4. For the converses, let X be a set. Construct the function f of bounded variation from Section 4.2 relative to X. Then the points of differentiability of f are precisely the reals which are Martin-Löf random relative to X. Thus, since both clause 2 and clause 3 ensure the existence of a point of differentiability of f, this shows the existence of such a real.

Some remarks must be made about the argument. One proves that f is absolutely continuous (and hence of bounded variation) by showing that it can be written as an integral, but to the author's knowledge the equivalence

$$f \text{ absolutely continuous} \Leftrightarrow f(x) = \int_0^x g \mathrm{d}\lambda \text{ for some } g \in L_1$$

has not yet been verified in WWKL_0. Yu [29, Theorem 3.5] shows that a form of the Lebesgue dominated convergence theorem (DCT) is equivalent to WWKL_0 over RCA_0, but does not establish the a.e. finiteness requirement needed for our choice of g (which we constructed in Section 4.2). By [1, Theorem 4.3], a weaker version of DCT (where all functions are assumed to already be integrable) is equivalent over RCA_0 to 2-WWKL, the statement that any Π^0_2 tree of positive measure has a path. Note that this is strictly stronger than WWKL.

7 The Jordan decomposition operator in type two computability

We now move to the setting of Weihrauch's *type 2 theory of effectivity* (TTE). We study recent work of Weihrauch and Jafarikhah [13] on the computability of the operator $\boldsymbol{x} \mapsto (\boldsymbol{x}^+, \boldsymbol{x}^-)$ which maps objects $\boldsymbol{x}$ (either continuous linear functionals, functions of bounded variation, or signed measures) to their minimal Jordan decompositions.

We will briefly review TTE. Let Σ be a finite alphabet. Computability of functions defined on Σ^ω is defined via (type 2) Turing machines which map sequences to sequences; the sequence being fed to the Turing machine symbol by symbol. We say that a function $F : \Sigma^\omega \to \Sigma^\omega$ is *computable* if there is a Turing machine that, given $p \in \mathsf{dom}(F)$ as a stream on an input tape, produces $F(p)$ on its output tape symbol by symbol, and that, given $p \in \Sigma^\omega \setminus \mathsf{dom}(F)$, does not write infinitely many symbols on its output tape.

A *representation* of a set X is a surjective function $\delta :\subseteq \Sigma^\omega \to X$. If $\delta(p) = x$, we call p a δ-*name* of x. Let M_0, M_1 be two sets represented by $\delta_0 :\subseteq \Sigma^\omega \to M_0$ and $\delta_1 :\subseteq \Sigma^\omega \to M_1$ respectively. A function $h :\subseteq \Sigma^\omega \to \Sigma^\omega$ is a (δ_0, δ_1)-*realisation* of a function $f :\subseteq M_0 \to M_1$ if for all $p \in \mathsf{dom}(\delta_0)$,

$$\delta_0(p) \in \mathsf{dom}(f) \;\Rightarrow\; \delta_1(h(p)) = f(\delta_0(p)).$$

The function f is called (δ_0, δ_1)-*computable* if it has a computable (δ_0, δ_1)-realiser, and f is called (δ_0, δ_1)-*continuous* if it has a continuous (δ_0, δ_1)-realiser with respect to the usual topology on Cantor space. The intuition for computability is that from a δ_0-name of $x \in \mathsf{dom}(f)$, a type 2 Turing machine can compute a δ_1-name of $f(x)$.

For the set $\mathbb{R}$ we use the Cauchy representation $\rho : \Sigma^\omega \to \mathbb{R}$ defined by $\rho(p) = x$ iff p encodes a sequence $(\alpha_i)_{i\in\mathbb{N}}$ of rationals such that $|x - \alpha_i| \leq 2^{-i}$ for all i. We also use representations $\rho_<$ and $\rho_>$ defined by $\rho_<(p) = x$ iff p encodes a sequence of rationals $(\alpha_i)_i$ such that $x = \sup_i \alpha_i$, and $\rho_>(p) = x$ iff p encodes a sequence of rationals $(\beta_i)_i$ for which $x = \inf_i \beta_i$. A function $f : \mathbb{R} \to \mathbb{R}$ is computable in the sense of Definition 2.5 iff it is (ρ, ρ)-computable. That is, from a Cauchy name of x one may compute a Cauchy name of $f(x)$. A further example is given by the variation function $\boldsymbol{v}_f$, which is $(\rho, \rho_<)$-computable. Note that if f is $(\rho, \rho_<)$-computable and $(\rho, \rho_>)$-computable then f is (ρ, ρ)-computable.

