Computability in Ordinal Ranks and Symbolic Dynamics

by

Linda Brown Westrick

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Mathematics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Theodore Slaman, Chair
Professor Marc Rieffel
Professor John Steel
Professor Christos Papadimitriou

Spring 2014

Computability in Ordinal Ranks and Symbolic Dynamics

Abstract

Computability in Ordinal Ranks and Symbolic Dynamics

by

Linda Brown Westrick

Doctor of Philosophy in Mathematics

University of California, Berkeley

Professor Theodore Slaman, Chair

Part 1: Computability in Ordinal Ranks

We analyze the computable part of three classical hierarchies from analysis and set theory. All results are expressed in the notation of Ash and Knight [1].

In the differentiability hierarchy defined by Kechris and Woodin [14], the rank of a differentiable function is an ordinal less than ω_1 which measures how complex it is to verify differentiability for that function. We show that for each recursive ordinal $\alpha > 0$, the set of Turing indices of computable $C[0,1]$ functions that are differentiable with rank at most α is $\Pi_{2\alpha+1}$-complete.

In the hierarchy defined by the transfinite process of Denjoy integration, the rank of a Denjoy-integrable function f is defined as the ordinal $\alpha < \omega_1$ at which the process of integrating f terminates. We show that the set of Turing indices of computable $C[0,1]$ functions of the form $\int f$, where f is Denjoy-integrable of rank 1, is Π_3-complete; and that for any recursive ordinal $\alpha > 1$, the set of indices of computable $C[0,1]$ functions of the form $\int f$, where f has rank at most α, is $\Sigma_{2\alpha}$-complete.

Finally, we give a new proof that for any recursive ordinal $\alpha > 1$, the set of indices for computable trees in $2^{<\omega}$ with no dead ends which vanish after α applications of the Cantor-Bendixson derivative is $\Sigma_{2\alpha}$-complete. This result, in a different notation, was originally due to Lempp [17].

The major contribution of Part 1 is a general theorem which lies at the core of all three results. We introduce the *limsup rank* which assigns an ordinal to each well-founded tree in Baire space. Trees of limsup rank α are seen to correspond in a computable way to objects of rank α in each of the three contexts discussed above. For each recursive ordinal $\alpha > 0$, the theorem provides a one-one reduction from $\emptyset_{(2\alpha)}$ to the set of Turing indices of trees of limsup rank at most α, where $\emptyset_{(\alpha)}$ is the canonical Σ_α complete set.

Part 2: Computability in Symbolic Dynamics

We consider various questions in the intersection of computability theory as applied to subshifts. In particular, we consider three subshift invariants: entropy, Medvedev degree, and effective dimension spectrum. The last one is a new invariant. We explore these invariants in the context of important examples of subshifts: density-r subshifts, shifts of finite type, subshifts consisting of shift-complex sequences, and Medvedev subshifts.

Building on the work of [13, 25, 28, 32], we show that the entropy and the Medevedev degree are independent invariants. To do this we construct subshifts with every combination of entropy and Medvedev degree that is not immediately prohibited, in both one and two dimensions. When the entropy is right-r.e. and the Medvedev degree is Π^0_1, the subshifts we produce are Π^0_1 in the one-dimensional case, and shifts of finite type in the two-dimensional case. When the entropy is in $[0, 1)$, we accomplish this using an alphabet with only two symbols.

We introduce the *dimension spectrum* of a subshift X as $\{\dim(x) : x \in X\}$, where dim is the effective dimension, and work towards a characterization of the possible dimension spectra. By [32], every dimension spectrum has a top element. Conditions are given under which the dimension spectrum of X is the interval $[0, \mathrm{ent}(X)]$, and examples are given where the dimension spectrum is bounded away from 0. We show that the dimension spectrum of a one-dimensional minimal subshift has a least element, and find the dimension spectrum of the minimal subshift from [2].

To my parents.

Contents

List of Figures

Acknowledgments

This dissertation would not have been possible without the help of many mentors, colleagues, students, friends, family and organizations. Here I would like to highlight the positive effects that many members of my community have had on this process.

First and foremost, I would like to thank my advisor, Theodore Slaman. He gave me freedom to work on whatever I found the most interesting, and offered his relevant advice and experience on each topic. His influence has shaped my mathematical values, taste and expectations, all for the better. He shared his sabbatical with his students, taking me and my cohort to the University of Buenos Aires for a semester, during which I leveled up mathematically and met many interesting people, many of whom also had a positive effect on this dissertation. His sense of humor, kind attitude, reasonable expectations, descriptions of life as a mathematician, and way of explaining his own mathematical thought process all combined to make working with him both enjoyable and productive.

My dissertation committee members, Theodore Slaman, Marc Rieffel, John Steel, and Christos Papadimitriou, have been always supportive of me and yet happy to let me choose my own directions. It has been a pleasure doing my graduate studies under their supervision.

I have received a great deal of excellent advice and mentoring from various people, from which this work has also benefitted. Rod Downey taught me to pick up the pace in my research style and encouraged me to publish the work in Chapter 1. When I was in the doldrums on that topic, he and Gerald Sussman and Jan Reimann listened to my interests and pointed me in the general direction of chaos theory, and in particular the papers of Simpson and Hochman & Meyerovitch that inspired Chapter 2. Noam Greenberg immediately saw the notation choice that would bring the results of Chapter 1 into clarity. Peter Cholak and Alexander Kechris encouraged me to work on the Denjoy rank, without which Chapter 1 would have felt incomplete. Stephen Simpson let me invite myself to Penn State and gave me a week of his mentorship and a great jump-start into the area of subshifts for Chapter 2. He suggested the concept of the effective dimension spectrum, and he also asked critical questions about the material of Chapter 1 which sharpened my understanding of the subject. Grigory Barenblatt, Joseph Miller, Russell Miller and Antonio Montalbán were wonderful people to talk to for questions on a variety of topics. I am grateful to these people and to all the other members of the research community, both in person and on MathOverflow, who have given me so many interesting things to think about.

The other students of my advisor have been another source of support. I would like to thank my cohort, Kelty Allen, Ian Haken, Ian Herbert, and Greg Igusa, for good times, for many interesting conversations about their work, and also for helping me survive in foreign countries. Johanna Franklin, Andrew Marks and Mariya Soskova (the first two were also students of Slaman) have been great role models and givers of advice about graduate school and beyond.

I'm also grateful for the academic and industry mentorship I've received from Eric Jonas, Gregory Marton, and Sasen Cain.

Throughout my time as a graduate student, I've been supported financially through a variety of sources, including a Siemens Foundation scholarship, a P.E.O. Scholar Award, a GSR through Theodore Slaman's NSF grant, and support from the UC-Berkeley Math Department during the semesters when I was a GSI (and one semester when I was not). I would like to thank all these organizations, as well as the University of Buenos Aires, the National University of Singapore, the John Templeton Foundation, the Association for Symbolic Logic, and the Isaac Newton Institute.

I'm grateful to the staff of the Math Department for all that they do, solving so many problems that I know about and surely even more that I do not know about. Thank you to Barbara Waller for knowing everything about everything, to Judie Filomeo for making my day brighter whenever we cross paths, and to Barbara Peavy for coordinating my experience as a GSI and making everything run smoothly.

I have been inspired by the vision of Hugo Ramirez and all my students in the Professional Development Program. I am honored to have had the experience of working with them in the last few years.

Some people contributed to the dissertation quite directly. Theodore Slaman provided a simpler proof of Lemma 1.2.2. Ian Haken and an anonymous referee carefully read a draft of the paper that became Sections 1.1-1.3, and provided valuable suggestions. Yael Degany taught me how to use new parts of tikz with examples tailored to the kinds of figures I proposed to make. Sai spent many hours helping me typeset my edits in the packed days leading up to the deadline. I would like to thank all these people for their help and generosity. It goes without saying that any errors, points of confusion, cosmetic or expository flaws are entirely my own responsibility.

Finally, I was supported throughout the process by the creativity and companionship of all my friends, my family and housemates, the community at FCCB, and Sai. Thank you all for seeing me through.

Chapter 1

Computability in Ordinal Ranks

In this chapter, we define an alternate rank on well-founded trees called the *limsup rank*, and use it to analyze three hierarchies from analysis and set theory: the Cantor-Bendixson rank, the Kechris-Woodin rank, and the Denjoy rank. All three hierarchies have, at their core, the same descriptive difficulty which is captured by the notion of the limsup rank. In Section 1.1 we introduce notation and concepts common to all the subsequent sections. In Section 1.2 we define the limsup rank on trees in $\mathbb{N}^{<\mathbb{N}}$ and prove the main theorem of this chapter, establishing the descriptive complexity of its initial segments. In this section we also find the descriptive complexity of the initial segments of the Cantor-Bendixson rank as a corollary of the main theorem about the limsup rank. This corollary was originally implicit in [17]. In Section 1.3 we introduce the Kechris-Woodin rank as first defined in [14], and also find the descriptive complexity of its initial segments as another consequence to the main theorem.[1] Finally, in Section 1.4 we introduce Denjoy integration from the constructive perspective, noting how this process may be naturally understood in terms of a rank function on the indefinite integrals which result from Denjoy integration. As in the previous sections, we use the main theorem to quantify the descriptive complexity of this hierarchy.

1.1 Preliminaries

This section provides background, essential definitions, methods previously used to construct functions of different ranks, and corollaries that are straightforward effectivizations of arguments in the literature. In Section 1.1 we establish some notation and review the basic facts about computable $C[0,1]$ functions. In Section 1.1 we introduce the recursive ordinals and use them to define Σ_α-completeness.

[1]Most of the work in Sections 1.1, 1.2 and 1.3 will also appear in the Journal of Symbolic Logic under the title "A lightface analysis of the differentiability rank", copyright held by the Association for Symbolic Logic.

Basic notions and encoding $C[0,1]$ functions

We use ϕ_e to denote to the eth Turing functional, and W_e refers to the domain of ϕ_e. We identify subsets $X \subseteq \mathbb{N}$ with their characteristic functions $X \in 2^\omega$. The jump of $X \in 2^\omega$ is written X', and the nth jump of X is written $X^{(n)}$. Turing reducibility is denoted by $\leq_T$ and one-reducibility by $\leq_1$. We use $\langle n_1, \ldots, n_k\rangle$ to denote a single integer which represents the tuple $(n_1, \ldots, n_k)$ according to some standard computable encoding. If $\tau = \langle m_1, \ldots, m_r\rangle$ and $\sigma = \langle n_1, \ldots, n_k\rangle$, let $\tau^\frown\sigma$ denote $\langle m_1, \ldots, m_r, n_1, \ldots, n_k\rangle$. If $T \subseteq \mathbb{N}^{<\mathbb{N}}$ is a tree, let T_n denote $\{\sigma : \langle n\rangle^\frown\sigma \in T\}$, the nth subtree of T. If T is well-founded, $|T|$ denotes its rank.

We identify the computable functions with the computable subsets of $\mathbb{N}$ that encode those functions. Following [14], all our functions are real-valued with domain $[0,1]$. For the encoding we use Simpson's definition from [33] because this encoding makes it straightforward to determine the degree of unsolvability of various statements. For example, we will observe that "ϕ_e encodes a computable $C[0,1]$ function" is Π_2. However, the exact details of the Simpson encoding are not needed beyond this section, and any of the many equivalent definitions for a computable real-valued function can be safely substituted.

In the following definition, $(a,r)\Phi(b,s)$ is shorthand for $\exists n((n,a,r,b,s,) \in \Phi)$, and $(a,r) < (a',r')$ means that $|a-a'| + r' < r$. The idea is that $(a,r)\Phi(b,s)$ should mean that $f(B(a,r)) \subseteq \overline{B(b,s)}$.

Definition 1.1.1. *A code for a continuous partial function f from $[0,1]$ to $\mathbb{R}$ is a set of quintuples $\Phi \subseteq \mathbb{N} \times \mathbb{Q} \cap [0,1] \times \mathbb{Q}^+ \times \mathbb{Q} \times \mathbb{Q}^+$ which satisfies:*

1. *if $(a,r)\Phi(b,s)$ and $(a,r)\Phi(b',s')$ then $|b-b'| \leq s+s'$*
2. *if $(a,r)\Phi(b,s)$ and $(a',r') < (a,r)$, then $(a',r')\Phi(b,s)$*
3. *if $(a,r)\Phi(b,s)$ and $(b,s) < (b',s')$, then $(a,r)\Phi(b',s')$.*

This set Φ is coded as a subset of $\mathbb{N}$ using the standard encoding. Some important facts can be seen from this definition. Firstly, it is Π_2 to check whether a given code $X \subseteq \mathbb{N}$ satisfies the above properties. Secondly, the codes satisfying the above might not represent total functions. That is, for some points x in $[0,1]$ and some ε there may not be an a, r, b such that $|x-a| < r$ and $(a,r)\Phi(b,\varepsilon)$. However if the code does represent a total function then, by the compactness of $[0,1]$, for each ε there is a finite set $\{(a_i, r_i, b_i, s_i)\}_{i<p}$ such that the (a_i, r_i) cover $[0,1]$ and for each i, $s_i \leq \varepsilon$ and $(a_i, r_i)\Phi(b_i, s_i)$. Therefore, "ϕ_e encodes a $C[0,1]$ function" is a Π_2 statement: ϕ_e is total, and the corresponding code satisfies Definition 1.1.1, and for all ε there is a finite cover as described above. Let f_e denote the $C[0,1]$ function encoded by ϕ_e. Note that any function encoded using this convention is, by necessity, continuous.

If f is any computable $C[0,1]$ function and z and x any rational numbers, the statement $f(x) > z$ is Σ_1, because $f(x) > z$ if and only if there are δ, b and ε such that $(x,\delta)\Phi(b,\varepsilon)$ and $b - \varepsilon > z$.

We will also freely make use of the fact that addition, multiplication, division, and composition of computable functions are computable. For details we refer the reader to [33].

Kleene's $\mathcal{O}$ and the notion of a Σ_α-complete set

Kleene's $\mathcal{O}$ is a way of encoding the recursive ordinals as natural numbers. First one defines a relation $<_\mathcal{O}$ on $\mathbb{N}$ as the least relation closed under the following properties:

1. $1 <_\mathcal{O} 2$.
2. If $a <_\mathcal{O} b$ then $b <_\mathcal{O} 2^b$.
3. If $\phi_e(n)$ is total and $\phi_e(n) <_\mathcal{O} \phi_e(n+1)$ for all n, then $\phi_e(n) <_\mathcal{O} 3 \cdot 5^e$ for all n.
4. If $a <_\mathcal{O} b$ and $b <_\mathcal{O} c$ then $a <_\mathcal{O} c$.

The field of this relation is called Kleene's $\mathcal{O}$. One can show that $\mathcal{O}$ is a Π^1_1-complete set, that $<_\mathcal{O}$ is well-founded, and for each $a \in \mathcal{O}$, the set $\{b : b <_\mathcal{O} a\}$ is well ordered and computably enumerable. (See [29] for details). Therefore, for each $a \in \mathcal{O}$ there is a well-defined ordinal $|a|_\mathcal{O} = \mathrm{ot}(\{b : b <_\mathcal{O} a\})$. In this situation a is called an *ordinal notation* for $|a|_\mathcal{O}$. If an ordinal has an ordinal notation in $\mathcal{O}$, it is called a *constructive ordinal*. Note that there are infinitely many ordinal notations corresponding to each constructive ordinal $\alpha \geq \omega$. There are only countably many constructive ordinals and these form an initial segment of the ordinals. The least nonconstructive ordinal is called ω_1^{CK}, "the ω_1 of Church and Kleene".

We will use the fact that it is computable to add ordinal notations in a way that is consistent with their corresponding ordinals.

The constructive ordinals have an important equivalent characterization. They are exactly the ranks of the recursive well-founded relations. This will be used to establish that the differentiability ranks of the computable functions are the constructive ordinals.

We recall the arithmetical hierarchy for $n < \omega$. A set X is said to be Σ_n (respectively Π_n) if $X \leq_1 \emptyset^{(n)}$ (respectively $\overline{\emptyset^{(n)}}$), and X is Σ_n-complete if $X \equiv_1 \emptyset^{(n)}$ (and similarly for Π_n-completeness).

The ordinal notations provide a natural way to extend the notion of the Turing jump through the ordinals less than ω_1^{CK}, giving rise to the hyperarithmetical hierarchy. Define $H_1 = \emptyset$, $H_{2^b} = (H_b)'$, and $H_{3 \cdot 5^e} = \{\langle x, n\rangle : x \in H_{\phi_e(n)}\}$. Spector [35] showed that if $|a|_\mathcal{O} = |b|_\mathcal{O}$, then $H_a \equiv_T H_b$. Therefore, $H_{2^a} \equiv_1 H_{2^b}$, and thus there is a well-defined notion of one-reducibility and completeness at the successor levels. We define the notions of Σ_α and Π_α for infinite ordinals following [1]:

Definition 1.1.2. *Let $\alpha < \omega_1^{CK}$ be an infinite ordinal and let $X \in 2^\omega$. Then X is Σ_α if $X \leq_1 H_{2^a}$ for any a such that $|a|_\mathcal{O} = \alpha$, and X is Σ_α-complete if $X \equiv_1 H_{2^a}$ for any such a. The Π_α and Π_α-complete sets are defined similarly.*

Note that using this definition, $(\emptyset^{(\omega)})'$ is a Σ_ω-complete set. There is a conflicting notational convention, found in [34, pg. 259], in which $(\emptyset^{(\omega)})'$ is classified $\Sigma_{\omega+1}$-complete, and the symbol Σ_ω is not defined. We prefer the notation of [1] because it is more consonant with the definition of the rank function. As will be seen, to determine whether the core rank-ascertaining process terminates at a limit stage, it is necessary to use a quantification over the results of the previous stages, not merely a unified presentation of them.

We fix a particular (but arbitrary) path $\mathcal{P}$ through $\mathcal{O}$ and define $\emptyset^{(\alpha)}$ for each $\alpha < \omega_1^{CK}$ by $\emptyset^{(\alpha)} = H_a$, where a is the unique $a \in \mathcal{P}$ such that $|a|_{\mathcal{O}} = \alpha$. (We call $\mathcal{P}$ a path through $\mathcal{O}$ if $\mathcal{P} \subseteq \mathcal{O}$ is $<_{\mathcal{O}}$-linearly ordered and contains an ordinal notation for each $\alpha < \omega_1^{CK}$.)

Because $\emptyset^{(\alpha+1)}$ is the canonical Σ_α-complete set when $\alpha > \omega$, we follow [7] in defining

$$\emptyset_{(\alpha)} = \begin{cases} \emptyset^{(\alpha)} & \text{if } \alpha < \omega \\ \emptyset^{(\alpha+1)} & \text{if } \alpha \geq \omega \end{cases}$$

so that $\emptyset_{(\alpha)}$ is always the canonical Σ_α-complete set. In addition, we identify α with the relevant ordinal notation, which in this paper is the notation $a \in P$ such that $H_a = \emptyset_{(\alpha)}$. (Thus infinite α are identified with the a such that $|a|_{\mathcal{O}} = \alpha + 1$). This choice greatly simplifies the presentation in Section 1.3 by removing the need to explicitly and constantly deal with the non-uniformity between the finite and the infinite discussed here.