The canonical representation of the product space $M_0 \times M_1$ is the representation $[\delta_0, \delta_1]$ given by

$$[\delta_0, \delta_1]\langle p, q\rangle = (\delta_0(p), \delta_1(q)).$$

The canonical representation can be extended to a representation δ on the product space $\Pi_{i=1}^{k+1} M_i$ by putting $\delta = [\delta^k, \delta_{k+1}]$, where δ^k is the canonical representation for $\Pi_{i=1}^{k} M_i$, and δ_{k+1} is a representation of M_{k+1}.

The representation δ_0 is *reducible* to δ_1, written $\delta_0 \leq \delta_1$, if there is a computable function $h :\subseteq \Sigma^\omega \to \Sigma^\omega$ where, for all $p \in \mathsf{dom}(\delta_0)$, one has $\delta_0(p) = \delta_1(h(p))$. Thus δ_0 is reducible to δ_1 if there is a computable function which sends a δ_0-name of an object p to a δ_1-name of p. The representations δ_0 and δ_1 are *equivalent* if $\delta_0 \leq \delta_1$ and $\delta_1 \leq \delta_0$.

By [4, Proposition 6.15] there is a representation $[\delta_0 \to \delta_1]$ of the set of all (δ_0, δ_1)-continuous functions such that

- the function $(f, x) \mapsto f(x)$ is $([[\delta_0 \to \delta_1], \delta_0], \delta_1)$-computable;
- if γ is a representation of a set of (δ_0, δ_1)-continuous functions and $(f, x) \mapsto f(x)$ is $([\gamma, \delta_0], \delta_1)$-computable, then $\gamma \leq [\delta_0 \to \delta_1]$.

Moreover, the representation $[\delta_0 \to \delta_1]$ is unique up to equivalence. Thus, from the (canonical) representation of a pair (f, x), a type 2 Turing machine can compute a representation of the image $f(x)$.

Table 2 below can be used to track the representations needed to compute the studied operators.

7.1 Decomposition of functions of bounded variation

Let $\mathbf{BV_0} \subseteq \mathbf{BV}$ be the set of all left continuous functions $g : [0, 1] \to \mathbb{R}$ of bounded variation such that $g(0) = 0$. These requirements allow an identification to be made between Borel measures and elements of $\mathbf{BV_0}$ which we will use for computation later.

We specify representations δ_{V} and $\delta_{\mathbf{BV}}$ of $\mathbf{BV_0}$ in the following manner. Let $g \in \mathbf{BV_0}$. We will require that δ_{V} encodes the image and pre-image of g on some countable dense subset of $[0, 1]$. To this end, we define $\delta_{\mathrm{V}}(p) = g$ iff p is a sequence $\langle\langle p_0, q_0\rangle, \langle p_1, q_1\rangle, \ldots\rangle$ which meets the five requirements of the following table.

1. $p_i, q_i \in \Sigma^\omega$ for all $i \in \mathbb{N}$
2. $\rho(p_0) = \rho(q_0) = 0$
3. $\rho(p_1) = 1$
4. $g(\rho(p_i)) = \rho(q_i)$ for all $i \in \mathbb{N}$
5. $A_p = \{\rho(p_i) : i \geq 2\}$ is a dense subset of $(0,1)$ on which g is continuous.

Let $\delta_{\mathbf{BV}}\langle p, q\rangle = g$ iff $\delta_{\mathrm{V}}(p) = g$ and $\rho(q) = \mathbf{V}(g)$. The representation $\delta_{\mathbf{BV}}$ encodes both the δ_{V}-name of the function g along with a Cauchy name of its total variation $\mathbf{V}(g)$.

Under these representations we can prove that the Jordan decomposition operator on $\mathbf{BV_0}$ is computable, in the sense that from a representation of $g \in \mathbf{BV_0}$, one may compute a representation of the minimal decomposition (g^+, g^-).

Theorem 7.1 (Jafarikhah and Weihrauch [13]). *The Jordan decomposition operator $g \mapsto (g^+, g^-)$ on $\mathbf{BV_0}$ is $(\delta_{\mathbf{BV}}, [\delta_{\mathrm{V}}, \delta_{\mathrm{V}}])$-computable.*

Proof. Let $\langle p, q\rangle$ be a $\delta_{\mathbf{BV}}$-name of g with $p = \langle\langle p_0, q_0\rangle, \langle p_1, q_1\rangle, \dots\rangle$. The minimal decomposition of g is given by $g^+ = (\boldsymbol{v}_g + g)/2$, and $g^- = (\boldsymbol{v}_g - g)/2$. We will compute a sequence $\langle\langle p_0, r_0\rangle, \langle p_1, r_1\rangle, \dots\rangle$ such that $\rho(r_k) = \boldsymbol{v}_g(\rho(p_k))$ for each k.

Put $r_0 := q_0$. Then $\rho(p_0) = \rho(r_0) = 0$. Since we are given $\rho(q) = \mathbf{V}(g)$ we may also put $r_1 = q$. For each k, let $x_k := \rho(p_k)$ and let

$$\Pi(k) = \{\sigma \in \mathbb{N}^{<\mathbb{N}} : \sigma(0) = 0, \sigma(|\sigma| - 1) = k, \text{ and } x_{\sigma(0)} < \cdots < x_{\sigma(|\sigma|-1)}\}.$$

Then since $A_p = \{x_k : k \in \mathbb{N}\}$ is a dense subset of $(0,1)$ on which g is continuous,

$$\boldsymbol{v}_g(x_k) = \sup_{\sigma \in \Pi(k)} \sum_{i=0}^{|\sigma|-1} |g(\sigma(i+1)) - g(\sigma(i))|.$$

Note that $\Pi(k)$ is c.e. relative to $p \oplus k$, so it has some $(p \oplus k)$-effective enumeration $(\sigma_i)_i$. Thus a $\rho_<$-name of $\boldsymbol{v}_g(x_k) = \sup_k \sum_{i=0}^{|\sigma_k|-1} |g(\sigma_k(i+1)) - g(\sigma_k(i))|$ is computable from p and k. One may similarly compute a $\rho_<$-name of $\mathbf{V}(g, [x_k, 1])$. Thus both $\boldsymbol{v}_g$ and $k \mapsto \mathbf{V}(g, [x_k, 1])$ are $(\rho, \rho_<)$-computable.

It suffices now to show that $\boldsymbol{v}_g$ is $(\rho, \rho_>)$-computable. There are sequences of rationals (encoded by elements of Σ^ω) $(\alpha_i)_i$ and $(\beta_i)_i$ computable from p and k such that

$$\mathbf{V}(g) \leq \alpha_i < \mathbf{V}(g) + 2^{-i-1} \text{ and } \mathbf{V}(g, [x_k, 1]) \geq \beta_i > \mathbf{V}(g, [x_k, 1]) - 2^{-i-1}.$$

Let $\gamma_i = \alpha_i - \beta_i$. Then $\boldsymbol{v}_g(x_k) \leq \gamma_i < \boldsymbol{v}_g(x_k) + 2^{-i}$, and hence $\inf_i \gamma_i = \boldsymbol{v}_g(x_k)$. Thus $\boldsymbol{v}_g(x_k)$ has a $\rho_>$-name computable from p and k, and so it has a ρ-name computable from p and k. We take r_k to be that name, which completes the construction of the desired sequence.

Since we have shown that $\boldsymbol{v}_g$ is (ρ, ρ)-computable from a $\delta_{\mathbf{BV}}$ name of g, we may complete the proof by observing that the minimal decomposition $((\boldsymbol{v}_g + g)/2, (\boldsymbol{v}_g - g)/2)$ is subsequently computable from a ρ-name of $\boldsymbol{v}_g$. □

7.2 Decomposition of signed Borel measures

In what follows all measures will be over the Borel σ-algebra on $[0,1]$. Let $\mathbf{BM}^+$ be the set of all non-negative bounded measures. We first define a representation of $\mathbf{BM}^+$, and then extend this to the encompassing space $\mathbf{BM}$.

Let $\mathrm{Int} = \{(a,b),[0,b),(a,1],[0,1] : a,b \in \mathbb{Q} \text{ and } 0 \leq a \leq b \leq 1\}$ be the set of all open subintervals of $[0,1]$ which have rational endpoints. Define $\delta_m : \Sigma^\omega \to \mathbf{BM}^+$ by $\delta_m\langle p,q\rangle = \mu$ iff $\rho(q) = \mu([0,1])$ and p is an enumeration of all $(a,I) \in \mathbb{Q} \times \mathrm{Int}$ such that $a < \mu(I)$. The intuition is that δ_m is the weakest representation from which one can both compute the variation norm of a measure, and enumerate the corresponding $<$ relation.