As we are in the business of establishing the Π_α-completeness of various sets, we will construct reductions to and from $\emptyset_{(\alpha)}$ for various values of α. All of our reductions will be either to some $\emptyset_{(\alpha)}$ or to index sets. Since all sets of these kinds permit padding, it will suffice to find many-one reductions, and this is what we do. We use the technique of effective transfinite recursion which is described in detail in [29]. For our purposes it can be stated as follows:

Theorem 1. *Let $I : \omega \to \omega$ be a recursive function, and suppose for all $e \in \mathbb{N}$ and all $x \in \mathcal{P}$, if $\phi_e(y)$ is defined for all $y \in \mathcal{P}$ such that $y <_{\mathcal{O}} x$, then $\phi_{I(e)}(x)$ is defined. Then for some c, $\phi_c(x)$ is defined for all $x \in \mathcal{P}$, and $\phi_c(x) = \phi_{I(c)}(x)$ for any x on which either converges.*

When we use this technique, the function I will be defined only implicitly.

1.2 A combinatorial theorem with applications to the Cantor-Bendixson rank

In this section, we define a rank on well-founded trees, the *limsup rank*, whose structure reflects the topological difficulties inherent in measuring Cantor-Bendixson rank, Kechris-Woodin rank, and Denjoy rank in the coming sections. In Section 1.2 we prove some unsurprising but needed tools. In Section 1.2 we give a reduction from canonical $\Sigma_{2\alpha}$-complete sets to trees of an appropriate rank. As a result we see that for all constructive $\alpha > 0$, $\{e : |T_e|_{ls} \leq \alpha\}$ is $\Sigma_{2\alpha}$-complete, where T_e is the eth computable tree in Baire space, represented as the set of its initial segments.

Definition 1.2.1. *For a well-founded tree $T \in \mathbb{N}^{<\mathbb{N}}$, define the limsup rank of the tree by*

$$|T|_{ls} = \max\left(\sup_n |T_n|_{ls}, (\limsup_n |T_n|_{ls}) + 1\right),$$

if T is nonempty, and $|T|_{ls} = 0$ if T is empty.

Note that reordering the subtrees does not change the limsup rank of the tree. A node can have a rank higher than all its children in one of two situations: either there is no child of maximal rank, or there are infinitely many maximal rank children. In the next sections, we will see that this mechanism corresponds exactly to the mechanism for constructing functions of increasing Cantor-Bendixson rank, differentiability rank, or Denjoy rank. Note that $|T|_{ls}$ is always a successor.

Proposition 1.2.1. *For all constructive $\alpha > 0$, $\{e : |T_e|_{ls} \leq \alpha\}$ is $\Sigma_{2\alpha}$.*

Proof. We have $T = \emptyset$ if and only if it fails to contain the root, so $\{T : |T|_{ls} \leq 0\}$ is Σ_0, under the Π_1 assumption that T is actually a tree. Assuming that $\{T : |T|_{ls} \leq \alpha\}$ is $\Sigma_{2\alpha}$ uniformly in T and α, we now examine the claim for $\alpha+1$. We claim that $|T|_{ls} \leq \alpha+1$ if and only if the set of $\sigma \in T$ for which $|T_\sigma|_{ls} > \alpha$ is finite. If this set is finite, then in particular $\limsup_n |T_n|_{ls} \leq \alpha$, and because $|T_n|_{ls} > \alpha + 2$ implies that there must be infinitely many σ extending n with $|T_n|_{ls} = \alpha+1$, we have $\sup_n |T_n|_{ls} \leq \alpha+1$ as well. (If the set is finite, then T must be well-founded, for none of the infinitely many nodes along an infinite path are ever assigned a rank.) On the other hand, if the set of σ for which $|T_\sigma|_{ls} > \alpha$ is infinite, then either T is not well-founded (in which case by convention $|T|_{ls} = \infty$) or if T is well-founded, then there must be a σ for which T_σ has infinitely many children satisfying $|T_{\sigma n}|_{ls} \geq \alpha + 1$, in which case $|T_\sigma|_{ls} \geq \alpha + 2$, and thus $|T|_{ls} \geq \alpha + 2$. Therefore,

$$|T|_{ls} \leq \alpha + 1 \iff \exists \sigma_1 \ldots \sigma_N \in T \forall \sigma \in T \setminus \{\sigma_i\}_{i \leq N} |T_\sigma|_{ls} \leq \alpha$$

Therefore, under the assumption that $\{T : |T|_{ls} \leq \alpha\}$ is $\Sigma_{2\alpha}$ uniformly in T and α for constructive α, we have $\{T : |T|_{ls} \leq \alpha+1\}$ is $\Sigma_{2\alpha+2}$ uniformly in T and α. In the limit case,

$$|T|_{ls} \leq \lambda \iff |T|_{ls} < \lambda \iff \exists \alpha < \lambda |T|_{ls} \leq \alpha.$$

The matrix of the above is uniformly computable in $\emptyset^{(\lambda)}$, so the statement is $\Sigma_\lambda = \Sigma_{2\lambda}$, uniformly in T and λ. □

In the next sections we show that this descriptive complexity is exact.

Technical Lemmas

The purpose of the next two lemmas is to specify exactly how to strip two quantifiers off most Π_α facts in a particularly nice way, a way which will be useful for the main argument which is coming up in Theorem 2. The lemmas are surely known, but proofs are provided for completeness.

The first lemma takes an arbitrary $\Pi_{\alpha+2}$ fact and rewrites it in a nice form, with unique witnesses and stable evidence. In the process, two computable reduction functions g_0 and g_s are defined which will be used in Theorem 2.

Lemma 1.2.2. *For any $\Pi_{\alpha+2}$ predicate P(x), there is a Π_α predicate $R(x, z, y)$ such that*

1. $P(x) \iff \forall z \exists y R(x, z, y)$
2. $R(x, z, y_1) \wedge R(x, z, y_2) \implies y_1 = y_2$ *(R has unique witnesses)*
3. *For* $z_1 < z_2$, $\neg\exists y R(x, z_1, y) \implies \neg\exists y R(x, z_2, y)$ *(R has stable evidence)*
4. $R(x, z, y) \implies z < y$

Proof. We may as well assume that $P(x)$ is "$x \notin \emptyset_{(\alpha+2)}$". For the case $\alpha = 0$, we define R using a computable, total $\{0, 1\}$-valued function g_0, and set $R(x, z, y) \iff g_0(x, z, y) = 1$.

Let e be a Π_2 index for $\emptyset''$, i.e. ϕ_e is total and $x \notin \emptyset'' \iff \forall v \exists w[\phi_e(x, v, w) = 1]$. Define

$$g_0(x, z, y) = \begin{cases} 1 & \text{if } y > z \text{ and for all } v < z \text{ there is } w < y \text{ such that} \\ & \phi_e(x, v, w) = 1 \text{ and } y \text{ is least such that this is true} \\ 0 & \text{otherwise.} \end{cases}$$

One may check that four conditions on R are satisfied.

For the case $\alpha > 0$, we define R using a computable, total function g_s and set $R(x, z, y) \iff g_s(x, z, y) \notin \emptyset_{(\alpha)}$. The construction that defines g_s uses movable markers to build Π_α sets with at most one element. At any moment there is one particular element being held which is linked to a potential least-witness, and this element will be held for as long as that witness seems viable.

Let e be a universal Π_3 index, i.e. ϕ_e^X is total for all X and

$$x \notin X''' \iff \forall u \exists v \forall w[\phi_e^X(x, u, v, w) = 1].$$

The intended oracle X is an inverse jump of $\emptyset_{(\alpha)}$, so that $X' = \emptyset_{(\alpha)}$ and $X''' = \emptyset_{(\alpha+2)}$. But the claims of the lemma also hold for an arbitrary X when we let $P(x)$ be $x \notin X'''$ and $R(x, z, y)$ be $g_s(x, z, y) \notin X'$.

Define $W^X_{g(x,z)}$ in stages according to the following dynamic process. At stage $s = 0$, let $W^X_{g(x,z),0} = \{n : n \leq z\}$, and let $t_1 = 0$. For each $s > 0$, let y^0_s and y^1_s be respectively the smallest and second smallest elements of $\overline{W^X_{g(x,z),s-1}}$. Check whether $(\forall u < z)(\exists v < t_s)(\forall w < s)[\phi^X_e(x,u,v,w) = 1]$. If this is so, put y^1_s into $W^X_{g(x,z),s}$, and set $t_{s+1} = t_s$. If this is not so, put y^0_s into $W^X_{g(x,z),s}$, and set $t_{s+1} = t_s + 1$.

Then define

$$W^X_{g_s(x,z,y)} = \begin{cases} \mathbb{N} & \text{if } y \in W^X_{g(x,z)} \\ \emptyset & \text{otherwise.} \end{cases}$$

This has the effect that $g_s(x,z,y) \in X' \leftrightarrow y \in W^X_{g(x,z)}$.

Now let us verify the claims of the lemma, in the more general case where $P(x)$ is $x \notin X'''$ and $R(x,z,y)$ is $g_s(x,z,y) \notin X'$.

First we address the second claim, that R has unique witnesses. For a given x, z, X, let us verify that there is at most one y such that $g_s(x,z,y) \notin X'$. Suppose y^0_s does not stabilize in the construction above. Then $\overline{W^X_{g(x,z)}}$ does not have a smallest element, so it is empty, so $W^X_{g(x,z)} = \mathbb{N}$. On the other hand if y^0_s stabilizes, then let s_0 be such that for all $s > s_0$, $y^0_{s_0} = y^0_s$. Then for all $s > s_0$, it must be that y^1_s is put into $W^X_{g(x,z),s}$, so $\overline{W^X_{g(x,z)}} = \{y^0_{s_0}\}$. Thus in either case, $W^X_{g_s(x,z,y)} = \mathbb{N}$ for all but at most one y, so $g_s(x,z,y) \in X'$ for all but at most one y.

For the first claim, suppose that $x \notin X'''$. This is true if and only if $\forall u \exists v \forall w[\phi^X_e(x,u,v,w) = 1]$ In that case, for all z, in the construction of $W^X_{g(x,z)}$, we see that t_s stabilizes, because there is a t for which $(\forall u < z)(\exists v < t)(\forall w)[\phi^X_e(x,u,v,w) = 1]$. And conversely, if t_s stabilizes for each z, then $x \notin X'''$. We have $\lim_s t_s$ exists exactly when $\lim_s y^0_s$ exists, since they always change together. And $\lim y^0_s = y$ exists exactly when $\overline{W^X_{g(x,z)}} = \{y\}$, which is equivalent to saying $g_s(x,z,y) \notin X'$. Thus $x \notin X'''$ if and only if $g_s(x,z,y) \notin X'$.

For the third claim, note that if $z_1 < z_2$ then $\lim_s t_s(z_1) \leq \lim_s t_s(z_2)$ where $t_s(z)$ refers to the t_s-values associated to the construciton of $W^X_{g(x,z)}$. Thus, if $W^X_{g(x,z_1)} = \emptyset$, then $W^X_{g(x,z_2)} = \emptyset$ as well.

Finally, for the last claim, if $y \in \overline{W^X_{g(x,z)}}$ then $y > z$ because $\{n : n \leq z\} \subseteq W^X_{g(x,z)}$ from the outset.

□

The next lemma explicitly splits up the queries to a $\emptyset^{(\lambda)}$ oracle that occur during the evaluation of a Π_λ question. The goal is to isolate the parts of the computation that can be done using a weaker oracle. In the proof we define a function g_l which will be used in Theorem 2.

Lemma 1.2.3. *Let λ be a limit ordinal, given as a uniform supremum $\lambda = \sup_n \beta_n$. For any Π_λ predicate $P(x)$ there is a sequence of predicates R_n such that*

$$P(x) \iff \bigwedge_n R_n(x)$$

where R_n is $\Pi_{2\beta_n}$ for each n. Furthermore, the R_n are uniformly computable from P and λ.

Proof. We may assume that each β_n is a successor ordinal, and that $P(x)$ is "$x \notin \emptyset_{(\lambda)}$". Now we define R_n by specifying a computable function g_l below and letting $R_n(x) \iff g_l(x, \lambda, n) \notin \emptyset_{(2\beta_n)}$.

Uniformly in any pair of constructive ordinals $\alpha < \beta$, there is a reduction from $\emptyset_{(\beta)}$ to $\emptyset_{(\alpha)}$. (See for example [1, Lemma 5.1].) And any standard encoding will have the property that $\langle z, n\rangle \geq n$. Therefore, $\emptyset^{(\lambda)} \upharpoonright n$ is uniformly computable from λ, n and $\emptyset_{(\beta_n)}$, in the sense that there is a partial recursive function $\sigma(\lambda, n, X)$ which halts and returns $\emptyset^{(\lambda)} \upharpoonright n$ if $X = \emptyset_{(\beta_n)}$.

Define $g(x, \lambda, n)$ by

$$W^X_{g(x,\lambda,n)} = \begin{cases} \emptyset & \text{if } \phi^{\sigma(\lambda,n,X)}_{x,n}(x)\uparrow \\ \mathbb{N} & \text{otherwise.} \end{cases}$$

Suppose that $x \notin \emptyset_{(\lambda)}$. This is true if and only if

$$\phi^{\emptyset^{(\lambda)}}_x(x)\uparrow \iff \forall n \phi^{\emptyset^{(\lambda)}\upharpoonright n}_{x,n}(x)\uparrow \iff \forall n[g(x,\lambda,n) \notin \emptyset_{(\beta_n+1)}].$$

Define $g_l(x, \lambda, n)$ so that $g_l(x, \lambda, n) \notin \emptyset_{(2\beta_n)} \iff g(x, \lambda, n) \notin \emptyset_{(\beta_n+1)}$. (Since β_n is a successor ordinal, $\beta_n + 1 \leq 2\beta_n$.) □

Recognizing trees of limsup rank α is $\Sigma_{2\alpha}$-hard

The following lemma contains the heart of the reduction. Given a Π_α fact, we must build a tree of the appropriate limsup rank. Each node of this tree will be associated with a finite set of Π_β assertions for different ordinals β. The behavior of the subtree below a node is as follows. If all the assertions are true, then the rank of the subtree should be large, on the order of the largest β from the set of assertions. But if some Π_β assertion is false, then the rank of the subtree should be small, of a similar height to that β.

The node achieves this behavior by selecting which assertions should be given to each of its child-nodes. The collection of Π_β assertions, if all true, could be viewed as having a generalized Skolem function which covers the first two quantifiers of every assertion in the collection. The previous two lemmas will ensure that this Skolem function, if it exists, is unique. The children try to guess fragments of this unique Skolem function, and each child is given a set of assertions which explore the fragment of the Skolem function that the child provided. The previous two lemmas will ensure that if infinitely many children can correctly guess a fragment of the generalized Skolem function, then (1) all the assertions of the parent are true and (2) these children, having guessed all the right witnesses, will achieve high rank.

On the other hand, if some assertion was false at the level of the parent node, then since the guesses are only fragments, finitely many children will still come up with lucky guesses which give them a pile of true assertions, some of which could be very large compared with the false assertion the parent had. Therefore, the children also each re-evaluate all of the non-maximal assertions from their parent node; this damps the sup of the ranks of the children.

As for damping the limsup, cofinitely many children will automatically dampen down their own ranks through exploring the false assertions generated by their Skolem guesses, which were doomed guesses in a situation where in fact no witnesses existed. Thus the limsup of the ranks of the children is damped. There is a subtlety here. If the limsup is supposed to be damped below some limit ordinal, it is not enough that each child get below that ordinal individually. They have to obey a common bound. That is why, in step (5) below, when α_i is a limit ordinal, M_i is chosen the way it is.

All of the complication that is to follow arises in order to deal with the limit case. When a node is given only one $\Pi_{\alpha+2}$ assertion, each of its children is simply given a single Π_α assertion. If α is finite, the resulting tree has finite height and just one assertion per node. On a first reading it may be helpful to have this special case in mind.

Here is another example, this one for the simplest limit case. If a node is given a single Π_ω assertion, that assertion may be broken up into assertions of size $\Pi_2, \Pi_4, \Pi_6, \Pi_8$, and so on, such that the original assertion is true if and only if all the sub-assertions are true. In that case, most of the children of the node end up totally empty, but of the ones that do not, the first one evaluates only the Π_2 assertion, the second one evaluates the Π_2 and Π_4 assertions, and so on. If all the assertions are true, then the childrens' ranks get bigger the more assertions they evaluate, causing the rank of the whole tree to reach $\omega + 1$. But if the Π_{2n} assertion is false for some n, then every child that evaluates that one has finite rank at most n, and every child that does not evaluate that one has rank at most n as well (because it only evaluates small assertions). So the tree as a whole gets rank at most $n + 1$.

Theorem 2. *Let $\alpha_1, \dots, \alpha_k > 0$ be constructive ordinals, and let $x_1, \dots, x_k$ be any natural numbers. Recursively in $\alpha_1, \dots, \alpha_k, x_1, \dots, x_k$, one may compute a well-founded tree T such that*

- $|T|_{ls} = \max_i \alpha_i + 1$ *if $x_i \notin \emptyset_{(2\alpha_i)}$ for all i*

- $|T|_{ls} \leq \alpha_i$ *whenever $x_i \in \emptyset_{(2\alpha_i)}$.*

Proof. In order to perform the induction we will actually prove something slightly stronger. If $x_i \in \emptyset_{(2\alpha_i)}$ for α_i a limit, given as $\alpha_i = \sup_n \beta_n$, then by Lemma 1.2.3 there is a least z such that $g_l(x_i, \alpha_i, z) \in \emptyset_{(2\beta_z)}$. In this case, we will ensure that T also satisfies $|T|_{ls} \leq \beta_z + 1$ for that least z.

Define T recursively as follows. Renumber the inputs so that $\alpha_1 \geq \cdots \geq \alpha_k$. (Since all the ordinal notations are comparable, this step is computable). The empty sequence is in T. To compute information about the nth child of the root, decode n as $n = \langle m_0, m_1, \dots, m_k \rangle$ and do the following:

1. Check that $m_0 < m_1 < \cdots < m_k$. If it is not, $T_n = \emptyset$.

2. For any i such that α_i is a limit, check that $m_i = m_{i-1} + 1$. If it does not, then $T_n = \emptyset$.

3. For any i such that $\alpha_i = 1$, check that $g_0(x_i, m_{i-1}, m_i) = 1$. If it does not, then $T_n = \emptyset$.

4. If $\alpha_1 = 1$, $T_n = \{\emptyset\}$.

5. Otherwise, we decide the subtree rooted at $\langle m_0, \ldots, m_k \rangle$ according to membership in the tree which we will now specify. Build a finite set $\mathcal{F}$ of ordinal-input pairs as follows.

   - Let $\mathcal{F}_1 = \{(\alpha_i, x_i) : \alpha_i < \alpha_1\}$

   - Let $\mathcal{F}_2 = \{(\beta, g_s(x_i, m_{i-1}, m_i)) : \alpha_i = \beta + 1 \text{ where } \beta > 0\}$

   - For each limit $\alpha_i = \sup_n \beta_n$, let $M_i \geq m_i$ be least such that for each $(\gamma, x) \in \mathcal{F}_1 \cup \mathcal{F}_2$, if $\gamma < \alpha_i$, then $\gamma \leq \beta_{M_i}$. (Again, this M_i may be effectively computed since the notations involved are all comparable.) For each $n \leq M_i$, let $(\beta_n, g_l(x_i, \alpha_i, n)) \in \mathcal{F}_3$.

   Let $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3$. Then T_n is defined recursively as the tree computed from the pairs in $\mathcal{F}$.

This completes the construction.

Observe that the resulting T is well-founded because each time we recurse, the size of the largest ordinal under consideration decreases. Let us verify the properties of this T. We proceed by induction on the size of $\max_i \alpha_i$.

For now on, consider the α_i to be numbered in order, so $\max_i \alpha_i = \alpha_1$.