Define a representation $\delta_{\mathbf{BM}}$ of $\mathbf{BM}$ by

$$\delta_{\mathbf{BM}}\langle p,q,r\rangle = \mu \text{ iff } \mu = \delta_m(p) - \delta_m(q) \text{ and } ||\mu||_m = \rho(r).$$

Notice that, similar to the representation $\delta_{\mathbf{BV}}$, we explicitly encode the norm of the object we are representing. Also note that while we encode μ as a difference of non-negative measures $\mu^+ - \mu^-$, the pair does not necessarily form the Jordan decomposition (which we required to be minimal).

For a continuous function $h \in C[0,1]$ and a function of bounded variation $g \in \mathbf{BV}$, the *Riemann-Stieltjes integral* $\int h\mathrm{d}g$ of h with respect to g is the unique real I such that for every $\varepsilon > 0$ there is $\delta > 0$ for which

$$\left| I - \sum_{i=1}^{n-1} h(t_{i+1})(g(t_{i+1}) - g(t_i)) \right| < \varepsilon$$

whenever $t_0 < \cdots < t_n$ is a partition of $[0,1]$ with $\max_i(t_{i+1} - t_i) < \delta$. The Riesz representation theorem asserts that every bounded linear functional $F \in C^*[0,1]$ can be represented by such a Riemann-Stieltjes integral (see, e.g., [16, Theorem 4.4.1]). Jafarikhah and Weihrauch have shown that an effective analogue of this theorem holds.

Theorem 7.2 (Effective Riesz representation theorem [12]). *The function $(F, ||F||) \mapsto g$ mapping every functional $F \in C^*[0,1]$ and its norm to the unique $g \in \mathbf{BV}$ such that $F(h) = \int h\mathrm{d}g$ is $([\delta_C \to \rho], \rho, \delta_{\mathrm{V}})$-computable.*

By the Riesz representation theorem there is a unique linear homeomorphism $\Psi_{\mathrm{VM}} : \mathbf{BV_0} \to \mathbf{BM}$ such that $\Psi_{\mathrm{VM}}(g) = \mu$ implies, for every $h \in C[0,1]$, that $\int h\mathrm{d}g = \int h\mathrm{d}\mu$. Let Ψ^+_{VM} be the restriction of Ψ_{VM} to non-decreasing functions. Given the appropriate representations, we show that the operator is computable.

Theorem 7.3 (Jafarikhah and Weihrauch [13]). *The operator Ψ^+_{VM} is $(\delta_{\mathrm{V}}, \delta_m)$-computable.*

Proof. Let $g = \delta_{\mathrm{V}}(p)$ be non-decreasing, and let A_p be its associated dense set. We compute a name of the measure μ corresponding to g. Note that p encodes a list of all pairs $(x, g(x))$ with $x \in A_p$. For $a, b \in \mathbb{Q}$ with $0 \leq a < b \leq 1$, one may computably enumerate all $c \in \mathbb{Q}$ such that $c < g(b') - g(a')$ for some $a', b' \in A_p$ with $a < a' < b' < b$. This determines the measure of the interval (a, b) since

$$\mu((a,b)) = \sup(g(b') - g(a')),$$

where the supremum is taken over all $a', b' \in \mathbb{Q}$ with $a < a' < b' < b$. Similarly, for rationals $a, b > 0$ we may computably enumerate all $c \in \mathbb{Q}$ with $c < \mu([0,b))$, and all $c \in \mathbb{Q}$ with $c < \mu((a,1])$. Moreover, $\mu([0,1])$ is explicitly provided as $g(1) = \rho(q_1)$. By printing q_1 and then alternating between the three preceding enumerations, one effectively enumerates all $(c, I) \in \mathbb{Q} \times \text{Int}$ such that $c < \mu(I)$. □

Operation	*Domain*	*Complexity*	*Theorem*
Ψ^+_{VM}	$\mathbf{BV_0}$	$(\delta_{\text{V}}, \delta_m)$	7.2
Ψ^+_{FV}	$C^*[0,1]$	$([\delta_{\text{C}} \to \rho], \delta_{\text{V}})$	7.5
Ψ^+_{MF}	$\mathbf{BM}$	$(\delta_m, [\delta_{\text{C}} \to \rho])$	7.6
$g \mapsto (g^+, g^-)$	$\mathbf{BV_0}$	$(\delta_{\mathbf{BV}}, [\delta_{\text{V}}, \delta_{\text{V}}])$	7.1
$F \mapsto (F^+, F^-)$	$C^*[0,1]$	$(\delta_{\text{CF}}, [[\delta_{\text{C}} \to \rho], [\delta_{\text{C}} \to \rho]])$	7.9
$\mu \mapsto (\mu^+, \mu^-)$	$\mathbf{BM}$	$(\delta_{\mathbf{BM}}, [\delta_m, \delta_m])$	7.10

Table 2: Various operations and the representations by which they are computable.

The operator Ψ_{VM} preserves Jordan decompositions: if (g^+, g^-) is the Jordan decomposition of $g \in \mathbf{BV_0}$, then

$$g^+ = \frac{\boldsymbol{v}_g + g}{2} \quad \text{and} \quad g^- = \frac{\boldsymbol{v}_g - g}{2}.$$

It follows that

$$\Psi_{\text{VM}}(g^+) = \frac{\boldsymbol{v}_{\Psi_{\text{VM}}(g)} + \Psi_{\text{VM}}(g)}{2} \quad \text{and} \quad \Psi_{\text{VM}}(g^-) = \frac{\boldsymbol{v}_{\Psi_{\text{VM}}(g)} - \Psi_{\text{VM}}(g)}{2},$$

where for a signed measure μ, $\boldsymbol{v}_\mu$ is its variation function (i.e., $\boldsymbol{v}_\mu(x) = \sup \sum_{I \in \pi} |\mu(I)|$, with the supremum taken over all partitions π of $[0, x]$ into finitely many intervals). This shows the following result.

Theorem 7.4. *If (g^+, g^-) is the (unique minimal) Jordan decomposition of g, then $(\Psi_{\text{VM}}(g^+), \Psi_{\text{VM}}(g^-))$ is the Jordan decomposition of $\Psi_{\text{VM}}(g)$.*

Note that since a Jordan decomposition consists of non-decreasing functions, we may replace $\Psi_{\text{VM}}(g^+)$ and $\Psi_{\text{VM}}(g^-)$ with the restrictions from Theorem 7.3. Thus if we can compute a Jordan decomposition of g, we can also compute a Jordan decomposition of $\Psi_{\text{VM}}(g)$.

7.3 Decomposition of continuous linear functionals

We introduce two representations for the space $C^*[0,1]$ of continuous linear functionals on $C[0,1]$. First note that there is a representation δ_C for the set $C[0,1]$, each function being named by a sequence of polygonal functions satisfying the effective Weierstraß condition. Then a functional $F : C[0,1] \to \mathbb{R}$ is continuous iff it is (δ_C, ρ)-continuous, see e.g. [28]. It follows that $[\delta_C \to \rho]$ is a representation of $C^*[0,1]$.

Recall that we equip $C^*[0,1]$ with the norm $||\cdot||$ defined by $||F|| = \sup\{F(h) : h \in C[0,1], ||h||_\infty \leq 1\}$. This norm is known to be $([\delta_C \to \rho], \rho_<)$-computable, but not $([\delta_C \to \rho], \rho)$-computable. This situation is analogous to what we studied in Section 3: with the more general TTE framework we see that the variation norm is $([\rho \to \rho], \rho_<)$-computable but not $([\rho \to \rho], \rho)$-computable. Indeed, by Theorem 3.1 there is a (ρ, ρ)-computable function whose total variation is not a computable real.

We proceed by defining a representation of $C^*[0,1]$ similar to the way that we defined one for $\mathbf{BV_0}$. That is, since the norm of a computable functional is not necessarily computable, we include it in the representation. Let $\delta_{CF} : \Sigma^\omega \to C'[0,1]$ be defined by

$$\delta_{CF}\langle p, q\rangle = F \text{ iff } [\delta_C \to \rho](p) = F \text{ and } \rho(q) = ||F||.$$

Hence $\langle p, q\rangle$ is a representation of $F \in C^*[0,1]$ if p is a $[\delta_C \to \rho]$-name of F, and q is a ρ-name of its norm.

Other instances of the Riesz representation theorem (e.g., the *Riesz-Markov-Kakutani representation theorem*) prove that there are linear homeomorphisms $\Psi_{\mathrm{FV}} : C^*[0,1] \to \mathbf{BV_0}$ and $\Psi_{\mathrm{MF}} : \mathbf{BM} \to C^*[0,1]$ such that

1. $\Psi_{\mathrm{FV}}(F) = g$ implies, for all $h \in C[0,1]$, that $F(h) = \int h \mathrm{d}g$, and
2. $\Psi_{\mathrm{MF}}(\mu) = F$ implies, for all $h \in C[0,1]$, that $\int h \mathrm{d}\mu = F(h)$.