In the base case, $\alpha_1 = \cdots = \alpha_k = 1$. If $g_0(x_i, m_{i-1}, m_i) = 0$ for any i, then $T = \{\emptyset\}$ and $|T|_{ls} = 1$ which is correct. If $g_0(x_i, m_{i-1}, m_i) = 1$ for all i, step (4) is encountered infinitely often and thus $|T|_{ls} = 2$, which is correct.

Now we consider the case $\alpha_1 > 1$. If, when computing subtree T_n, the algorithm makes it to step 5, then we call n a *recursing child*.

By induction we may always assume that for each child of the root n, $|T_n|_{ls} \leq \alpha_1$. This follows because $|T_n|_{ls} \leq 1$ for non-recursing children n, and for recursing children n, the ordinals considered in order to decide subtree T_n are all less than α_1. Therefore it is always true that $|T|_{ls} \leq \alpha_1 + 1$.

Case 1: The rank should be large

Suppose that for all i, $x_i \notin \emptyset_{(2\alpha_i)}$. Let us see that in this case $|T|_{ls} = \alpha_1 + 1$ is attained. Recall that a child of the root n is decoded as $n = \langle m_0, \ldots, m_k \rangle$. For each choice of m_0, a certain child of the root is obtained by inductively choosing m_i as follows according to the nature of α_i. The functions g_0 and g_s are as defined in Lemma 1.2.2.

1. If $\alpha_i = 1$, choose m_i so that $g_0(x_i, m_{i-1}, m_i) = 1$,

2. If $\alpha_i = \beta + 1$ with $\beta > 0$, choose m_i so that $g_s(x_i, m_{i-1}, m_i) \notin \emptyset_{(2\beta)}$

3. If α_i is a limit, choose $m_i = m_{i-1} + 1$.

Let n_j be the child so constructed starting with $m_0 = j$. By the definitions of g_0 and g_s, each m_i described above exists, is unique, and satisfies $m_i > m_{i-1}$.

One can check that n_j is a recursing child, and so T_{n_j} is formed using a finite set of ordinal-index pairs (γ, z). Notice that the choices of m_i above, together with the fact that for all i, $x_i \notin \emptyset_{(2\alpha_i)}$, guarantee that $z \notin \emptyset_{(2\gamma)}$ for each of these pairs (γ, z). Therefore, $|T_{n_j}|_{ls}$ will be determined by the largest ordinal under consideration in the construction of T_{n_j}. Now if $a_1 = \beta + 1$, then one of the pairs under consideration in the construction of T_{n_j} is $(\beta, g_s(x_1, m_0, m_1))$, and β is maximal among ordinals considered for T_{n_j}. Therefore by the inductive hypothesis, for each j we have $|T_{n_j}|_{ls} = \beta + 1 = \alpha_1$. Since there are infinitely many child subtrees where this rank is obtained, $\limsup_n |T_n|_{ls} = \alpha_1$ and thus $|T|_{ls} = \alpha_1 + 1$ as required. On the other hand, if $\alpha_1 = \sup_n \beta_n$ is a limit, then $(\beta_{M_1}, g_l(x_1, \alpha_1, M_1))$ is used when assembling T_{n_j}, and β_{M_1} is maximal among ordinals considered, because if $\alpha_i < \alpha_1$, then $\beta_{M_1} \geq \alpha_i$, and if $\alpha_i = \alpha_1$, then $M_i = M_1$ (since their selection algorithms are identical). Therefore, by the inductive hypothesis,

$$|T_{n_j}|_{ls} = \beta_{M_1} + 1 > \beta_j + 1$$

because $M_1 \geq m_1 > m_0 = j$. Since $\lim_j \beta_j = \alpha_1$, we have

$$\lim_j |T_{n_j}|_{ls} \geq \lim_j (\beta_j + 1) = \alpha_1$$

as well. Therefore, $\limsup_n |T_n|_{ls} = \alpha_1$ and $|T|_{ls} = \alpha_1 + 1$ as required. Therefore, if for all i, $x_i \notin \emptyset_{(2\alpha_i)}$, then $|T|_{ls} = \alpha_1 + 1$.

Case 2: The rank should be small

On the other hand, suppose that $x_i \in \emptyset_{(2\alpha_i)}$ for some i. Fix an index r at which this occurs. We will show that $|T|_{ls} \leq \alpha_r$.

Subcase 2.1 Suppose $\alpha_r = \beta_r + 1$. By Lemma 1.2.2 let z_r be such that

$$(\forall z > z_r)(\forall y > z)[g_s(x_r, z, y) \in \emptyset_{(2\beta_r)}]$$

if $\beta_r > 0$, or such that $(\forall z > z_r)(\forall y > z)[g_0(x_r, z, y) = 0]$ if $\beta_r = 0$. One may check that for any child $n = \langle m_0, \ldots, m_k \rangle$ such that $m_{r-1} > z_r$, if n is recursing, then included in consideration for T_n is $(\beta_r, g_s(x_r, m_{r-1}, m_r))$ where $g_s(x_r, m_{r-1}, m_r) \in \emptyset_{(2\beta_r)}$; and if n is not recursing, $T_n = \emptyset$. Therefore by induction, $|T_n|_{ls} \leq \beta_r < \alpha_r$ for such n.

Now let us consider recursing children n such that $m_{r-1} \leq z_r$. There are only finitely many ways $m_0 < \cdots < m_{r-1} \leq z_r$ to begin such children. Fix one such beginning. We claim that for all but at most one choice of the remaining $m_r < \cdots < m_k$, $|T_n|_{ls} < \alpha_r$. That one choice, if it exists, is constructed inductively as in the previous case. That is, for each $i \geq r$, choose m_i to satisfy

1. If $\alpha_i = 1$, satisfy $g_0(x_i, m_{i-1}, m_i) = 1$,
2. If $\alpha_i = \beta + 1$ with $\beta > 0$, satisfy $g_s(x_i, m_{i-1}, m_i) \notin \emptyset_{(2\beta)}$, and
3. If α_i is a limit, let $m_i = m_{i-1} + 1$.

If these m_i exist, they are unique. Suppose we deviate from this recipe in the case of α_i a limit. Then T_n is empty. Suppose we deviate from this one way in the case of $\alpha_i = 1$, and let $g_0(x_i, m_{i-1}, m_i) = 0$. Then by step (3), T_n is empty. Suppose we deviate from this one way in the case of $\alpha_i = \beta + 1$, and include $(\beta, g_s(x_i, m_{i-1}, m_i))$ in the assembling of T_n, where $g_s(x_i, m_{i-1}, m_i) \in \emptyset_{(2\beta)}$. Then by the inductive hypothesis we are guaranteed $|T_n|_{ls} \leq \beta < \alpha_i \leq \alpha_r$. Therefore, considering all children n, there are at most finitely many such that $|T_n|_{ls} \geq \alpha_r$. Therefore, $\limsup_n |T_n|_{ls} \leq \beta_r$.

It remains to show that for each recursing child n, $|T_n|_{ls} \leq \alpha_r$. There are two possibilities. If $\alpha_1 > \alpha_r$, then (α_r, x_r) is included in consideration for T_n, and thus by the inductive hypothesis $|T_n|_{ls} \leq \alpha_r$. On the other hand, if $\alpha_1 = \alpha_r = \beta_r + 1$, then β_r is maximal among ordinals considered for T_n, so by the inductive hypothesis $|T_n|_{ls} \leq \beta_r + 1 = \alpha_r$. Therefore, if α_r is a successor, then $|T|_{ls} \leq \alpha_r$.

Subcase 2.2: Suppose $\alpha_r = \sup_n \beta_n$ is a limit. Using Lemma 1.2.3, let z_r be least such that $g_{ls}(x_r, \alpha_r, z_r) \in \emptyset_{(2\beta_{z_r})}$. Let us consider children $n = \langle m_0, \ldots, m_k \rangle$ such that $m_r \geq z_r$. For each of these n, the pair $(\beta_{z_r}, g_l(x_r, \alpha_r, z_r))$ is used in assembling T_n. So for each such n, $|T_n|_{ls} \leq \beta_{z_r}$.

On the other hand, there are the n such that $m_r < z_r$. There are only finitely many ways $m_0 < \cdots < m_r < z_r$ to begin such an n. We claim that for each such beginning, there is at most one sequence $m_{r+1}, \ldots, m_k$ which completes n in such a way that $|T_n|_{ls} > \beta_{z_r}$. The strategy is exactly the same as in the successor case. See (1)-(3) above.

In each case, if such an m_i exists, it is unique. If we deviate from this plan in the case of α_i a limit or $\alpha_i = 1$, then one may check that T_n is empty. If we deviate in the case of $\alpha_i = \beta+1$ with $\beta > 0$, then we include $(\beta, g_s(x_i, m_{i-1}, m_i)) \in \mathcal{F}_2$, where $g_s(x_i, m_{i-1}, m_i) \in \emptyset_{(2\beta)}$. So to start with, $|T_n|_{ls} \leq \beta$, and if $\beta \leq \beta_{z_r}$ then $|T_n|_{ls}$ is small enough. But if $\beta > \beta_{z_r}$, then this bound is insufficient. In that case, recall that during the creation of $\mathcal{F}_3$ which was used to assemble T_n, we defined M_r to satisfy $M_r \geq m_r$ and $\beta_{M_r} \geq \gamma$ for each $(\gamma, z) \in \mathcal{F}_2$ such that $\gamma < \alpha_r$. Because $\alpha_i < \alpha_r$, $\beta < \alpha_r$. So $\beta_{M_r} \geq \beta > \beta_{z_r}$, so $M_r > z_r$. So in particular, $(\beta_{z_r}, g_l(x_r, \alpha_r, z_r))$ was included when assembling T_n. Therefore, $|T_n|_{ls} \leq \beta_{z_r}$. Therefore, for all but finitely many n, $|T_n|_{ls} \leq \beta_{z_r}$.

It remains to show that for each individual n, $|T_n|_{ls} \leq \beta_{z_r} + 1$.

We now consider two cases. Suppose $\alpha_1 > \alpha_r$. Then for each n, the pair (α_r, x_r) is again under consideration for T_n. But the new leading ordinal is smaller, so by induction, $|T_n|_{ls} \leq \beta_{z_r} + 1$ for each n. On the other hand, if $\alpha_1 = \alpha_r$, then $\alpha_1 = \alpha_2 = \cdots = \alpha_r$, so $M_1 = M_2 = \cdots = M_r$, since the algorithm which selects M_i is the same for each $i = 1, \ldots, r$. One may check that β_{M_r} is the largest ordinal under consideration in the assembling of T_n. If $\beta_{M_r} \geq \beta_{z_r}$, then $(\beta_{z_r}, g_l(x_r, \alpha_r, z_r))$ is included, and $|T_n|_{ls} \leq \beta_{z_r}$. On the other hand, if

$\beta_{M_r} < \beta_{z_r}$, then since β_{M_r} is largest,

$$|T_n|_{ls} \leq \beta_{M_r} + 1 \leq \beta_{z_r}.$$

Therefore, $\sup_n |T_n|_{ls} \leq \beta_{z_r} + 1$. Therefore, if α_r is a limit with $g_l(x_r, \alpha_r, z_r) \in \emptyset_{(2\beta_{z_r})}$, then $|T|_{ls} \leq \beta_{z_r} + 1 < \alpha_r$. This completes the proof. □

Application to the Cantor-Bendixson rank

As a first application of the result in the previous section, we completely analyze the descriptive complexity of the Cantor-Bendixson hierarchy in Cantor space.

Let $T \subseteq 2^{<\omega}$ be a tree with no dead ends. The Cantor-Bendixson derivative $D(T)$ is defined as the tree without dead ends whose paths are exactly those not isolated in T. Formally,

$$D(T) = \{\sigma \in T : \forall k \exists \tau_1 \dots \tau_k \in T, \sigma \prec \tau_i \text{ and } \tau_i | \tau_j \text{ when } i \neq j\}$$

where $\tau_i|\tau_j$ means that these nodes are incompatible. Define $D^0(T) = T$, $D^{\alpha+1}(T) = D(D^\alpha(T))$, and $D^\lambda(T) = \cap_{\alpha<\lambda} D^\alpha(T)$ for λ a limit.

Definition 1.2.2. *The Cantor-Bendixson rank of a tree T, denoted $|T|_{CB}$, is the least α such that $D^\alpha(T) = \emptyset$, if such exists. Otherwise we say $|T|_{CB} = \infty$.*

For some e, the tree encoded by ϕ_e is not well-defined or has dead ends. We note that $\{e : \phi_e \text{ codes a tree in } 2^{<\omega} \text{ with no dead ends}\}$ is Π_2, because it can be written as "ϕ_e is total, and codes a tree, and $(\forall \sigma \in T)(\exists \tau \in T)[\sigma \prec \tau]$." Let T_e denote the tree coded by ϕ_e whenever this is defined. We will drop the subscript when it is clear from context. If $\sigma \in T$, then T_σ denotes $\{\tau : \sigma^\frown\tau \in T\}$. Note that if C is any finite prefix-free collection of nodes σ such that $\cup_{\sigma \in C}[\sigma]$ covers $[T]$, then

$$D^\alpha(T) = \bigcup_{\sigma \in C} \sigma^\frown D^\alpha(T_\sigma)$$

where $\sigma^\frown T$ denotes $\{\sigma^\frown\tau : \tau \in T\}$, $[\sigma]$ denotes $\{X \in 2^\omega : \sigma \prec X\}$, and $[T]$ denotes $\{X \in 2^\omega : \forall n X \restriction n \in T\}$. In other words, since the path space is totally disconnected, the Cantor-Bendixson derivative may be performed on a finite open partition of the space without error.

Corollary 1.2.4. *(Lempp, 1987) For each constructive $\alpha > 1$,the sets*

$$\{e : T_e \text{ has no dead ends and } |T|_{CB} \leq \alpha\}$$

are $\Sigma_{2\alpha}$-complete.

Proof. First we show that when e is restricted to the Π_2 set of codes for no-dead-end trees, "$|T_e|_{CB} \leq \alpha$" is $\Sigma_{2\alpha}$ for all constructive α. The proof is by effective transfinite recursion. Because checking whether a tree is empty can be accompanied by checking the root, the statement "$|T_e|_{CB} = 0$" is Δ_0 when the domain of e is restricted to the Π_2 set of codes for no-dead-end trees.

A tree has $D^{\alpha+1}(T) = \emptyset$ if and only if $D^{\alpha}(T)$ has only finitely many branches. If $D^{\alpha}(T)$ has at least k branches, then by going up to a height n at which the branches have separated, we may find at least k-many σ of length n such that $D^{\alpha}(T_\sigma) \neq \emptyset$. And if there are k incomparable σ such that $D^{\alpha}(T_\sigma) \neq \emptyset$, then $D^{\alpha}(T)$ has at least k branches. Thus an equivalent condition to "$D^{\alpha+1}(T) = \emptyset''$ is: "There is a k such that for all n, there are at least $(2^n - k)$-many σ of length n for which $D^{\alpha}(T_\sigma) = \emptyset$." Assuming $D^{\alpha}(T) = \emptyset$ is $\Sigma_{2\alpha}$ uniformly in α and T, this shows that $D^{\alpha+1}(T) = \emptyset$ is $\Sigma_{2\alpha+1}$.

If λ is a limit, a tree has $D^{\lambda}(T) = \emptyset$ if and only if there is an $\alpha < \lambda$ such that $D^{\alpha}(T) = \emptyset$, by compactness. Assuing $D^{\alpha}(T) = \emptyset$ is uniformly $\Sigma_{2\alpha}$ and the set of $\alpha < \lambda$ is c.e., we have $D^{\lambda}(T) = \emptyset$ if and only if $\exists \alpha < \lambda[D^{\alpha}(T) = \emptyset]$, a Σ_λ statement. Note $\Sigma_\lambda = \Sigma_{2\lambda}$ for λ a limit.

Therefore, the statement "T is a no-dead-ends tree and $D^{\alpha}(T) = \emptyset$" is $\Sigma_{2\alpha}$ for $\alpha > 1$, uniformly in T and α.

Now we use the main theorem to provide a new proof that having Cantor-Bendixson rank at most α is $\Sigma_{2\alpha}$-complete.

Theorem 2 gives a reduction from the canonical $\Sigma_{2\alpha}$ complete set to trees in Baire space of limsup rank at most α. Here we show that the following familiar reduction f takes trees in Baire space of limsup rank α to trees in Cantor space of Cantor-Bendixson rank α.

The intuitive idea is that each node of a tree $T \subseteq \mathbb{N}^{<\mathbb{N}}$ corresponds to a path in $f(T) \subseteq 2^{<\omega}$ with the topological clustering of the paths provided by the hierarchical structure of T.

Define f by

$$f(T) = \{0^{n_0}10^{n_1}1 \cdots 0^{n_k}10^m : (n_0, \ldots, n_k) \in T\}.$$

The intention is that for all m, $0^m \in f(T)$ if and only if the empty node is in T.

We claim that $|T|_{ls} = |f(T)|_{CB}$. The proof is by induction on the usual rank of T.

If $T = \emptyset$ then also $f(T) = \emptyset$, so $|f(T)|_{CB} = |T|_{ls} = 0$.

Suppose $|T|_{ls} = \alpha + 1$. (This is the only case because the limsup rank is always a successor.) Then there is an N such that for $n > N$, $|T_n|_{ls} \leq \alpha$. Because $|S|_{CB} = |\sigma^\frown S|_{CB}$ for all $\sigma \in 2^{<\omega}$ and $S \subseteq 2^{<\omega}$, we have by induction that for $n \leq N$, $|0^n1^\frown f(T_n)|_{CB} \leq \alpha + 1$, and for $n > N$, we have $|0^n1^\frown f(T_n)|_{CB} \leq \alpha$. We have

$$\begin{aligned} D^{\alpha}(T) &= \bigcup_{n \leq N} 0^n1^\frown D^{\alpha}(f(T)_{0^n1}) \bigcup 0^{N+1} D^{\alpha}(f(T)_{0^{N+1}}) \\ &= \bigcup_{n \leq N} 0^n1^\frown D^{\alpha}(f(T_n)) \bigcup 0^{N+1} D^{\alpha}(f(T)_{0^{N+1}}). \end{aligned}$$

Because each member of the finite union on the left is a tree with finitely many branches, the left side contributes finitely many branches. On the right, $D^{\alpha}(f(T)_{0^{N+1}})$ has either one branch (0^ω) or is empty. It has no branches starting $0^n1\dots$ because $D^\alpha(f(T_n)) = \emptyset$ for all $n > N$. Since $D^\alpha(f(T))$ has finitely many paths, $|f(T)|_{CB} \leq \alpha + 1$.

Now we need $|f(T)|_{CB} \geq \alpha + 1$. If $\limsup_n |T_n|_{ls} = \alpha$, then for every $\beta < \alpha$, there are infinitely many n such that $|T_n| > \beta$. By induction we have that for all $\beta < \alpha$ there infinitely many n for which $|0^n 1^\frown f(T_n)|_{CB} > \beta$. Therefore, for all $\beta < \alpha$, 0^ω is not isolated in $D^\beta(f(T))$, so $0^\omega \in D^\alpha(f(T))$, and $|f(T)|_{CB} > \alpha$.

In the other case, $|T_n|_{ls} = \alpha + 1$ for some n. Then by induction $|0^n 1^\frown f(T_n)|_{CB} = \alpha + 1$, and the result follows. □

1.3 Applications to the Kechris-Woodin rank

The set of differentiable $C[0,1]$ functions is not Borel, but it can be represented hierarchically as an increasing union of Borel sets. The Kechris-Woodin rank, denoted $|\cdot|_{KW}$, on differentiable functions is defined in [14] using an *ordinal rank*, a mapping from differentiable functions to countable ordinals, whose range is unbounded below ω_1. It decomposes the set $\mathcal{D}$ of differentiable $C[0,1]$ functions as

$$\mathcal{D} = \bigcup_{\alpha<\omega_1} \{f : |f|_{KW} < \alpha\}$$

where each constituent of the union is Borel.

Our contribution is a finer-grained, recursion-theoretic analysis of this hierarchy. The lightface situation mirrors the boldface situation in many ways. We begin with the observation (a corollary of the work in [14]) that the set D of integer codes for computable differentiable $C[0,1]$ functions is a Π^1_1-complete set, and it decomposes as

$$D = \bigcup_{\alpha<\omega_1^{CK}} \{c : c \text{ codes } f \text{ with } |f|_{KW} < \alpha\}$$

where each constituent of the union is hyperarithmetic. Our results pinpoint the exact location of each constituent set in the hyperarithmetical hierarchy.

Theorem 3. *For each nonzero $\alpha < \omega_1^{CK}$, the set*

$$\{c : c \textit{ codes } f \textit{ with } |f|_{KW} < \alpha + 1\}$$

is $\Pi_{2\alpha+1}$-complete.

Here and throughout we use the notational convention of Ash and Knight [1] for a Σ_α set, discussed in Section 1.1. We also analyze the limit case:

Theorem 4. *For each limit $\lambda < \omega_1^{CK}$, the set $\{c : c \textit{ codes } f \textit{ with } |f|_{KW} < \lambda\}$ is Σ_λ-complete.*

The study of differentiation through the lens of computable analysis has typically involved restricting attention to the continuously differentiable functions. The definition of a computable function proposed by Grzegorczyk and Lacombe, and further developed by Pour-El and Richards and others (see [9], [16], [27]), has no notion of computability for a discontinuous function. Therefore, restricting differentiation to the continuously differentiable functions is a strategy for making questions such as "Is differentiation computable?" meaningful. The fact that $f \mapsto f'$ is not computable was first demonstrated by Myhill [26], who constructed a computable function whose continuous derivative is not computable.

At the other end of the spectrum, computable functions that are not everywhere differentiable have been studied. Brattka, Miller and Nies (to appear) have used randomness notions to characterize the points at which all computable almost everywhere differentiable functions must be differentiable. However, as far as the author is aware, the everywhere differentiable functions with discontinuous derivatives have not yet been studied in the setting of computable analysis.

Previously, Cenzer and Remmel [3] showed that $\{e : f_e \text{ is continuously differentiable}\}$ is Π^0_3-complete, which is the same as the $\alpha = 1$ case of our Theorem 7. They also showed that $\{e : f_e \text{ is continuously differentiable with } f'_e \text{ computable}\}$ is Σ^0_3-complete. Again, only continuously differentiable functions were considered. By contrast, our aim is to provide a clearer picture of the structure of the unrestricted set of everywhere differentiable functions.

In Section 1.3 we define Kechris and Woodin's differentiability rank. In Section 1.3 we familiarize the reader with the building blocks used in [14] to construct functions of arbitrary rank, as these essential elements are taken for granted in what follows. In Section 1.3 we establish more notation that is used throughout the paper. Finally, in Section 1.3 we present some necessary facts about computable differentiable functions that can be obtained by effectivizing existing work. In Section 1.3 we redefine the differentiability rank in a more computationally convenient way, and use this definition to demonstrate $\{c : c \text{ codes } f \text{ with } |f|_{KW} < \alpha + 1\}$ is $\Pi_{2\alpha+1}$. Finally, in section 1.3, we we address the question of completeness to prove both theorems above.

Preliminaries

Kechris and Woodin's differentiability rank

Kechris and Woodin [14] define a rank on differentiable $C[0,1]$ functions as follows. Let $\Delta_f(x,y)$ denote the secant slope

$$\Delta_f(x,y) = \frac{f(x) - f(y)}{x - y}.$$

They define a "derivative" operation, which is given below. This operation starts with a closed set of points P and removes from it some points at which f seems to be differentiable. A point x is removed if the oscillation of f' near x is no more than the given ε.

Definition 1.3.1. *Given a closed set P, a function $f \in C[0,1]$ and $\varepsilon > 0$,*

$$P'_{f,\varepsilon} = \{x \in P : \forall \delta > 0 \exists p < q, r < s \in B(x,\delta) \cap [0,1] \text{ with } [p,q] \cap [r,s] \cap P \neq \emptyset \text{ and } |\Delta_f(p,q) - \Delta_f(r,s)| \geq \varepsilon\}$$

where all the quantifiers range over rational numbers.

If P is closed, then P' is closed as well, so for each $f \in C[0,1]$ and each $\varepsilon > 0$ one defines the following inductive hierarchy:

$$P^0_{f,\varepsilon} = [0,1]$$
$$P^{\alpha+1}_{f,\varepsilon} = (P^\alpha_{f,\varepsilon})'_{f,\varepsilon}$$
$$P^\lambda_{f,\varepsilon} = \cap_{\alpha<\lambda} P^\alpha_{f,\varepsilon} \text{ for a limit } \lambda$$

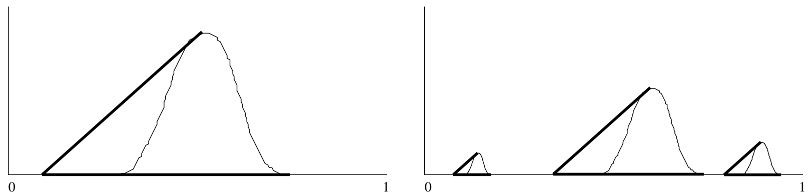


Figure 1.1: (a) A continuously differentiable bump with one secant of slope zero and one secant of positive slope. (b) Resized copies of this bump with proportions preserved.

Kechris and Woodin showed that for any $f \in C[0,1]$, f is differentiable if and only if $\forall n \exists \alpha < \omega_1 (P^{\alpha}_{f,1/n} = \emptyset)$. Considering the supremum of all such α, they make the following definition:

Definition 1.3.2. *For each differentiable $f \in C[0,1]$, define $|f|_{KW}$ to be the least ordinal α such that $\forall \varepsilon P^{\alpha}_{f,\varepsilon} = \emptyset$.*

For example, if f is any continuously differentiable function, then $|f|_{KW} = 1$, the least possible. To see that $P^1_{f,\varepsilon} = \emptyset$ for any such f and any ε, let δ be s.t. $|f'(z) - f'(y)| < \varepsilon$ whenever $|z - y| < \delta$. Then for any x and any $p < q, r < s \in B(x, \delta/2)$, the Mean Value Theorem provides $y \in [p,q]$ and $z \in [r,s]$ such that $f'(y) = \Delta_f(p,q)$ and $f'(z) = \Delta_f(r,s)$, so $|\Delta_f(p,q) - \Delta_f(r,s)| < \varepsilon$ and $x \notin P^1_{f,\varepsilon}$. A common example of a differentiable function whose derivative is not continuous is $x^2 \sin(1/x)$, and this function has differentiability rank 2.

Basic building blocks

Kechris and Woodin show that for each ordinal α, there is a function with rank α, and in order to show this they construct an explicit f with that rank. This section gives a summary of the building blocks that they used to produce an example of a function living at each level of their hierarchy. We will use the same building blocks in a more complicated construction in Section 1.3.

The most natural way of constructing a function while controlling its rank is to build it up recursively from smaller pieces. Our basic building block is a simple continuously differentiable bump (Figure 1.1).

Observe a certain pair of secants made by the existence of the bump, one with slope zero and one with positive slope. We build functions out of resized copies of this same bump, always preserving the proportions to keep the corresponding slopes uniform. In Definition 1.3.1 there is a free parameter ε, and one compares various secants to see if their slope difference is at least ε. Therefore, by choosing a single sufficiently small value for ε, all the secant pairs induced by the bumps are made visible for the purposes of the rank-ascertaining process. We will sometimes refer to ε as the *oscillation sensitivity* because it sets the threshold above which oscillations in the value of the derivative matter.

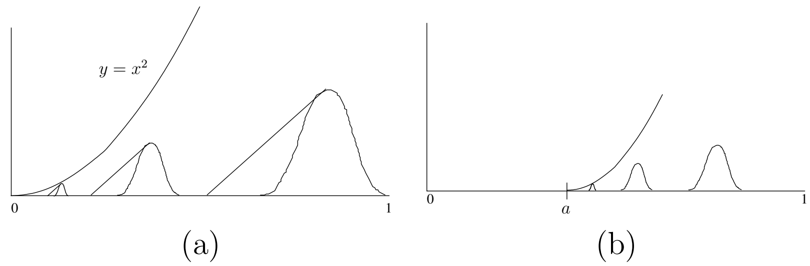


Figure 1.2: (a) A simple differentiable function of rank 2. (b) A shifted and resized copy of this function, which fits in a small neighborhood of the point a and keeps a alive through the first iteration.

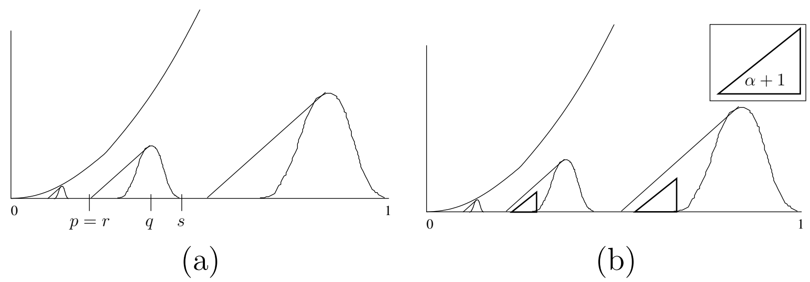


Figure 1.3: (a) Points p, q, r, s as used in Definition 1.3.1. (b) A differentiable function of rank $\alpha + 2$. The triangle represents a function of rank $\alpha + 1$.

A simple rank 2 function is pictured in Figure 1.2. To keep 0 from being removed at the first iteration, we put a bump (and thus a disagreeing pair of secants) in every neighborhood of 0. To ensure the function remains differentiable at 0 despite all the oscillation, we make the bumps small enough to fit inside an envelope of x^2. The resulting rank 2 function can itself be proportionally shrunk and used as a building block in functions of larger rank.

The reason 0 is removed at the second iteration, despite infinitely many pairs of disagreeing secants, is that P^1 contains no points which lie in the intersection $[p, q] \cap [r, s]$, where p, q, r, s are the endpoints of the intervals defining the disagreeing secant pair as shown in Figure 1.3. But if we have a rank $\alpha + 1$ function to use as a building block (the rank must be a successor for reasons discussed below), we can make 0 survive the $(\alpha + 1)$st iteration. By putting a shrunken copy of our rank $\alpha + 1$ function in $[p, q] \cap [r, s]$ as shown in Figure 1.3, we construct a function of rank $\alpha + 2$. We say that we have put the rank $\alpha + 1$ function in the *shadow* of each bump. In fact, it would suffice to put a rank $\alpha + 1$ function in the shadow of infinitely many of the bumps, and this is done later in the paper.

Next we describe how to make functions of rank $\lambda + 1$ and rank λ, where λ is a limit ordinal. We say that an oscillation sensitivity ε *witnesses* the rank of a function f if $|f|_{KW} = \alpha$ and $P^{\beta}_{\varepsilon} \neq \emptyset$ for all $\beta < \alpha$. Note that if a function has successor rank, there is always an ε

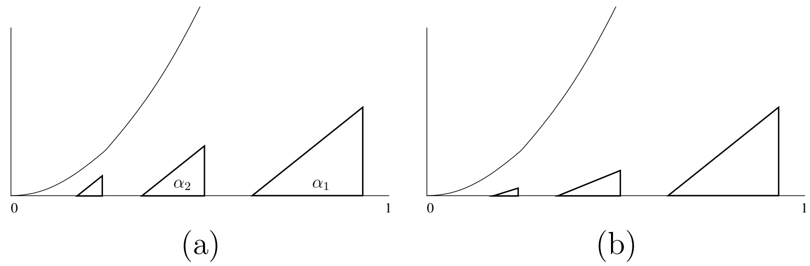


Figure 1.4: (a) A function of rank $\lambda + 1$ for λ a limit ordinal. (b) A function of rank λ.

that witnesses this, but if the function has limit rank, there cannot be a witness.

Suppose we have a sequence of functions, with ranks cofinal in λ, whose ranks are all witnessed at a uniform sensitivity ε. As shown in Figure 1.4, a function of rank $\lambda + 1$ can be made by putting proportionally shrunken copies of functions of increasing rank in each neighborhood of 0. The rank of the resulting function is witnessed by the same ε.

By recursively applying the $\alpha + 2$ step and the $\lambda + 1$ step, we can build functions of any successor rank. To make a function of rank λ, we must start with a sequence of functions with uniformly bounded derivatives, whose ranks are cofinal in λ. Because the derivatives are uniformly bounded, their possible secant slope differences are also uniformly bounded by the Mean Value Theorem. Again we use shrunken copies of functions from the sequence, but in addition to shrinking the nth function proportionally, we also scale it vertically by a factor of $\frac{1}{n}$. In the resulting function, as x approaches 0 the nearby secant slope differences approach zero, which has the effect of ensuring that 0 is removed at the first iteration no matter what the oscillation sensitivity.

Functions whose ranks are limit ordinals do not make good building blocks for more complicated functions because there is no ε that witnesses their rank. If we construct a rank $\lambda + 1$ function f, there needs to be a ε such that $P^{\lambda}_{\varepsilon,f} \neq \emptyset$. If we used a rank λ function g as a building block, then by compactness there would have to be some $\beta < \lambda$ such that $P^{\beta}_{\varepsilon,g} = \emptyset$. So a function of rank β would have been equally unhelpful. That explains why, in our construction of the rank $\alpha+2$ function above, we needed to use a function with successor rank $\alpha + 1$ as a building block.

Notation

The following notations are used throughout.

Definition 1.3.3. *For each ordinal α, let D_α denote the set of all indices e such that $f_e \in C[0,1]$ is differentiable with $|f_e|_{KW} < \alpha$. Define $D = \cup_{\alpha<\omega_1} D_\alpha$.*

For any function $f \in C[0,1]$, we write $f[a,b]$ to denote the function which is identically 0 outside of $[a,b]$, and for $x \in [a,b]$, $f[a,b](x) = (b-a)f(\frac{x-a}{b-a})$. Note that if f is continuous and $f(0) = f(1) = 0$, then $f[a,b]$ is continuous; it is computable when f, a, and b are and differentiable when f is differentiable and $f'(0) = f'(1) = 0$.

Similarly, for any real number $c \in [0,1]$ and any interval $[a,b]$, let $c[a,b] = a + c(b-a)$. This notation comes in handy when talking about scaled down versions of functions, because $(b-a)f(c) = f[a,b](c[a,b])$. Also, this scaling preserves a function's proportions ($f[a,b]'(c[a,b]) = (b-a)f'(c)\frac{1}{b-a} = f'(c)$), so $||f'|| = ||f[a,b]'||$ for any interval $[a,b]$.

Facts about D

In section 1.3, we described the major components of Kechris and Woodin's construction of an explicit f with $|f|_{KW} = \alpha$ for each α. When $\alpha < \omega_1^{CK}$, their construction by transfinite recursion easily effectivizes. Therefore their argument also shows that for each constructive α, there is a computable differentiable f with rank α.

On the other hand, every computable differentiable function has constructive rank. This follows from work in the same paper by Kechris and Woodin.

Definition 1.3.4. *Let $\mathcal{D}$ denote the set of differentiable functions in $C[0,1]$.*

Definition 1.3.5. *For each function $f \in C[0,1]$ and each $\varepsilon \in \mathbb{Q}^+$, define a tree S_f^ε on $A = \{\langle p,q\rangle \ : 0 \le p < q \le 1$ and $p, q \in \mathbb{Q}\}$ as follows:*

$$(\langle p_1,q_1\rangle,\ldots,\langle p_n,q_n\rangle) \in S_f^\varepsilon \iff \forall i \le n(q_i - p_i \le 1/i) \text{ and } \cap_{i=1}^n [p_i,q_i] \neq \emptyset \\ \text{and } \forall i < n(|\Delta_f(p_{i+1},q_{i+1}) - \Delta_f(p_i,q_i)| \ge \varepsilon).$$

Kechris and Woodin showed that for all $f \in C[0,1]$, $f \in \mathcal{D}$ if and only if $\forall \varepsilon \in \mathbb{Q}^+(S_f^\varepsilon$ is well-founded). That makes possible the following alternative rank definition:

Definition 1.3.6. *Let $f \in \mathcal{D}$. Define $|f|^* = \sup\{|S_f^\varepsilon| + 1 : \varepsilon \in \mathbb{Q}^+\}$.*

Lemma 1.3.1. *If $f \in \mathcal{D}$ is computable, then $|f|^*$ is constructive.*

Proof. Note that the tree S_f^ε would be computable if one did not have to verify that $|\Delta_f(p_{i+1},q_{i+1}) - \Delta_f(p_i,q_i)| \ge \varepsilon$, a Π_1 statement. In fact such a strong statement is not needed, and to get around it we use a computable approximation. For any computable $g \in C[0,1]$, rational $p \in [0,1]$, and rational $\delta > 0$, the notation $[g(p)]_\delta$ refers to a standard δ-approximation of $g(p)$, which is a rational number z such that $|g(p) - z| < \delta$. (For specificity we could say $[g(p)]_\delta$ is the b component of the smallest $\langle n,p,r,b,\delta/2\rangle$ in the computable code for g.) Given a computable f, consider the following collection of trees $\widetilde{S}_f^\varepsilon$, which are the same as the S_f^ε defined above, except for the use of a computable approximation:

$$(\langle p_1,q_1\rangle,\ldots,\langle p_n,q_n\rangle) \in \widetilde{S}_f^\varepsilon \iff \forall i \le n(q_i - p_i \le 1/i) \text{ and } \cap_{i=1}^n [p_i,q_i] \neq \emptyset \\ \text{and } \forall i < n([|\Delta_f(p_{i+1},q_{i+1}) - \Delta_f(p_i,q_i)|]_{\varepsilon/4} \ge \varepsilon).$$

The $\widetilde{S}^{\varepsilon}_f$ are computable trees, uniformly in f and ε. Furthermore, for each ε, $S^{2\varepsilon}_f \subseteq \widetilde{S}^{\varepsilon}_f \subseteq S^{\varepsilon/2}_f$, so $|S^{2\varepsilon}_f| \leq |\widetilde{S}^{\varepsilon}_f| \leq |S^{\varepsilon/2}_f|$. Therefore, although $|f|^*$ is defined in terms of S^{ε}_f, it is also true that $|f|^* = \sup\{|\widetilde{S}^{\varepsilon}_f| + 1 : \varepsilon \in \mathbb{Q}^+\}$. Since $\widetilde{S}^{\varepsilon}_f$ are defined uniformly in ε, the tree

$$\widetilde{S} = \{\langle\varepsilon\rangle^\frown\sigma : \varepsilon \in \mathbb{Q}^+, \sigma \in \widetilde{S}^{\varepsilon}_f\}$$

is also computable, and $|f|^* = |\widetilde{S}|$. Therefore $|f|^*$ is constructive. □

Theorem 5 ([14]). *Let $f \in \mathcal{D}$. Then if f is linear, $|f|_{KW} = 1$, and if f is not linear, $|f|^* = \omega|f|_{KW}$.*

Therefore, for each computable f, $|f|_{KW} \in \mathcal{O}$. Thus

$$D = \bigcup_{\alpha<\omega_1^{CK}} D_\alpha.$$

By the standard definition of differentiability, D is a Π^1_1 set. Mazurkiewicz [22] gave a reduction from well-founded trees to differentiable functions. This reduction, reproduced in [14], easily effectivizes, and therefore also serves as a reduction from $\mathcal{O}$ to D. Thus we know that D is Π^1_1-complete. We will generate functions from well-founded trees using a method similar to that of Mazurkiewicz. By constructing the trees carefully we can obtain finer grained results.