Let Ψ^+_{FV} and Ψ^+_{MF} denote the restriction of the corresponding operators to non-negative functionals and measures, respectively.

Theorem 7.5. *The operator Ψ^+_{FV} is $([\delta_C \to \rho], \delta_{\mathrm{V}})$-computable.*

Proof. For any non-negative functional F one has $||F|| = F(\mathbb{1})$. Hence we may apply the effective Riesz representation theorem 7.2, which establishes the statement. □

Theorem 7.6 (Weihrauch [27])**.** *The operator Ψ^+_{MF} is $(\delta_m, [\delta_C \to \rho])$-computable.*

We also have invariance theorems for Ψ_{FV} and Ψ_{MF} analogous to Theorem 7.4. As we remarked earlier, this ensures that whenever we can effectively decompose an object, we can also effectively decompose its image under Ψ.

Theorem 7.7. *Let $F \in C^*[0,1]$.*

1. *If $F^+, F^- \in C^*[0,1]$ are non-negative and $F = F^+ - F^-$, then (F^+, F^-) is the Jordan decomposition of F iff $||F|| = ||F^+|| + ||F^-||$.*

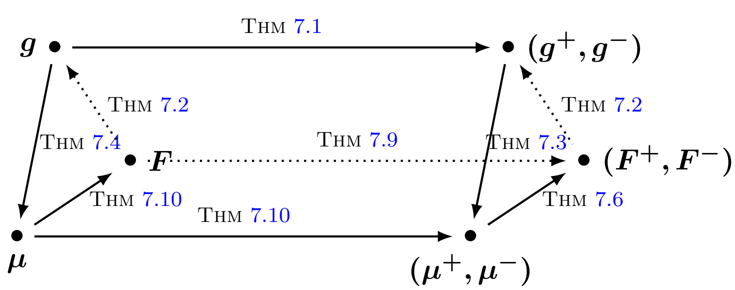
2. *If* (F^+, F^-) *is the Jordan decomposition of* F*, then* $(\Psi_{\mathrm{FV}}(F^+), \Psi_{\mathrm{FV}}(F^-))$ *is the Jordan decomposition of* $\Psi_{\mathrm{FV}}(F)$*.*

Note that item 2 follows directly from item 1 and the Riesz representation theorem: if (F^+, F^-) is the Jordan decomposition of F, then $||F|| = ||F^+|| + ||F^-||$, so $||\Psi_{\mathrm{FV}}(F)||_{\mathbf{BV}} = ||\Psi^+_{\mathrm{FV}}(F^+)||_{\mathbf{BV}} + ||\Psi^+_{\mathrm{FV}}(F^-)||_{\mathbf{BV}}$, whereby $(\Psi^+_{\mathrm{FV}}(F^+), \Psi^+_{\mathrm{FV}}(F^-))$ is the Jordan decomposition of $\Psi_{\mathrm{FV}}(F)$. The following theorem is similar.

Theorem 7.8. *Let* $\mu \in \mathbf{BM}$*.*

1. *If* $\mu^+, \mu^- \in \mathbf{BM}^+$ *and* $\mu = \mu^+ - \mu^-$*, then* (μ^+, μ^-) *is the Jordan decomposition of* μ *iff* $||\mu||_m = ||\mu^+||_m + ||\mu^-||_m$*.*
2. *If* (μ^+, μ^-) *is the Jordan decomposition of* μ*, then* $(\Psi_{\mathrm{MF}}(\mu^+), \Psi_{\mathrm{MF}}(\mu^-))$ *is the Jordan decomposition of* $\Psi_{\mathrm{MF}}(\mu)$*.*

Our aim now is to establish the relationships between the nodes in the following diagram. The label $\boldsymbol{g}$ represents an element of $\mathbf{BV_0}$, $\boldsymbol{\mu}$ an element of $\mathbf{BM}$, and $\boldsymbol{F}$ an element of $C^*[0,1]$. The corresponding pairs $(\boldsymbol{x}^+, \boldsymbol{x}^-)$ represent their respective Jordan decompositions, and arrows (dashed or undashed) from $\boldsymbol{x}$ to $\boldsymbol{y}$, indicate that from a name of $\boldsymbol{x}$, one can compute a name of $\boldsymbol{y}$.



Note that the converse arrows from the right triangle to the left triangle are trivial: from a representation of a decomposition one may easily recover a representation of the original function as the difference of the two components.