Having differentiability rank at most α is $\Pi_{2\alpha+1}$

In this section, we show that "$|f|_{KW} < \alpha + 1$" is a $\Pi_{2\alpha+1}$ statement. This follows from a mostly straightforward translation of the definition of differentiability rank into the formal language. The only obstacle is that the original definition needs to be slightly optimized. In Section 1.3 we give an equivalent definition of differentiability rank which uses fewer quantifiers. In Section 1.3 we formalize the sentence "$|f|_{KW} \leq \alpha + 1$".

An equivalent rank function

In [14] the rank is defined using a "derivative operation" $P'_{f,\varepsilon}$ on sets P. To prove our result we use an almost identical operation $P^*_{f,\varepsilon}$ defined below. The only difference between this definition and the definition of $P'_{f,\varepsilon}$ is that $\geq$ is replaced with $>$. This is done in order to make the statement $[0,1]^*_{f,\varepsilon} = \emptyset$ a Σ_2 statement (instead of Σ_3), and this is necessary for the base case of Proposition 1.3.4.

Definition 1.3.7. *Given a closed set P, a function f and $\varepsilon > 0$,*

$$P^*_{f,\varepsilon} = \{x \in P : \forall\delta > 0 \exists p < q, r < s \in B(x,\delta) \cap [0,1] \text{ with } [p,q] \cap [r,s] \cap P \neq \emptyset \text{ and } |\Delta_f(p,q) - \Delta_f(r,s)| > \varepsilon\}$$

where all the quantifiers range over rational numbers.

It is easy to see that $P^*_{f,\varepsilon}$ is a closed subset of P, so it makes sense to define a rank function using it. We define a hierarchy of closed sets analogously to [14]:

Definition 1.3.8. ($\tilde{P}^{\alpha}_{f,\varepsilon}(I)$ hierarchy) *Fix a continuous function f, a rational $\varepsilon > 0$, and a closed set $I \subseteq [0,1]$. Define $\tilde{P}^{0}_{f,\varepsilon}(I) = I$. Then for each ordinal α, define $\tilde{P}^{\alpha+1}_{f,\varepsilon}(I) = (\tilde{P}^{\alpha}_{f,\varepsilon}(I))^*_{f,\varepsilon}$. If λ is a limit ordinal, define $\tilde{P}^{\lambda}_{f,\varepsilon}(I) = \cap_{\alpha<\lambda}\tilde{P}^{\alpha}_{f,\varepsilon}(I)$.*

In the special case $I = [0,1]$, we write $\tilde{P}^{\alpha}_{f,\varepsilon}$ instead of $\tilde{P}^{\alpha}_{f,\varepsilon}([0,1])$. Sometimes the function f may also be omitted from the notation if it is clear from context.

The rank of a differentiable function f is defined in [14] to be the smallest ordinal α such that for all ε, $P^{\alpha}_{\varepsilon} = \emptyset$. The next lemma shows our $\tilde{P}^{\alpha}_{\varepsilon}$ hierarchy is similar enough to preserve the notion.

Lemma 1.3.2. *For any differentiable $f \in C[0,1]$, $\varepsilon > 0$ and ordinal α,*

$$\tilde{P}^{\alpha}_{\varepsilon} \subseteq P^{\alpha}_{\varepsilon/2} \subseteq \tilde{P}^{\alpha}_{\varepsilon/4}.$$

Proof. The proof is by induction on α. When $\alpha = 0$ all these sets coincide. Next we observe that both $'$ and * have the property that if $P \subseteq Q$, then for any ε, $P'_{\varepsilon} \subseteq Q'_{\varepsilon}$ and $P^*_{\varepsilon} \subseteq Q^*_{\varepsilon}$. Also it is easy to observe that for all ε and all P, $P^*_{\varepsilon} \subseteq P'_{\varepsilon/2} \subseteq P^*_{\varepsilon/4}$. So when $\alpha = \beta + 1$, if we assume $\tilde{P}^{\beta}_{\varepsilon} \subseteq P^{\beta}_{\varepsilon/2} \subseteq \tilde{P}^{\beta}_{\varepsilon/4}$ we have

$$\tilde{P}^{\alpha}_{\varepsilon} = (\tilde{P}^{\beta}_{\varepsilon})^*_{\varepsilon} \subseteq (P^{\beta}_{\varepsilon/2})^*_{\varepsilon} \subseteq (P^{\beta}_{\varepsilon/2})'_{\varepsilon/2} = P^{\alpha}_{\varepsilon/2}$$
$$P^{\alpha}_{\varepsilon/2} = (P^{\beta}_{\varepsilon/2})'_{\varepsilon/2} \subseteq (\tilde{P}^{\beta}_{\varepsilon/4})'_{\varepsilon/2} \subseteq (\tilde{P}^{\beta}_{\varepsilon/4})^*_{\varepsilon/4} = \tilde{P}^{\alpha}_{\varepsilon/4}$$

Finally, when λ is a limit, $\cap_{\alpha<\lambda}\tilde{P}^{\alpha}_{\varepsilon} \subseteq \cap_{\alpha<\lambda}P^{\alpha}_{\varepsilon/2} \subseteq \cap_{\alpha<\lambda}\tilde{P}^{\alpha}_{\varepsilon/4}$ follows because $\tilde{P}^{\alpha}_{\varepsilon} \subseteq P^{\alpha}_{\varepsilon/2} \subseteq \tilde{P}^{\alpha}_{\varepsilon/4}$ holds for all $\alpha < \lambda$. □

From Lemma 1.3.2 it is clear that for all α,

$$\forall\varepsilon P^{\alpha}_{\varepsilon} = \emptyset \iff \forall\varepsilon \tilde{P}^{\alpha}_{\varepsilon} = \emptyset,$$

and thus the notion of rank defined using the P^{α}_{ε} hierarchy coincides with the notion of rank defined using the $\tilde{P}^{\alpha}_{\varepsilon}$ hierarchy.

The formal statements "$|\mathbf{f}|_{\mathbf{KW}} \leq \alpha + \mathbf{1}$"

Before we can use the previous section's definition to formalize "$|f|_{KW} \leq \alpha+1$", we need the following lemma. Briefly, the lemma holds because membership in $\tilde{P}^{\alpha}_{\varepsilon}(I)$ is a local property.

Lemma 1.3.3. *Fix f and ε. For any closed $I \subseteq [0,1]$, any closed interval $[i,j]$, and any α,*

$$[i,j] \cap \tilde{P}^{\alpha}_{\varepsilon}(I) = \bigcap_{d>0} \tilde{P}^{\alpha}_{\varepsilon}([i-d, j+d] \cap I).$$

Proof. On the one hand, suppose that $x \notin [i,j] \cap \tilde{P}^{\alpha}_{\varepsilon}(I)$. If $x \notin [i,j]$ then eventually $x \notin [i-d, j+d]$. So assume that $x \in [i,j]$. Then $x \notin \tilde{P}^{\alpha}_{\varepsilon}(I)$, so x could not be in $\tilde{P}^{\alpha}_{\varepsilon}([i-d, j+d] \cap I)$ for any d, since $\tilde{P}^{\alpha}_{\varepsilon}([i-d, j+d] \cap I) \subseteq \tilde{P}^{\alpha}_{\varepsilon}(I)$ for all α.

For the other direction we proceed by induction on α. The relationship certainly holds when $\alpha = 0$. Suppose $\alpha = \beta + 1$ and suppose that $x \in [i,j] \cap \tilde{P}^{\alpha}_{\varepsilon}(I)$. We wish to show that $x \in \tilde{P}^{\alpha}_{\varepsilon}([i-d, j+d] \cap I)$, so fix δ, and we will proceed to find our witnesses. Since $x \in \tilde{P}^{\alpha}_{\varepsilon}(I)$, let $p < q, r < s \in B(x, \min(\delta, d/2)) \cap I$ be such that $[p,q] \cap [r,s] \cap \tilde{P}^{\beta}_{\varepsilon}(I) \neq \emptyset$ and $|\Delta_f(p,q) - \Delta_f(r,s)| > \varepsilon$. Then because $x \in [i,j]$, we have these same $p,q,r,s \in B(x,\delta) \cap [i-d, j+d] \cap I$, and in fact, because p,q,r,s are within $d/2$ of $[i,j]$, we have $p,q,r,s \in [i-d/2, j+d/2]$. If we can show that $[p,q] \cap [r,s] \cap \tilde{P}^{\beta}_{\varepsilon}([i-d, j+d] \cap I) \neq \emptyset$ then we are done.

Let $z \in [p,q] \cap [r,s] \cap \tilde{P}^{\beta}_{\varepsilon}(I)$. By the induction hypothesis,

$$z \in \bigcap_{\zeta > 0} \tilde{P}^{\beta}_{\varepsilon}([\max(p,r) - \zeta, \min(q,s) + \zeta] \cap I).$$

So in particular

$$z \in \tilde{P}^{\beta}_{\varepsilon}([\max(p,r) - d/2, \min(q,s) + d/2] \cap I) \subseteq \tilde{P}^{\beta}_{\varepsilon}([i-d, j+d] \cap I).$$

This completes the proof for the successor case.

Finally, if α is a limit ordinal, we have

$$\begin{aligned}
[i,j] \cap \tilde{P}^{\alpha}_{\varepsilon}(I) &= \bigcap_{\beta < \alpha} [i,j] \cap \tilde{P}^{\beta}_{\varepsilon}(I) \\
&= \bigcap_{\beta < \alpha} \bigcap_{d > 0} \tilde{P}^{\beta}_{\varepsilon}([i-d, j+d] \cap I) \\
&= \bigcap_{d > 0} \bigcap_{\beta < \alpha} \tilde{P}^{\beta}_{\varepsilon}([i-d, j+d] \cap I) \\
&= \bigcap_{d > 0} \tilde{P}^{\alpha}_{\varepsilon}([i-d, j+d] \cap I).
\end{aligned}$$

□

The definition of the rank of a function f uses transfinite recursion in order to calculate $P^{\alpha}_{f,\varepsilon}$ for each α while holding ε fixed. Thus, knowing the expressive complexity of "$|f|_{KW} \leq 1$" does not give us a foothold into the expressive complexity of "$|f|_{KW} \leq 2$", because "$|f|_{KW} \leq \alpha$" does not appear as a sub-expression of "$|f|_{KW} \leq \alpha + 1$". The sub-expression which does persist, and on which it is almost appropriate to transfinitely recurse, is "$[i,j] \cap \tilde{P}^{\alpha} = \emptyset$", where $[i,j]$ is some arbitrary interval. Lemma 1.3.3 lets us express this intersection in statements of the form "$\tilde{P}^{\alpha}([i,j]) = \emptyset$", and so this last expression is a useful core concept. Its expressive complexity is $\Sigma_{2\alpha}$, as seen in the next proposition.

Proposition 1.3.4. *Let $\alpha > 0$ be a constructive ordinal, $\varepsilon, i, j \in \mathbb{Q}$ with $\varepsilon > 0$ and $0 \leq i < j \leq 1$. The set of e such that $\tilde{P}^{\alpha}_{f_e,\varepsilon}([i,j]) = \emptyset$ is $\Sigma_{2\alpha}$, uniformly in α, ε, i and j.*

Proof. We carry along an arbitrary index e and oscillation sensitivity ε, so to reduce clutter we write f instead of f_e, and $\tilde{P}^{\alpha}$ instead of $\tilde{P}^{\alpha}_{f,\varepsilon}$.

In general, when $\alpha = \beta + 1$,

$$\tilde{P}^{\alpha}([i,j]) = [i,j] \setminus \bigcup\{I : \forall p,q,r,s \in I \left([p,q] \cap [r,s] \cap \tilde{P}^{\beta}([i,j]) = \emptyset \vee |\Delta_f(p,q) - \Delta_f(r,s)| \leq \varepsilon\right)\}$$

where I ranges over intervals open in $[i,j]$. Since the I are closed under taking subsets, it suffices to let I range over intervals open in $[i,j]$ with rational endpoints. So $\tilde{P}^{\alpha}([i,j]) = \emptyset$ if and only if these I cover $[i,j]$. If the I do cover, then by compactness there is a rational δ such that for all $x \in [i,j]$, $B(x,\delta) \subseteq I$ for some I. Thus there is a δ such that for any open interval U with rational endpoints where $\mathrm{diam}(U) < \delta$, $U \subseteq I$ for some I. On the other hand, if the I do not cover, then there cannot be any such δ. Thus if $\alpha = \beta + 1$,

$$\tilde{P}^{\alpha}([i,j]) = \emptyset \iff \exists \delta > 0 \forall c \in [i,j] \forall p,q,r,s \in B(c,\delta) \cap [i,j] \left([p,q] \cap [r,s] \cap \tilde{P}^{\beta}([i,j]) = \emptyset \vee |\Delta_f(p,q) - \Delta_f(r,s)| \leq \varepsilon\right)$$

where all quantifiers range over the rationals.

When $\beta = 0$, $[p,q] \cap [r,s] \cap \tilde{P}^{\beta}([i,j]) = \emptyset \Leftrightarrow [p,q] \cap [r,s] \cap [i,j] = \emptyset$, so the above statement is Σ_2 uniformly in e, ε, i, and j.

When $\beta > 0$, we have

$$[p,q] \cap [r,s] \cap \tilde{P}^{\beta}([i,j]) = \emptyset \iff \exists \zeta \tilde{P}^{\beta}([\max(p,r) - \zeta, \min(q,s) + \zeta] \cap [i,j]) = \emptyset.$$

which follows from Lemma 1.3.3 and compactness. Thus with the assumption that $\tilde{P}^{\beta}([c,d]) = \emptyset$ is $\Sigma_{2\beta}$ uniformly in all variables, then $\tilde{P}^{\beta+1}([i,j]) = \emptyset$ is $\Sigma_{2\beta+2}$, uniformly in all variables.

Finally, suppose that α is a limit, given as a uniform supremum $\alpha = \sup_n \beta_n$. Then by compactness and the definition of $\tilde{P}^{\alpha}$ for α a limit,

$$P^{\alpha}([i,j]) = \emptyset \iff \exists n \tilde{P}^{\beta_n}([i,j]) = \emptyset.$$

So assuming that $\tilde{P}^{\beta_n}([i,j]) = \emptyset$ is uniformly $\Sigma_{2\beta_n}$ in all variables including n, we see that $\tilde{P}^{\alpha}([i,j]) = \emptyset$ is uniformly Σ_{α}, which is the same as $\Sigma_{2\alpha}$ since α is a limit. □

Proposition 1.3.5. *For any constructive $\alpha > 0$, $D_{\alpha+1}$ is $\Pi_{2\alpha+1}$, uniformly in α.*

Proof. We have

$$e \in D_{\alpha+1} \iff f_e \in C[0,1] \wedge \forall \varepsilon[\tilde{P}^{\alpha}_{f_e,\varepsilon} = \emptyset]$$

where ε ranges over positive rationals. Recall that "$f_e \in C[0,1]$" is Π_2, and by Proposition 1.3.4, $\tilde{P}^{\alpha}_{f_e,\varepsilon} = \emptyset$ is $\Sigma_{2\alpha}$. Thus the right hand side is a $\Pi_{2\alpha+1}$ statement, uniformly in α and e. □

Having differentiability rank at most α is $\Pi_{2\alpha+1}$-complete

In this section, we provide a many-one reduction in the other direction, from $\overline{\emptyset_{(2\alpha+1)}}$ to $D_{\alpha+1}$ using Theorem 7.

To set up this step, we happen to need only a certain class of $C[0,1]$ functions which can be structurally represented by well-founded trees according to a recipe reminiscent of Mazurkiewicz's original reduction. This allows us to construct a function of the right rank through an intermediate step of constructing a tree with the right structure.

In Section 1.3 we construct special $C[0,1]$ functions which reflect the structure of well-founded trees on $\mathbb{N}^{<\mathbb{N}}$. In Section 1.3, we show that the limpsup rank of the tree agrees with the differentiability rank of the functions that the tree generates, when a fixed arbitrary value for ε is used.

Section 1.3 combines the results of the previous sections with the additional ingredient of varying ε to obtain the final result.

Making differentiable functions out of well-founded trees

The idea of this section is to set up countably many closed disjoint intervals in $[0,1]$, put the intervals in bijective correspondence with $\mathbb{N}^{<\mathbb{N}}$, and then given a tree $T \subseteq \mathbb{N}^{<\mathbb{N}}$, define f_T as a sum of continuously differentiable bumps supported on each of the intervals which correspond to $\sigma \in T$. These functions are structurally similar to the ones described in Section 1.3. If $S = \{\rho : \sigma^\frown\rho \in T\}$ then a shrunken version of f_S can be found in f_T. Furthermore, if $\tau \supset \sigma$, then the bump corresponding to τ is in the shadow of the bump corresponding to σ. The intervals are arranged so that the resulting f_T has a differentiability rank which can be computed from T in a way that is described in the next section.

In the following definition, the choices of the constants $\frac{1}{2}$ and $\frac{1}{4}$ and the bounds on p and p' are arbitrary, but consistent with each other. The requirement $b_n - a_n < (a_n - \frac{1}{4})^2$ is what keeps f_T everywhere differentiable.

Definition 1.3.9. *Let $p : [0,1] \to \mathbb{R}$ be a computable function satisfying*

1. *p is continuously differentiable*
2. $p(\frac{1}{2}) = \frac{1}{2}$
3. $p(0) = p(1) = p'(0) = p'(1) = 0$
4. $||p|| < 1$ *and* $||p'|| < 2$

Let $\{[a_n, b_n]\}_{n\in\mathbb{N}}$ be any computable sequence of intervals with rational endpoints satisfying

1. *Each interval is contained in $(\frac{1}{4}, \frac{1}{2})$*
2. *$b_{n+1} < a_n < b_n$ for each n.*
3. $\lim_{n\to\infty} a_n = \frac{1}{4}$

4. $b_n - a_n < (a_n - \frac{1}{4})^2$ *for each* n

Then for any well-founded tree $T \in \mathbb{N}^{<\mathbb{N}}$*, define* f_T *as follows.*

1. *If* T *is empty,* $f_T \equiv 0$.
2. *Otherwise,* $f_T = p[\frac{1}{2}, 1] + \sum_{n=0}^{\infty} f_{T_n}[a_n, b_n]$

Recall that $f[a, b]$ denotes a copy of f proportionally resized to have domain $[a, b]$, and that T_n denotes $\{\sigma : \langle n \rangle^\frown \sigma \in T\}$, the nth subtree of T. Now we verify that the above definition produces well-defined computable differentiable functions.