Theorem 7.9 (Jafarikhah and Weihrauch [13])**.** *The Jordan decomposition operator* $F \mapsto (F^+, F^-)$ *on* $C^*[0,1]$ *is* $(\delta_{\mathrm{CF}}, [[\delta_C \to \rho], [\delta_C \to \rho]])$*-computable.*

Proof. Given a δ_{CF}-name of F, by Theorem 7.2 we may compute a δ_{V}-name of $g \in \mathbf{BV_0}$ such that $F(h) = \int h dg$. We may subsequently compute a $(\delta_{\mathrm{V}}, \delta_{\mathrm{V}})$-name of its Jordan decomposition (g^+, g^-) by Theorem 7.1.

By Theorem 7.3 we can compute a (δ_m, δ_m)-name of a pair of non-negative measures (μ^+, μ^-) such that $\int h d\mu^{\pm} = \int h dg^{\pm}$. By Theorem 7.4, (μ^+, μ^-) is the Jordan decomposition of a signed measure μ. Finally, by Theorem 7.6 we may compute a $([\delta_C \to \rho], [\delta_C \to \rho])$-name of a pair (F^+, F^-) of non-negative functionals such that $F^{\pm}(h) = \int h d\mu^{\pm}$, and then by Theorem 7.7, (F^+, F^-) is the Jordan decomposition of F. □

Finally we establish computability of the Jordan decomposition operator for signed measures with respect to the representation $\delta_{\mathbf{BM}}$.

Theorem 7.10 (Jafarikhah and Weihrauch [13]). *The Jordan decomposition operator* $\mu \mapsto (\mu^+, \mu^-)$ *on* **BM** *is* $(\delta_{\mathbf{BM}}, [\delta_m, \delta_m])$*-computable.*

Proof. We first represent μ as a functional F, which will establish the arrow $\boldsymbol{\mu} \to \boldsymbol{F}$ in the diagram above. Suppose $\delta_{\mathbf{BM}}\langle p, q, r\rangle = \mu$. Then μ can be written as the difference $\mu^+ - \mu^-$ for some $\mu^+, \mu^- \in \mathbf{BM}^+$ such that $\delta_m(p) = \mu^+$, $\delta_m(q) = \mu^-$, and $||\mu||_m = \rho(r)$. By Theorem 7.6 we can compute names of $F^+ := \Psi_{\mathrm{MF}}(\mu^+)$ and $F^- := \Psi_{\mathrm{MF}}(\mu^-)$ which, by Theorem 7.8, form the Jordan decomposition of the functional $F := F^+ - F^-$. Since $||F|| = ||F^+|| + ||F^-|| = F^+(\mathbb{1}) + F^-(\mathbb{1}) = \int \mathbb{1}\mathrm{d}\mu^+ + \int \mathbb{1}\mathrm{d}\mu^- = ||\mu||_m$, we may also compute a δ_{CF}-name of F. A succession of computations in the order $\mu \mapsto F \mapsto g \mapsto (g^+, g^-) \mapsto (\mu^+, \mu^-)$ then yields the desired Jordan decomposition. □

We have shown that from a representation of an object which encodes its norm, we may compute a representation of its Jordan decomposition. Jafarikhah and Weihrauch also demonstrate that the converse holds: consider the following representations, where g ranges over $\mathbf{BV_0}$, μ ranges over **BM**, and F ranges over $C^*[0,1]$.

- $\gamma_{\mathrm{VJ}}\langle p, q\rangle = g$ iff $(\delta_{\mathrm{V}}(p), \delta_{\mathrm{V}}(q))$ is the Jordan decomposition of g,
- $\gamma_{\mathrm{MJ}}\langle p, q\rangle = \mu$ iff $(\delta_m(p), \delta_m(q))$ is the Jordan decomposition of μ,
- $\gamma_{\mathrm{FJ}}\langle p, q\rangle = F$ iff $([\delta_C \to \rho](p), [\delta_C \to \rho](q))$ is the Jordan decomposition of F.

Then $\delta_{\mathrm{CF}} \equiv \gamma_{\mathrm{FJ}}$, $\delta_{\mathbf{BV}} \equiv \gamma_{\mathrm{VJ}}$, and $\delta_{\mathbf{BM}} \equiv \gamma_{\mathrm{MJ}}$. It follows that representing an object from one of the three Banach spaces via some countable description together with its norm is effectively equivalent to representing it as its minimal Jordan decomposition.