Proposition 1.3.6. *For any well-founded computable tree* $T \in \mathbb{N}^{<\mathbb{N}}$*:*

1. f_T *is uniformly computable in* T
2. f_T *is differentiable*
3. $f_T(0) = f_T(1) = f'_T(0) = f'_T(1) = 0$
4. $||f_T|| < 1$ *and* $||f'_T|| < 2$

Proof. Proceeding by induction on the rank of the tree, in the base case all four properties are satisfied. Assume they hold for all trees of rank less than $|T|$. Then the sequence f_{T_n} is uniformly computable with each $||f_{T_n}|| < 1$. Then on any interval whose closure does not contain $\frac{1}{4}$, f_T is equal to a uniformly determined finite sum of computable functions, and is thus computable. And for ε sufficiently small, we claim that $|f_T\left((\frac{1}{4} - \varepsilon, \frac{1}{4} + \varepsilon)\right)| < \varepsilon^2$. This follows because $||f_{T_n}|| < 1$ by induction and because the intervals $[a_n, b_n]$ are disjoint, so for any x in such an interval we have the bound

$$|f_T(x)| = f_{T_n}[a_n, b_n](x) \leq ||f_{T_n}||(b_n - a_n) \leq 1 \cdot (a_n - \frac{1}{4})^2 \leq (x - \frac{1}{4})^2 < \varepsilon^2.$$

Therefore f_T is uniformly computable in T. Similarly, assuming f_{T_n} are each differentiable with $f_{T_n}(0) = f_{T_n}(1) = f'_{T_n}(0) = f'_{T_n}(1) = 0$, then each $f_{T_n}[a_n, b_n]$ is differentiable. Then f_T is certainly differentiable at any point $x \neq \frac{1}{4}$, since on some neighborhood of that point f_T is equal to a finite sum of differentiable functions. On the other hand, in the vicinity of $\frac{1}{4}$, f_T satisfies $|f_T(x)| \leq (x - \frac{1}{4})^2$, so f_T is differentiable at $\frac{1}{4}$ as well. Because $f_T \upharpoonright [0, \frac{1}{4}] \equiv 0$ and $p(1) = p'(1) = 0$, we have $f_T(0) = f_T(1) = f'_T(0) = f'_T(1) = 0$. Finally, $||f_T|| < 1$ and $||f'_T|| < 2$ by induction, because $||p|| < 1$,$||p'|| < 2$, and $||f_{T_n}|| < 1$,$||f'_{T_n}|| < 2$, for each n, and the shrunken copies $p[\frac{1}{2}, 1]$ and $f_{T_n}[a_n, b_n]$ have disjoint support. □

We close this section with some comments about why this f_T is defined as it is, using the concepts from Section 1.3. Note that for every nonempty S,

$$\Delta_{f_S}(0, \frac{3}{4}) = \frac{1}{3} \text{ and } \Delta_{f_S}(0, \frac{1}{2}) = 0.$$

Now for each n, $f_{T_n}[a_n, b_n]$ is a proportionally shrunken copy of f_{T_n}, so unless T_n is empty, $f_{T_n}[a_n, b_n]$ contributes a bump and its pair of secants with slopes 0 and $\frac{1}{3}$. Thus a tree with infinitely many children of the root has infinitely many pairs of these disagreeing secants. If we construct T_n so that f_{T_n} has a large rank, the nth disagreeing pair of secants will be visible for many iterations of the rank-ascertaining process for f_T, because $P^{\alpha}_{f_T} \cap [a_n, \frac{a_n+b_n}{2}]$ will be nonempty for many iterations. If we construct T so that f_{T_n} has large rank for infinitely many n, these disagreeing pairs of secants can make a contribution to raising the Kechris-Woodin rank of f_T. This can happen in two ways: if $|f_{T_n}|_{KW} = \alpha + 1$ for infinitely many n, then $|f_T| \geq \alpha + 2$. And if the ranks of the f_{T_n} are unbounded below a limit ordinal λ, then $|f_T|_{KW} = \lambda + 1$.

Th relationship between the limsup rank and the Kechris-Woodin rank

We will now show that when a function is generated from a tree in the way described above, its Kechris-Woodin rank can be read right off the tree. Furthermore, we will see that this function's rank can already be witnessed at a fixed oscillation sensitivity $\varepsilon = \frac{1}{4}$. That is, the rank of f_T is always a successor, and when $|f_T|_{KW} = \alpha+1$, then $\tilde{P}^{\alpha}_{f,\frac{1}{4}} \neq \emptyset$. The limsup rank on trees corresponds to the differentiability rank of the functions they generate, $|T|_{ls} = |f_T|_{KW}$.

The following two straightforward lemmas which we will use later are woven into the proof of Fact 3.5 in [14]. For the purposes of exposition, we state and prove them here.

Lemma 1.3.7. *If $U \subseteq [0, 1]$ is open and $f \restriction U = g \restriction U$, then for all α and ε, $P^{\alpha}_{f,\varepsilon} \cap U = P^{\alpha}_{g,\varepsilon} \cap U$.*

Proof. By induction on α. The base and limit cases are trivial. Suppose that $P^{\alpha}_{f,\varepsilon} \cap U = P^{\alpha}_{g,\varepsilon} \cap U$. Fix $x \in U$ and let λ be small enough that $B(X, \lambda) \subseteq U$. Then $x \in P^{\alpha+1}_{f,\varepsilon}$ if and only if for all $\delta < \lambda$ there are $p, q, r, s \in B(x, \delta)$ such that $|\Delta_f(p, q) - \Delta_f(r, s)| \geq \varepsilon$ and $[p, q] \cap [r, s] \cap P^{\alpha}_{f,\varepsilon} \neq \emptyset$. Since $p, q, r, s \in B(x, \delta) \subseteq B(x, \lambda) \subseteq U$, we have $[p, q] \cap [r, s] \cap P^{\alpha}_{f,\varepsilon} \neq \emptyset$ if and only if $[p, q] \cap [r, s] \cap P^{\alpha}_{g,\varepsilon} \neq \emptyset$. Thus $x \in P^{\alpha+1}_{f,\varepsilon}$ if and only if $x \in P^{\alpha+1}_{g,\varepsilon}$. □

Recall that for any function $f \in C[0, 1]$, we write $f[a, b]$ to denote a proportionally shrunken verson of f. By definition, $f[a, b]$ is the function which is identically 0 outside of $[a, b]$, and for $x \in [a, b]$, $f[a, b](x) = (b - a)f(\frac{x-a}{b-a})$. Similarly, for any real number $c \in [0, 1]$ and any interval $[a, b]$, let $c[a, b] = a + c(b - a)$. The point is that c is to f as $c[a, b]$ is to $f[a, b]$.

Lemma 1.3.8. *Let $f \in C[0, 1]$ be a differentiable function satisfying $f(0) = f(1) = f'(0) = f'(1) = 0$. Let $[a, b] \subseteq [0, 1]$ be an interval with rational endpoints. Then $|f|_{KW} = |f[a, b]|_{KW}$. Furthermore, for any ordinal α and for all $x \in [0, 1]$,*

1. $x \in P^{\alpha}_{f,\varepsilon} \implies x[a, b] \in P^{\alpha}_{f[a,b],\varepsilon}$
2. $x[a, b] \in P^{\alpha}_{f[a,b],\varepsilon} \implies x \in P^{\alpha}_{f,\varepsilon/2}$

Proof. Proceeding by induction, it is clear that the both items holds when $\alpha = 0$. The limit case is also trivial.

Assume the first item holds for some α. If $x \in P^{\alpha+1}_{f,\varepsilon}$, then the collection of all the tuples p, q, r, s which witness this can be mapped to a collection of tuples $p[a,b], q[a,b], r[a,b], s[a,b]$ which witness $x[a,b] \in P^{\alpha+1}_{f[a,b],\varepsilon}$. That proves the first item.

On the other hand, suppose the second item holds for some α. If $x[a,b] \in P^{\alpha+1}_{f[a,b],\varepsilon}$ and $x \in (0,1)$ (i.e. x is not an endpoint), then as above corresponding witnesses can always be chosen for sufficiently small neighborhoods of x, so $x \in P^{\alpha+1}_{f,\varepsilon} \subseteq P^{\alpha+1}_{f,\varepsilon/2}$. Last we consider the endpoint case: suppose $a \in P^{\alpha+1}_{f[a,b],\varepsilon}$ (and the case $b \in P^{\alpha+1}_{f[a,b],\varepsilon}$ is of course just the same). Because $a = 0[a,b]$ and $f'(0) = 0$, let λ be small enough that for all distinct $p, q \in B(a, \lambda)$ with $p \leq a \leq q$, $|\Delta_{f[a,b]}(p,q)| < \varepsilon/4$. Then for each $\delta > 0$, there are $p, q, r, s \in B(a, \min(\lambda, (b-a)\delta))$ such that $|\Delta_{f[a,b]}(p,q) - \Delta_{f[a,b]}(r,s)| \geq \varepsilon$ and $[p,q] \cap [r,s] \cap P^{\alpha}_{f[a,b],\varepsilon} \neq \emptyset$. Then without loss of generality, $|\Delta_{f[a,b]}(p,q)| \geq \varepsilon/2$, so $a < p < q$. If also $a < r < s$, then we are done since the corresponding $\frac{p-a}{b-a}$, etc. can be used as the witness for δ. It is impossible that $r < s < a < p < q$ because $[p,q] \cap [r,s] \neq \emptyset$. In the last case, if $r < a < s$, this implies that $|\Delta_{f[a,b]}(r,s)| < \varepsilon/4$, so $|\Delta_{f[a,b]}(p,q)| \geq 3\varepsilon/4$. But then also $|\Delta_{f[a,b]}(a,s)| < \varepsilon/4$, and thus $|\Delta_{f[a,b]}(p,q) - \Delta_{f[a,b]}(a,s)| \geq \varepsilon/2$. Also $[p,q] \cap [a,s] = [p,q] \cap [r,s]$, and there is some $y \in [p,q] \cap [a,s] \cap P^{\alpha}_{f[a,b],\varepsilon}$, and by induction $\frac{y-a}{b-a} \in P^{\alpha}_{f,\varepsilon/2}$. Therefore $\frac{p-a}{b-a}, \frac{q-a}{b-a}, 0, \frac{s-a}{b-a}$ will do, and thus $x \in P^{\alpha+1}_{f,\varepsilon/2}$.

Finally, note that by the previous lemma, $P^{\alpha}_{f[a,b],\varepsilon} \cap ([0,1] \setminus [a,b]) = \emptyset$ for any $\alpha > 0$. Therefore, $|f|_{KW} = |f[a,b]|_{KW}$. □

The next proposition shows that for any well-founded tree T, the differentiable function f_T defined in the previous section has rank $|f_T|_{KW} = |T|_{ls}$, and that the rank of f_T is witnessed at oscillation sensitivity $\varepsilon = \frac{1}{4}$.

Proposition 1.3.9. *For any nonempty well-founded tree $T \in \mathbb{N}^{<\mathbb{N}}$,*

1. *$|T|_{ls}$ is a successor,*
2. *The function f_T is differentiable with $|f_T|_{KW} = |T|_{ls}$, and*
3. *Letting $|T|_{ls} = \alpha + 1$, we have $P^{\alpha}_{f,\frac{1}{4}} \neq \emptyset$.*

Proof. The proof is by induction on the usual rank of the tree. If T is just a root (smallest option for the rank of the the tree since the statement is for nonempty trees only) then f_T is just $p[\frac{1}{2}, 1]$, so it is continuously differentiable with $|f_T|_{KW} = 1$. For each n, $T_n = \emptyset$ so $|T_n|_{ls} = 0$ so $\sup_n |T_n|_{ls} = \limsup_n |T_n|_{ls} = 0$, so $|T|_{ls} = 1$.

If T is more than a root, assume the lemma holds for each of the subtrees T_n. First we show that $|f_T|_{KW} \geq |T|_{ls}$. Fix n and let $|T_n|_{ls} = \alpha + 1$. Then by the inductive hypothesis $|f_{T_n}|_{KW} = \alpha + 1$ and $P^{\alpha}_{f_{T_n},\frac{1}{4}} \neq \emptyset$. By Lemma 1.3.8, $x \in P^{\alpha}_{f_{T_n},\frac{1}{4}} \implies x[a,b] \in P^{\alpha}_{f_{T_n}[a_n,b_n],\frac{1}{4}}$, so $P^{\alpha}_{f_{T_n}[a_n,b_n],\frac{1}{4}} \neq \emptyset$. Because the $[a_n, b_n]$ are closed and disjoint from each other and from $[\frac{1}{2}, 1]$, there is an $\varepsilon > 0$ such that $f_T \upharpoonright (a_n - \varepsilon, b_n + \varepsilon) = f_{T_n}[a_n, b_n] \upharpoonright (a_n - \varepsilon, b_n + \varepsilon)$, and therefore using Lemma 1.3.7, $P^{\alpha}_{f_T,\frac{1}{4}} \cap (a_n - \varepsilon, b_n + \varepsilon) = P^{\alpha}_{f_{T_n}[a_n,b_n],\frac{1}{4}} \cap (a_n - \varepsilon, b_n + \varepsilon) \neq \emptyset$. Therefore $P^{\alpha}_{f_T,\frac{1}{4}} \neq \emptyset$ and thus $|f_T|_{KW} \geq \alpha + 1$. So $|f_T|_{KW} \geq \sup_n |T_n|_{ls}$.

Now let us show that $|f_T|_{KW} \geq (\limsup_n |T_n|_{ls}) + 1$. Let $\alpha = \limsup_n |T_n|_{ls}$. We will show that $\frac{1}{4} \in P^{\alpha}_{f_T,\frac{1}{4}}$. First we show that for any $\beta < \alpha$, $\frac{1}{4} \in P^{\beta}_{f_T,\frac{1}{4}}$. There are infinitely many n such that $|T_n|_{ls} > \beta$, so $P^{\beta}_{f_{T_n},\frac{1}{4}} \neq \emptyset$ for infinitely many n by the inductive hypothesis, so $P^{\beta}_{f_{T_n}[a_n,b_n],\frac{1}{4}} \neq \emptyset$ for infinitely many n by Lemma 1.3.8. By Lemma 1.3.7, $P^{\beta}_{f_{T_n}[a_n,b_n],\frac{1}{4}} \subseteq P^{\beta}_{f_T,\frac{1}{4}}$. Because $\lim_{n\to\infty} a_n = \frac{1}{4}$, and infinitely many $[a_n, b_n]$ contain an element of $P^{\beta}_{f_T,\frac{1}{4}}$, $\frac{1}{4}$ is a limit point of $P^{\beta}_{f_T,\frac{1}{4}}$. Because this set is closed, $\frac{1}{4}$ must be in it as well. Thus $\frac{1}{4} \in P^{\beta}_{f_T,\frac{1}{4}}$ for all $\beta < \alpha$. If α is as limit, this implies $\frac{1}{4} \in P^{\alpha}_{f_T,\frac{1}{4}}$, so $|f_T|_{KW} > \alpha$ if α is a limit. Now suppose α is a successor. Let $\alpha = \beta + 1$. Let U be a neighborhood of $\frac{1}{4}$, and let n be chosen such that $[a_n, b_n] \subseteq U$ and $P^{\beta}_{f_{T_n}[a_n,b_n],\frac{1}{4}} \neq \emptyset$. Then $\Delta_{f_T}(a_n, \frac{3}{4}[a_n, b_n]) = \frac{1}{3}$ and $\Delta_{f_T}(a_n, \frac{1}{2}[a_n, b_n]) = 0$, and $[a_n, \frac{3}{4}[a_n, b_n]] \cap [a_n, \frac{1}{2}[a_n, b_n]] \cap P^{\beta}_{f_T,\frac{1}{4}} \neq \emptyset$. Therefore $\frac{1}{4} \in P^{\beta+1}_{f_T,\frac{1}{4}}$, so again $|f_T|_{KW} > \alpha$. This completes the claim that $|f_T|_{KW} \geq |T|_{ls}$.

Now let us show that $|f_T|_{KW} \leq |T|_{ls}$. First, let $\alpha = \sup_n |T_n|_{ls}$. Note that $\alpha > 0$ because the case of T being only a root was already considered separately. For each n, $|T_n|_{ls} \leq \alpha$, so by induction $|f_{T_n}|_{KW} = |f_{T_n}[a_n, b_n]|_{KW} \leq \alpha$. So for each n and ε we have $P^{\alpha}_{f_{T_n}[a_n,b_n],\varepsilon} = \emptyset$ and also $P^{\alpha}_{p[\frac{1}{2},1],\varepsilon} = \emptyset$. Cover $[0, 1] \setminus \{\frac{1}{4}\}$ with open intervals U such that each U intersects at most one of the $[a_n, b_n]$ or $[\frac{1}{2}, 1]$. Then for each such interval and each ε, $P^{\alpha}_{f_T,\varepsilon} \cap U = P^{\alpha}_{f_{T_n}[a_n,b_n],\varepsilon} \cap U = \emptyset$, or $P^{\alpha}_{f_T,\varepsilon} \cap U = P^{\alpha}_{p[\frac{1}{2},1],\varepsilon} \cap U = \emptyset$, respectively. Therefore, for all ε, $P^{\alpha}_{f_T,\varepsilon} \subseteq \{\frac{1}{4}\}$. If $\limsup_n |T_n|_{ls} = \sup_n |T_n|_{ls}$, then $|T|_{ls} = \alpha + 1$, so this is enough: $P^{\alpha+1}_{f_T,\varepsilon} = \emptyset$ for all ε.

On the other hand, suppose $\limsup_n |T_n|_{ls} < \sup_n |T_n|_{ls}$. Then $\alpha = |T|_{ls} = \sup_n |T_n|_{ls}$ is a successor, because the induction hypothesis guarantees $|T_n|_{ls}$ is always a successor, and therefore if the sup were a limit, it would be equal to the limsup. Let $\alpha = \beta + 1$. Then eventually $|T_n|_{ls} \leq \beta$. Let V be an open neighborhood of $\frac{1}{4}$ such that $[a_n, b_n] \cap V \neq \emptyset$ implies $|T_n|_{ls} \leq \beta$. Covering $V \setminus \{\frac{1}{4}\}$ with open intervals U as before, we find $P^{\beta}_{f_{T_n}[a_n,b_n],\varepsilon} \cap U = \emptyset$ for each such $U \subseteq V$ and each ε, so $P^{\beta}_{f_T,\varepsilon} \cap V \subseteq \{\frac{1}{4}\}$, so $P^{\beta+1}_{f_T,\varepsilon} \cap V = \emptyset$. Therefore $P^{\beta+1}_{f_T,\varepsilon} = \emptyset$. Thus $|f_T|_{KW} \leq \sup_n |T_n|_{ls}$. □

Recognizing functions of rank α is $\Pi_{2\alpha+1}$-hard

In this section we obtain the final result by consideration of what can be encoded into the oscillation sensitivity ε at which a function's rank is witnessed. For this last step, it is necessary to consider functions again instead of trees, because with the trees we only produce functions made of bumps with all the same proportions. In the next theorem, we use functions made of increasingly shallow bumps, and encode the last jump into the uncertainty of how small ε will have to be in order for the bumps which determine the function's rank to be detectable.

Theorem 6. *Uniformly in a constructive ordinal $\alpha > 0$ and x, one may find a computable $f \in C[0,1]$ satisfying*

- $x \notin \emptyset_{(2\alpha+1)} \to |f|_{KW} \leq \alpha$
- $x \in \emptyset_{(2\alpha+1)} \to |f|_{KW} = \alpha + 1$

Proof. Given α, x, compute f as follows. Similar to earlier, let $\{[a_n, b_n]\}_{n\in\mathbb{N}}$ be any computable sequence of intervals with rational endpoints satisfying

- Each interval is contained in $(0,1)$
- $b_{n+1} < a_n < b_n$ for each n.
- $\lim_{n\to\infty} a_n = 0$
- $b_n - a_n < a_n^2$

Let g be a computable function satisfying for all x and X,

$$x \in X'' \iff \exists s[g(x,s) \notin X'].$$

Then

$$x \in \emptyset_{(2\alpha+1)} \iff \exists s[g(x,s) \notin \emptyset_{(2\alpha)}].$$

For any s, let $T(s)$ be the tree guaranteed by Theorem 2 with input $(\alpha, g(x,s))$. Thus $|T(s)|_{ls} = \alpha + 1$ if $g(x,s) \notin \emptyset_{(2\alpha)}$ and $|T(s)|_{ls} \leq \alpha$ otherwise. Then define

$$f = \sum_{s=0}^{\infty} \frac{1}{s+1} f_{T(s)}[a_s, b_s].$$

Recall that Proposition 1.3.6 guarantees that $||f_{T(s)}|| < 1$, so

$$||\frac{1}{s+1} f_{T(s)}[a_s, b_s]|| < \frac{b_s - a_s}{s+1} < \frac{a_s^2}{s+1}.$$

On neighborhoods bounded away from 0, f is a uniformly presented sum of finitely many computable differentiable functions, but f lives in the envelope of x^2, so it is computable near 0 as well. Thus f is computable and differentiable.

Suppose $x \notin \emptyset_{(2\alpha+1)}$. Then for each s, $g(x,s) \in \emptyset_{(2\alpha)}$, so $|T(s)|_{ls} \leq \alpha$, so $|f_{T(s)}|_{KW} \leq \alpha$. For each $z \neq 0$, there is a neighborhood U of z which intersects exactly one of the $[a_s, b_s]$. Because $P^{\alpha}_{f_{T(s)}[a_s,b_s],\varepsilon} = \emptyset$ for all ε, and $f_{T(s)}[a_s, b_s]$ coincides with f on U, Lemma 1.3.7 implies that $z \notin P^{\alpha}_{f,\varepsilon}$ for any ε. On the other hand, fix ε and let $z = 0$. Then for any s, by Proposition 1.3.6, $||f'_{T(s)}|| < 2$, so

$$||\frac{1}{s+1} f'_{T(s)}[a_s, b_s]|| = \frac{1}{s+1}||f'_{T(s)}|| < \frac{2}{s+1}.$$

Let S be large enough that $\frac{4}{S+1} < \varepsilon$. Then for all $p, q, r, s \in [0, b_S)$,

$$\begin{aligned} |\Delta_f(p,q) - \Delta_f(r,s)| &\leq |\Delta_f(p,q)| + |\Delta_f(r,s)| \\ &\leq 2||f' \upharpoonright [0, b_S)|| < \frac{4}{S+1} < \varepsilon, \end{aligned}$$

so $0 \notin P^{\beta}_{f,\varepsilon}$ for any $\beta > 0$. Therefore $P^{\alpha}_{f,\varepsilon} = \emptyset$ for all ε and $|f|_{KW} \leq \alpha$.

On the other hand, suppose that $x \in \emptyset_{(2\alpha+1)}$. Let s be such that $g(x, s) \notin \emptyset_{(2\alpha)}$. Then $T(s)$ has rank $\alpha + 1$. So $|f_{T(s)}|_{KW} = \alpha + 1$, and this rank is visible at oscillation sensitivity $\varepsilon = \frac{1}{4}$ by Proposition 1.3.9. So also $|\frac{1}{s+1} f_{T(s)}|_{KW} = |\xi^{-1}(a)|_{\mathcal{O}} + 1$, and this rank is visible at oscillation sensitivity $\varepsilon = \frac{1}{4(s+1)}$. Therefore by Lemmas 1.3.8 and 1.3.7,

$$\emptyset \neq P^{\alpha}_{\frac{1}{s+1} f_{T(s)}[a_s,b_s], \frac{1}{4(s+1)}} \subseteq P^{\alpha}_{f, \frac{1}{4(s+1)}}.$$

Thus $|f|_{KW} \geq \alpha + 1$. Also, for each s, $|f_{T(s)}|_{KW} \leq \alpha + 1$, and $0 \notin P^{\beta}_{f,\varepsilon}$ for any ε and any $\beta > 0$, so just as above, $|f|_{KW} \leq \alpha + 1$ always. So in fact $|f|_{KW} = \alpha + 1$. □

Therefore, we have the following:

Theorem 7. *For each nonzero $\alpha < \omega_1^{CK}$, $D_{\alpha+1}$ is $\Pi_{2\alpha+1}$-complete.*

Proof. By Proposition 1.3.5, $D_{\alpha+1} \leq_m \overline{\emptyset_{(2\alpha+1)}}$. By Theorem 6, $\overline{\emptyset_{(2\alpha+1)}} \leq_m D_{\alpha+1}$. □

Theorem 8. *For any limit ordinal $\lambda < \omega_1^{CK}$, D_λ is Σ_λ-complete.*

Proof. First we show that D_λ is Σ_λ. Given $\lambda = \sup_n \beta_n$, we have $e \in D_\lambda \iff \exists n[e \in D_{\beta_n+1}]$. Each $e \in D_{\beta_n+1}$ is $\Pi_{2\beta_n+1}$ by Proposition 1.3.5, so D_λ is Σ_λ.

Now we show that D_λ is Σ_λ-complete by giving an appropriate reduction. We claim that

$$x \in \emptyset_{(\lambda)} \iff |f_T|_{KW} < \lambda,$$

where T is the tree constructed in Theorem 2 from input (λ, x). That lemma guarantees first that $x \notin \emptyset_{(\lambda)}$ implies $|T|_{ls} = \lambda + 1$. Conversely, if $x \in \emptyset_{(\lambda)}$ we have $|T|_{ls} \leq \lambda$. But by Proposition 1.3.9, the limsup rank of a tree is always a successor, so in fact $x \in \emptyset_{(\lambda)}$ implies $|T|_{ls} < \lambda$. Thus $x \in \emptyset_{(\lambda)} \iff |T|_{ls} < \lambda \iff |f_T|_{KW} < \lambda$. □

1.4 Applications to the Denjoy rank

Having considered differentiability on $[0,1]$, we now consider integration. First we consider the Lebesgue integral. A function $F : [a,b] \to \mathbb{R}$ is absolutely continuous if for all ε there is a δ such that whenever $(a_i, b_i)_{i<k}$ is a finite sequence of disjoint subintervals of $[a,b]$ with $\sum_i |b_i - a_i| < \delta$ then $\sum_i |F(b_i) - F(a_i)| < \varepsilon$. The following is well known:

Theorem 9. *Let $F : [a,b] \to \mathbb{R}$. The following are equivalent:*

1. *F is absolutely continuous.*
2. *There is a Lebesgue integrable function g such that $\int_a^x g(x)dx = F(x) + F(a)$ for $x \in [a,b]$*
3. *F is a.e. differentiable and F' is Lebesgue integrable and $\int_a^x F'(x)dx = F(x) + F(a)$ for $x \in [a,b]$.*

We wish to consider the descriptive complexity of the set of Lebesgue integrable functions, but a problem one immediately encounters is how to represent the function to be integrated. Considering only the computable f trivializes the problem because every computable function is continuous and every continuous function on $[0,1]$ is Lebesgue integrable. Our solution is to consider instead AC, the image of Lebesgue integration, and a subset of the continuous functions. In fact, one may observe from Theorem 9 that the a.e. equivalence classes of the Lebesgue integrable functions are in one-to-one correspondence with their indefinite integrals in AC satisfying $F(a) = 0$. This motivates the question: What is the descriptive complexity of $\{e : F_e \in AC\}$, where F_e is the eth computable function? In Section 1.4 we show this set is Π_3 complete.

Lebesgue was dissatisfied with his integral because there are everywhere differentiable functions that are not absolutely continuous, and thus they cannot be recovered from their derivatives using Lebesgue integration. For example, $x^2 \sin(\frac{1}{x^2})$ has this property.

In 1912, Denjoy devised a transfinite process, generalizing Lebesgue integration, with the goal of being able to recover a primitive from every derivative. This process, known as the *narrow Denjoy integral*, has many equivalent definitions, among them the integrals of Perron, Kurzweil and Henstock, which we will not discuss here. This process succeeds in integrating every derivative, and was greeted by Lebesgue with great enthusiasm.

The process of narrow Denjoy integration in fact integrates more than just derivatives of everywhere differentiable functions. It also successfully recovers the primitive whenever applied to the derivative of a *nearly everywhere differentiable* function, which is a function differentiable at all but countably many points; and for some a.e. differentiable functions with uncountably many points of non-differentiability, it also recovers the primitive. However, it is not possible in principle to recover every a.e. differentiable function from its derivative because, as we will see below, there is an a.e. differentiable function f such that $f' = 0$ almost everywhere, but f is not a.e. equivalent to a constant function. However, there is a characterization of the image of narrow Denjoy integration analogous to Theorem 9 above. To get to the characterization theorem, some definitions are required.

Definition 1.4.1. *The* oscillation *of a function F on an interval I, denoted $\omega(F, I)$, is* $\sup_{x,y\in I} |F(y) - F(x)|$.

Definition 1.4.2. *Let $F : [a, b] \to \mathbb{R}$ and $E \subseteq [a, b]$ a closed set.*

1. *We say F is* absolutely continuous in the restricted sense *on E, and write $F \in AC_*(E)$, if for each $\varepsilon > 0$ there exists a $\delta > 0$ such that whenever $(a_i, b_i)_{i<k}$ is a finite sequence of disjoint subintervals of $[a, b]$ with $a_i, b_i \in E$ and $\sum_i |b_i - a_i| < \delta$ then $\sum_i \omega(F, [a_i, b_i]) < \varepsilon$.*

2. *We say F is* generalized absolutely continuous in the restricted sense *on E, and write $F \in ACG_*(E)$ if $F \restriction E$ is continuous on E and E can be written as a countable union of sets on each of which F is AC_*. We write $F \in ACG_*$ if the set E is clear from context.*

And here is the analogous theorem. The third equivalent condition is included for the completeness of the analogy, but we will not prove it. After stating the theorem we will give the constructive definition of Denjoy integration and prove the equivalence of conditions 1 and 2.

Theorem 10. *Let $F : [a, b] \to \mathbb{R}$. The following are equivalent.*

1. *$F \in ACG_*$.*

2. *There is a narrow Denjoy integrable function g such that $\int_a^x g(x)dx = F(x) + F(a)$ for $x \in [a, b]$.*

3. *F is a.e. differentiable and its derivative f is narrow (respectively wide) Denjoy integrable and $\int_a^x f(x)dx = F(x) + F(a)$ for $x \in [a, b]$.*

In Section 1.4 we develop the theory of Denjoy integration. In Section 1.4 we show that the set of absolutely continuous functions is Π_3-complete. These are exactly the indefinite Lebesgue integrals. In Section 1.4 we show that the functions of ACG_* of rank at most α is $\Sigma_{2\alpha}$, and in Section 1.4 we use the main theorem from Section 1.2 to prove this set is $\Sigma_{2\alpha}$-complete.

Preliminaries

Now we define Denjoy integration. Though we rely heavily on Saks [30] and Gordon [6], the level-by-level analysis in this section is not duplicated in either of those resources, though it must be known. For another approach to the fine analysis of this hierarchy, see [36]. We begin with the constructive definition of Denjoy integration and the transfinite process which carries out the integration. Some lemmas are needed in order to facilitate the proof of parts 1 and 2 of Theorem 10, showing that the set of indefinite Denjoy integrals coincides with ACG_*. The relationship is established via a transfinite process which can be applied to a continuous function to determine whether it is in ACG_*, and which exactly parallels the integration process applied to its a.e. derivative. Finally, we discuss the Lusin (N) property and the Banach (S) property, two closely related properties of ACG_* functions, which play a role in the subsequence analysis.

Definition 1.4.3. *Let $f : [a, b] \to \mathbb{R}$ be measurable.*

1. *We say f is Denjoy$_1$-integrable if f is Lebesgue integrable. In this case, for any closed $I \subseteq [a, b]$, the definite Denjoy$_1$-integral of f on I, denoted $D_1 \int_I f$, is equal to the Lebesgue integral of f on I.*

2. *Assume that Denjoy$_\alpha$ integration has been defined. Let $E \subseteq [a, b]$ be the set of points x such that f is not Denjoy$_\alpha$-integrable in any open neighborhood of x. We say f is Denjoy$_{\alpha+1}$-integrable if*

   a) For each open interval (c, d) contiguous to E in $[a, b]$, and each $x \in (c, d)$, $\lim_{y\to d^-} D_\alpha \int_{[x,y]} f$ and $\lim_{y\to c^+} D_\alpha \int_{[y,x]} f$ both exist.

   b) The restriction $f \restriction E$ is Lebesgue integrable on E.

   c)
   $$\sum_{(c,d)\in[a,b]\setminus E} \sup_{x,y\in(c,d)} D_\alpha \int_{[x,y]} f < \infty,$$
   where the sum ranges over intervals (c, d) contiguous to E in $[a, b]$, including intervals of the form $[a, d)$ and $(c, b]$ when $a, b \notin E$.

   In this case, for any closed $I \subseteq [a, b]$, define
   $$D_{\alpha+1} \int_I f = \int_{E\cap I} f + \sum_{(c,d)\in I\setminus E} \lim_{\substack{x\to c^+ \\ y\to d^-}} D_\alpha \int_{[x,y]} f.$$

3. *For λ a limit, f is Denjoy$_\lambda$ integrable if f is Denjoy$_\alpha$ integrable for some $\alpha < \lambda$. In this case, we define $D_\lambda \int_I f = D_\alpha \int_I f$ for such an α.*

We say that f is Denjoy integrable if f is Denjoy$_\alpha$-integrable for some α.

One may verify that Denjoy$_\alpha$-integration satisfies the following property.

Proposition 1.4.1. *If f is Denjoy$_\alpha$-integrable on (a, b) and on (c, d) and the union of these is an interval, then f is Denjoy$_\alpha$-integrable on $(a, b) \cup (c, d)$.*

One can see already suggested in this definition a transfinite process which could be applied to a function in an attempt to Denjoy-integrate it. At each step of the process, one defines a closed set P^α and also adds to a growing set of facts of the form $F(y) - F(x) = z$ where x and y are in the same connected component of $[0,1] \setminus P^\alpha$. By repeating the steps below transfinitely many times one either integrates the function f (in which case $P^\alpha = \emptyset$ for some α) or determines that f cannot be integrated ($P^\alpha = P^{\alpha+1} \neq \emptyset$). First we need some definitions.

Definition 1.4.4. *Let $f : [a,b] \to \mathbb{R}$ be measurable, and let E be a closed set. We say x is a* point of non-summability *of f on E if there is no neighborhood of x on which $f \restriction E$ is Lebesgue integrable.*

Definition 1.4.5. *Let $E \subseteq [a,b]$ be a closed set and let $F(x) - F(y)$ be defined for each $x, y \in [c,d]$ whenever (c,d) is contiguous to E in $[a,b]$. We say x is a* point of divergence *of F on E if for all neighborhoods I of x, $\sum_{(c,d)\in[I]\setminus E} \omega(F,(c,d)) = \infty$.*

To integrate a given $f : [a,b] \to \mathbb{R}$, define a sequence P^α of closed sets and a set of differences $F(y) - F(x)$ recursively by iterating the following steps, stopping when an α is reached such that $P^\alpha = P^\alpha + 1$. This must happen at a countable stage α because the P^α are closed and decreasing. Initialize with $P^0 = [0,1]$. Then repeat:

- At a successor stage $\alpha + 1$:
  1. If (a,b) is a connected component of $[0,1]\setminus P^\alpha$ and $y \in (a,b)$, define $F(y)-F(a) = \lim_{x\to a} F(y) - F(x)$, and similarly for $F(b) - F(y)$ and $F(b) - F(a)$. If any of these limits do not exist, f is not Denjoy integrable.
  2. $P^{\alpha+1} = \{x \in P^\alpha : x \text{ is a point of non-summability of } f \text{ or point of divergence of } F \text{ on } P^\alpha\}$
  3. For each $x < y$ in a connected component of $[0,1] \setminus P^{\alpha+1}$, define
  $$F(y) - F(x) = \int_{(x,y)\cap P^\alpha} f + \sum_{(c,d)\in(x,y)\setminus P^\alpha} F(d) - F(c).$$
- At a limit stage λ: $P^\lambda = \cap_{\alpha<\lambda} P^\alpha$.

Definition 1.4.6. *For any measurable $f : [a,b] \to \mathbb{R}$, let the sequence of sets P^α be defined as in the preceding paragraph. In case of ambiguity about the function, the notation P^α_f is used.*

Proposition 1.4.2. *Let $f : [a,b] \to \mathbb{R}$ be measurable. Let the sets P^α and the differences $F(y) - F(x)$ be defined as above. Then P^α is the set of points x for which f is not Denjoy$_\alpha$-integrable in any neighborhood of x. In particular, f is Denjoy$_\alpha$-integrable if and only if $P^\alpha = \emptyset$. Whenever $F(y) - F(x)$ is defined during stage α we have $F(y) - F(x) = D_\alpha \int_x^y f$, and whenever f is Denjoy$_\alpha$ integrable on $[x,y]$ then $F(y) - F(x)$ is defined by the end of stage α.*

Proof. We proceed by induction on α. If f is Denjoy$_1$-integrable, then f is Lebesgue integrable, so $P^1 = \emptyset$. If $P^1 = \emptyset$, then f is Lebesgue integrable in a neighborhood of every point of $[a, b]$. By compactness, f is Lebesgue integrable on $[a, b]$. Therefore, f is Denjoy$_1$-integrable.

Assume the proposition is known up to α, and let f be given. Let E be the set of points $z \in [a, b]$ for which f is not Denjoy$_\alpha$-integrable in any neighborhood of z. If $x, y \in (c, d)$ where (c, d) is contiguous to E in $[a, b]$, then (by compactness and Proposition 1.4.1) f is Denjoy$_\alpha$-integrable on $[x, y]$, so $P^\alpha \cap [x, y] = \emptyset$, by the inductive hypothesis. Therefore, $P^\alpha \subseteq E$. On the other hand, if I is a closed interval containing a point $x \in E$ in its interior, then f is not Denjoy$_\alpha$-integrable on I, so by the inductive hypothesis, $P^\alpha \cap I \neq \emptyset$. By taking the intersection of the closed P^α with smaller and smaller I, by compactness we see that $x \in P^\alpha$. Therefore, in fact $P^\alpha = E$.

Now, for any $x \in E$, f is Denjoy$_{\alpha+1}$-integrable in a neighborhood of x if and only if E satisfies conditions (2a,b,c) in Definition 1.4.3 in a neighborhood U of x. The conditions are satisfied on a neighborhood U of x if and only if $P^{\alpha+1} \cap U = \emptyset$.

Therefore, $P^{\alpha+1}$ is exactly the set of points x for which f is not Denjoy$_{\alpha+1}$-integrable in any neighborhood of x. If $[x, y]$ is a closed set for which $P^{\alpha+1} \cap [x, y] = \emptyset$, then by compactness, f is Denjoy$_{\alpha+1}$-integrable on $[x, y]$.

By the inductive hypothesis, $D_\alpha \int_x^y f$ exists whenever $[x, y] \cap P^\alpha = \emptyset$ and is equal to $F(y) - F(x)$. therefore, for (c, d) contiguous to P^α, $\lim_{\substack{x \to c^+ \\ y \to d^-}} F(y) - F(x)$ exists if and only if $\lim_{\substack{x \to c^+ \\ y \to d^-}} D_\alpha \int_x^y f$ exists, and when this happens, the newly defined $F(d) - F(c)$ is equal to $D_{\alpha+1} \int_c^d f$. Assuming these limits exist for all (c, d) contiguous to P^α, $D_{\alpha+1} \int_x^y f$ exists whenever $[x, y] \cap P^{\alpha+1} = \emptyset$, and is equal to $\int_{[x,y] \cap P^\alpha} f + \sum_{(c,d) \in [x,y] \setminus P^\alpha} F(d) - F(c)$; and this is exactly the condition in which $F(y) - F(x)$ is defined at stage $\alpha + 1$ to the appropriate value.