Logic sometimes makes monsters.
Henri Poincaré (1854 - 1912)

References

[1] Jeremy Avigad, Edward T. Dean, and Jason Rute. Algorithmic randomness, reverse mathematics, and the dominated convergence theorem. *Annals of Pure and Applied Logic*, 163(12):1854–1864, 2012.

[2] Robert G. Bartle. *The elements of integration and Lebesgue measure.* Wiley-Interscience, 1995.

[3] Vladimir I. Bogachev. *Measure theory.* Vol. 1. Springer Science & Business Media, 2007.

[4] Vasco Brattka, Peter Hertling, and Klaus Weihrauch. A tutorial on computable analysis. In *New computational paradigms*, pages 425–491. Springer, 2008.

[5] Vasco Brattka, Joseph Miller, and André Nies. Randomness and differentiability. *Transactions of the American Mathematical Society*, 2015.

[6] Douglas S. Bridges. A constructive look at functions of bounded variation. *Bulletin of the London Mathematical Society*, 32(3):316–324, 2000.

[7] Adam R. Day. On the strength of two recurrence theorems. arXiv:1305.5858, 2013.

[8] Osvald Demuth. The differentiability of constructive functions of weakly bounded variation on pseudo numbers. *Comment. Math. Univ. Carolin*, 16(3):583–599, 1975.

[9] Rod Downey, Denis R. Hirschfeldt, Joseph Miller, and André Nies. Relativizing Chaitin's halting probability. *Journal of Mathematical Logic*, 5(02):167–192, 2005.

[10] Rod Downey and Denis R. Hirschfeldt. *Algorithmic randomness and complexity.* Springer, 2010.

[11] Cameron Freer, André Nies, and Frank Stephan. Algorithmic aspects of Lipschitz functions. *Computability*, 3(1), 2014.

[12] Tahereh Jafarikhah and Klaus Weihrauch. The Riesz representation operator on the dual of $C[0,1]$ is computable. *J. UCS*, 19(6):750–770, 2013.

[13] Tahereh Jafarikhah and Klaus Weihrauch. Computable Jordan decomposition of linear continuous functionals on $C[0,1]$. *Logical Methods in Computer Science (LMCS)*, 10(3), 2014.

[14] Camille Jordan. Sur la série de Fourier. *CR Acad. Sci. Paris*, 92(5):228–230, 1881.

[15] Ker-I Ko. *Complexity theory of real functions.* Springer Science & Business Media, 2012.

[16] Erwin Kreyszig. *Introductory functional analysis with applications.* Vol. 81. Wiley New York, 1989.

[17] Hong Lu and Klaus Weihrauch. Computable Riesz representation for the dual of $C[0,1]$. *Mathematical Logic Quarterly*, 53(4-5):415–430, 2007.

[18] Kenshi Miyabe, André Nies, and Jing Zhang. Using almost-everywhere theorems from analysis to study randomness. arXiv:1411.0732, 2014.

[19] André Nies. *Computability and randomness.* Vol. 51. Oxford University Press, 2009.

[20] André Nies. Differentiability of polynomial time computable functions. In *STACS*, pages 602–613, 2014.

[21] André Nies (editor). Logic Blog 2013. arXiv:1403.5719, 2014.

[22] Noopur Pathak, Cristóbal Rojas, and Stephen G. Simpson. Schnorr randomness and the Lebesgue differentiation theorem. *Proceedings of the American Mathematical Society*, 142(1):335–349, 2014.

[23] Jason Rute. Randomness, martingales and differentiability. 2011.

[24] Claus P. Schnorr. *Zufälligkeit und Wahrscheinlichkeit: Eine algorithmische Begründung der Wahrscheinlichkeitstheorie.* Springer, 1971.

[25] Stephen G. Simpson. *Subsystems of second order arithmetic.* Cambridge University Press, 2009.

[26] Stephen G. Simpson and Keita Yokoyama. A nonstandard counterpart of WWKL_0. *Notre Dame Journal of Formal Logic*, 52(3):229–243, 2011.

[27] Klaus Weihrauch. Computability on the probability measures on the Borel sets of the unit interval. *Theoretical Computer Science*, 219(1):421–437, 1999.

[28] Klaus Weihrauch. *Computable analysis: An introduction.* Springer Science & Business Media, 2012.

[29] Xiaokang Yu. Lebesgue convergence theorems and reverse mathematics. *Mathematical Logic Quarterly*, 40(1):1–13, 1994.

[30] Xizhong Zheng and Robert Rettinger. Effective Jordan decomposition. *Theory of Computing Systems*, 38(2):189–209, 2005.