This completes the successor case.

By compactness, $P^\lambda = \emptyset$ for λ a limit if and only if $P^\alpha = \emptyset$ for some $\alpha < \lambda$. By the inductive hypothesis, this is true if and only if f is Denjoy$_\alpha$-integrable for some $\alpha < \lambda$, which is exactly the statement that f is Denjoy$_\lambda$-integrable.

□

Now we prove the equivalences 1 and 2 from Theorem 10. We will need to build up some facts about ACG_*.

Theorem 11 (Theorem 6.10 of [6]). *Let $F : [a, b] \to \mathbb{R}$, let $E \subseteq [a, b]$ be closed, and suppose that $F \upharpoonright E$ is continuous on E. Then F is ACG_* on E if and only if for every nonempty perfect subset $E' \subseteq E$, there is an open interval I such that $E' \cap I \neq \emptyset$ and F is AC_* on $E' \cap I$.*

Definition 1.4.7. *Let $F : [a, b] \to \mathbb{R}$, and $E \subseteq [a, b]$ a closed set. Then let F_E denote the function satisfying*

1. *$F_E(x) = F(x)$ for $x \in E$, and*

2. *for (c,d) contiguous to E, let $F_E \restriction (c,d)$ satisfy $F_E(c) = F(c), F_E(d) = F(d)$, $F_E(\frac{c+d}{2}) = \omega(F,[c,d]) + \min(F(c),F(d))$, and F_E linear between these three landmarks.*

Note that $\omega(F_E,[c,d]) = \omega(F,[c,d])$ for (c,d) contiguous to E.

Proposition 1.4.3. *Let $F : [a,b] \to \mathbb{R}$ be bounded, and let $E \subseteq [a,b]$ a closed set. Then F is AC_* on E if and only if F_E is absolutely continuous on $[a,b]$.*

Proof. If F_E is AC on $[a,b]$, then because $F \restriction E = F_E \restriction E$, the fact that F_E is AC on $[a,b]$ directly implies that F is AC_* on E.

On the other hand, suppose F is AC_* on E. It follows that

$$\sum_{(c,d)\in[x,y]\setminus E} \omega(F,[c,d]) < \infty.$$

Let ε be given. Let δ^* be small enough that

$$\sum_{\substack{(c,d)\in[x,y]\setminus E \\ \text{s.t. } d-c<\delta^*}} \omega(F,[c,d]) < \varepsilon.$$

and also small enough to witness that F is AC_* on E for ε. Let

$$C = \bigcup_{\substack{(c,d)\in[a,b]\setminus E \\ \text{s.t. } d-c\geq\delta^*}} (c,d).$$

Because C is a finite union of open intervals and F_E is piecewise linear on C, we can let $\delta < \delta^*$ be also small enough that $\delta|F'(x)| < \varepsilon$ for all $x \in C$. We claim that δ witnesses that F_E is AC on $[a,b]$ for 4ε.

Let $(a_i,b_i)_{i<n}$ be a sequence of finitely many disjoint intervals in $[a,b]$ with $\sum_i |b_i - a_i| < \delta$. If for any i, $(a_i,b_i) \cap E \neq \emptyset$, let c be the least element of $E \cap [a_i,b_i]$ and d the greatest. These numbers may coincide with each other or with either endpoint. We may as well replace (a_i,b_i) by the three intervals $(a_i,c),(c,d)$ and (d,b_i), since doing so can only increase $\sum_i |F_{E*}(b_i) - F_{E*}(a_i)|$. Therefore, without loss of generality, let us assume that for each (a_i,b_i), either $a_i,b_i \in E$, or (a_i,b_i) is disjoint from E. (Of course, both could be true). If (a_i,b_i) is disjoint from E, then either $(a_i,b_i) \subseteq C$, or $(a_i,b_i) \cap C = \emptyset$. Using these categories, we calculate

$$\begin{aligned}\sum_i |F_E(b_i)-F_E(a_i)| &\le \sum_{i:a_i,b_i\in E} |F_E(b_i)-F_E(a_i)| + \sum_{\substack{i:a_i\notin E \text{ or } b_i\notin E\\ \text{and } (a_i,b_i)\cap C=\emptyset}} |F_E(b_i)-F_E(a_i)| \\ &\quad + \sum_{\substack{i:a_i\notin E \text{ or } b_i\notin E\\ \text{and } (a_i,b_i)\subseteq C}} |F_E(b_i)-F_E(a_i)| \\ &\le \sum_{i:a_i,b_i\in E} |F(b_i)-F(a_i)| + \sum_{\substack{(c,d)\in[a,b]\setminus E\\ \text{s.t. } d-c<\delta}} 2\omega(F_E,[c,d]) \\ &\quad + \sum_{\substack{i:a_i\notin E \text{ or } b_i\notin E\\ \text{and } (a_i,b_i)\subseteq C}} |b_i-a_i| \max_{x\in C}|F'(x)| \\ &\le \varepsilon + \sum_{\substack{(c,d)\in[a,b]\setminus E\\ \text{s.t. } d-c<\delta}} 2\omega(F,[c,d]) + \left(\sum_{\substack{i:a_i\notin E \text{ or } b_i\notin E\\ \text{and } (a_i,b_i)\subseteq C}} |b_i-a_i| \right) \max_{x\in C}|F'(x)| \\ &\le \varepsilon + 2\varepsilon + \delta \max_{x\in C}|F'(x)| \le 4\varepsilon.\end{aligned}$$

Therefore, F_E is AC on $[a,b]$. □

Proposition 1.4.4. *Let $F:[a,b]\to\mathbb{R}$ be ACG_* and let $E\subseteq[a,b]$ be closed. Let $E'=\{x\in E : F$ is not AC_* on E in any neighborhood of $x\}$. Then $E=\overline{E\setminus E'}$ where the overline denotes the topological closure.*

Proof. Clearly $E\supseteq\overline{E\setminus E'}$. In the other direction, let $x\in E$ and suppose for contradiction that $x\notin\overline{E\setminus E'}$. Then there is an open neighborhood $U\ni x$ such that $U\cap\overline{E\setminus E'}=\emptyset$. Also, $x\in E'$, since $x\in E$ and $x\notin E\setminus E'$.

Because F is not AC_* on E in any neighborhood of x, x is not isolated in E. Let $[z,w]\subseteq U$ such that $x\in(z,w)$.

Then $E\cap[z,w]$ is closed, nonempty because $x\in[z,w]$.

If z or w in E, assume without loss of generality that they are not isolated in $E\cap[z,w]$, because if they are isolated, just chose a smaller interval $[z,w]$ so that $z,w\notin E$.

Now there are two cases. If $E\cap[z,w]$ has an isolated point y, then since $y\neq z$, $y\neq w$, in fact y is isolated in E, so $y\in E\setminus E'$, a contradiction. In case 2, $E\cap[z,w]$ is perfect. But then by Theorem 11, there is an open I such that F is AC_* on the nonempty $E\cap[z,w]\cap I$. Again, this is a contradiction. □

Proposition 1.4.5. *Suppose f is Denjoy$_{\alpha+1}$-integrable on an interval (a,b). Then $F(x)=D_{\alpha+1}\int_a^x f$ is AC_* on $P^\alpha\cap(a,b)$.*

Proof. Let ε be given. Because $f\restriction P^\alpha\cap(a,b)$ is Lebesgue integrable, $G(x)=\int_{P^\alpha\cap(a,x)} f$ is AC. Let δ^* witness the absolute continuity of G for $\frac{\varepsilon}{2}$.

Because $\sum_{(c,d)\in(a,b)\setminus P^\alpha}\omega(F,[c,d])<\infty$, let $\delta<\delta^*$ be such that

$$\sum_{\substack{(c,d)\in(a,b)\setminus P^\alpha\\ \text{s.t. } d-c<\delta}}\omega(F,[c,d])<\frac{\varepsilon}{2}.$$

Let $(x_i,y_i)i<k$ be disjoint with $\sum_i y_i-x_i<\delta$, and $x_i,y_i\in P^\alpha$.
Then

$$\begin{aligned}\sum_i\omega(F,[x_i,y_i]) &= \sum_i\sup_{x,y\in[x_i,y_i]}|F(y)-F(x)|\\ &\le\sum_i\sup_{x,y\in[x_i,y_i]}\left|\int_{P^\alpha\cap(x,y)}f\right|+\sum_{(c,d)\in(x,y)\setminus P^\alpha}|F(d)-F(c)|\\ &\le\sum_i\sup_{x,y\in[x_i,y_i]}|G(y)-G(x)|+\sum_{\substack{(c,d)\in(a,b)\setminus P^\alpha\\ \text{s.t. } d-c<\delta}}\omega(F,[c,d])\\ &\le\frac{\varepsilon}{2}+\frac{\varepsilon}{2}.\end{aligned}$$

□

Now we are ready to prove the equivalence of the first two conditions in Theorem 10.

Proof of Theorem 10, parts 1 and 2. Suppose that $F:[a,b]\to\mathbb{R}$ is ACG_*. We will construct a Denjoy integrable $g:[a,b]\to\mathbb{R}$ such that $D\int_{[x,y]}g=F(y)-F(x)$ for all $x,y\in[a,b]$.

Define a sequence of sets $\widetilde{P}^\alpha$ by induction as follows. Let $\widetilde{P}^0=[0,1]$, $\widetilde{P}^{\alpha+1}=\{x\in\widetilde{P}^\alpha:$ F is not AC_* on $\widetilde{P}^\alpha$ in any neighborhood of $x\}$, and $\widetilde{P}^\lambda=\cap_{\alpha<\lambda}\widetilde{P}^\alpha$ for λ a limit. If $\widetilde{P}^\alpha$ is non-empty, then $\widetilde{P}^{\alpha+1}$ is strictly contained in it, because if $\widetilde{P}^\alpha$ has an isolated point then this point is not present in $\widetilde{P}^{\alpha+1}$, and if $\widetilde{P}^\alpha$ is perfect, then Theorem 11 guarantees the strict containment of $\widetilde{P}^{\alpha+1}$. Therefore, there is an α for which $\widetilde{P}^\alpha=\emptyset$.

For each α, let $F_\alpha=F_{\widetilde{P}^\alpha}$. Then for every $z\in[x,y]$ disjoint from $\widetilde{P}^{\alpha+1}$, F_α is absolutely continuous on a neighborhood of z by Proposition 1.4.3. By the compactness of $[x,y]$, F_α is absolutely continuous on $[x,y]$.

For each α, let $g_\alpha:[a,b]\setminus\widetilde{P}^{\alpha+1}\to\mathbb{R}$ be the a.e. unique Lebesgue integrable function such that for all $[x,y]\subseteq[a,b]\setminus\widetilde{P}^{\alpha+1}$, $\int_{[x,y]}g_\alpha=F_\alpha(y)-F_\alpha(x)$. Let $g(x)=g_\alpha(x)$ for $x\in\widetilde{P}^\alpha\setminus\widetilde{P}^{\alpha+1}$. We claim that the sequence of P^α obtained in the process of integrating g is exactly the sequence $\widetilde{P}^\alpha$ used in its construction, and thus g is Denjoy$_\alpha$ integrable for exactly the α for which $P^\alpha=\emptyset$, and that furthermore $D_\alpha\int_{[x,y]}g=F(y)-F(x)$ for all $[x,y]\subseteq[a,b]$. The claim is proved by induction on α, using the following three properties as the inductive claim for α:

1. g is Denjoy$_\alpha$-integrable on $[x,y]$ when $\widetilde{P}^\alpha\cap[x,y]=\emptyset$. (This implies $P^\alpha\subseteq\widetilde{P}^\alpha$.)
2. $D_\alpha\int_x^y g=F(y)-F(x)$ for $[x,y]\cap\widetilde{P}^\alpha=\emptyset$.

3. For all $\beta < \alpha$, $\widetilde{P}^\beta = P^\beta$ where P^β are the sets obtained in the attempt to integrate g according to Definition 1.4.6.

In the base case $\alpha = 1$, and the inductive claim holds trivially. Now we assume it is true for α and consider $\alpha + 1$. Addressing the third property first, suppose $z \in \widetilde{P}^\alpha \setminus \widetilde{P}^{\alpha+1}$. If α is a limit, then $z \in \widetilde{P}^\beta = P^\beta$ for all $\beta < \alpha$, so $z \in P^\alpha$. If $\alpha = \beta + 1$, then $\widetilde{P}^\beta = P^\beta$. Let us show that $z \in P^\alpha$. Letting (x, y) be an interval on which g is D_α-integrable, we show $z \notin (x, y)$. For any open interval $(x, y) \ni z$, F is not AC^* on $P^\beta \cap (x, y)$. Without loss of generality, assume $[x, y] \cap \widetilde{(P)}^{\alpha+1} = \emptyset$. Then by Proposition 1.4.5, because g is Denjoy$_\alpha$-integrable on (x, y), the function $G(u) = D_\alpha \int_x^u g$ is AC_* on $P^\beta \cap [x, y]$. For $u, w \in (x, y)$, we have

$$D_\alpha \int_u^w g = D_\alpha \int_u^{u^*} g + D_\alpha \int_{u^*}^{w^*} g + D_\alpha \int_{w^*}^{w} g$$

where $u^* = \min[u, w] \cap \widetilde{P}^\alpha$ and $w^* = \max[u, w] \cap \widetilde{P}^\alpha$.

By the inductive hypothesis, $D_\alpha \int_x^y g = F(y) - F(x)$ whenever $[x, y] \cap \widetilde{P}^\alpha = \emptyset$. Together with the continuity of F, that implies $D_\alpha \int_u^{u^*} g = F(u^*) - F(u)$ and $D_\alpha \int_{w^*}^{w} g = F(w) - F(w^*)$.

Consider now the middle term

$$\begin{aligned}
D_\alpha \int_{u^*}^{w^*} g &= \int_{P^\beta \cap [u^*, w^*]} g + \sum_{(c,d) \in [u^*, w^*] \setminus P^\beta} F(d) - F(c) \\
&= \left(\int_{\widetilde{P}^\alpha \cap [u^*, w^*]} g + \sum_{(c',d') \in [u^*, w^*] \setminus \widetilde{P}^\alpha} \int_{P^\beta \cap [c', d']} g \right) + \left(\sum_{(c',d') \in [u^*, w^*] \setminus \widetilde{P}^\alpha} \sum_{(c,d) \in (c',d') \setminus P^\beta} F(d) - F(c) \right) \\
&= \int_{\widetilde{P}^\alpha \cap [u^*, w^*]} g_\alpha + \sum_{(c',d') \in [u^*, w^*] \setminus \widetilde{P}^\alpha} D_\alpha \int_{c'}^{d'} g \\
&= \int_{\widetilde{P}^\alpha \cap [u^*, w^*]} g_\alpha + \sum_{(c',d') \in [u^*, w^*] \setminus \widetilde{P}^\alpha} F(d') - F(c') \\
&= \int_{\widetilde{P}^\alpha \cap [u^*, w^*]} g_\alpha + \sum_{(c',d') \in [u^*, w^*] \setminus \widetilde{P}^\alpha} F_\alpha(d') - F_\alpha(c') \\
&= \int_{u^*}^{w*} g_\alpha = F(w^*) - F(u^*).
\end{aligned}$$

Therefore, $F(w) - F(u) = G(w) - G(u)$ for these arbitrary u, w, so F and G are equal up to a constant on $[x, y]$, so F is AC_* on $[x, y] \cap P^\beta$ because G was. Therefore $\widetilde{P}^\alpha \cap (x, y) = \emptyset$, so z is not in (x, y). Since this was done for an arbitrary (x, y), we conclude that $z \in P^\alpha$.

Now let x, y be such that $[x, y] \cap \widetilde{P}^{\alpha+1} = \emptyset$. We just saw that $P^\alpha = \widetilde{P}^\alpha$, so F is AC_* on $P^\alpha \cap [x, y]$. By the inductive hypothesis applied to each $[x, y]$ disjoint from P^α, $D_\alpha \int_{[x,y]} g = F(y) - F(x)$, so by the continuity of F, $\lim_{\substack{x \to c \\ y \to d}} D_\alpha \int_{[x,y]} g = F(d) - F(c)$, where

(c, d) is an interval contiguous to P^α which contains $[x, y]$. Because g_α is Lebesgue integrable and coincides with g on $P^\alpha \setminus \widetilde{P}^{\alpha+1}$, we see that $g \restriction P^\alpha \cap [x, y]$ is Lebesgue integrable. Because F is AC_* on $P^\alpha \cap [x, y]$, criterion (2c) of Definition 1.4.3 holds. Therefore, g is Denjoy$_{\alpha+1}$-integrable on $[x, y]$. Letting c^* be the least element of P^α in $[x, y]$ and d^* the greatest, we have

$$\begin{aligned} D_{\alpha+1} \int_{[x,y]} g &= \int_{[x,y] \cap P^\alpha} g + \sum_{(c,d) \in [x,y] \setminus P^\alpha} F(d) - F(c) \\ &= \int_{[c^*,d^*] \cap P^\alpha} g_\alpha + F(y) - F(d^*) + F(c^*) - F(x) + \sum_{(c,d) \in [c^*,d^*] \setminus P^\alpha} \int_{[c,d]} g_\alpha \\ &= F(y) - F(x) + \left(\int_{[c^*,d^*]} g_\alpha \right) - (F(d^*) - F(c^*)) \\ &= F(y) - F(x). \end{aligned}$$

Now for the limit case, assume the claim holds for all $\alpha < \lambda$ where λ is a limit. Then for all $\beta < \lambda$, $\beta < \alpha$ for some $\alpha < \lambda$, so $\widetilde{P}^\beta = P^\beta$. Then $[x, y] \cap \widetilde{P}^\lambda = \emptyset$ implies $[x, y] \cap \widetilde{P}^\alpha = \emptyset$ for some $\alpha < \lambda$, which by the inductive hypothesis implies that g is Denjoy$_\alpha$-integrable with the proper values on $[x, y]$, so g is Denjoy$_\lambda$-integrable with the proper values on $[x, y]$.

We have shown that if $F : [a, b] \to \mathbb{R}$ is ACG_* then there is a Denjoy integrable g such that $D \int_{[x,y]} g = F(y) - F(x)$ for all $[x, y] \subseteq [a, b]$. In the other direction, let us assume the existence of a Denjoy integrable g, and show that $F(x) = D \int_{[0,x]} g$ is ACG_*. Let P^α be the sets from Definition 1.4.6 which result when integrating g. Then $[a, b]$ is the union of the following countable collection of sets:

$$K_{\alpha,p,q} = [p, q] \cap P^\alpha \text{ for } p, q \in \mathbb{Q} \text{ and } [p, q] \text{ disjoint from } P^{\alpha+1}.$$

We verify that F is AC_* on each of these sets. Because $P^{\alpha+1} = \emptyset$ on $[p, q]$, $g \restriction [p, q]$ is Denjoy$_{\alpha+1}$-integrable there. Then $\lim_{\substack{x \to c^+ \\ y \to d^-}} D_\alpha \int_{[x,y]} g = \lim_{\substack{x \to c^+ \\ y \to d^-}} F(y) - F(x) = F(d) - F(c)$ for each (c, d) contiguous to P^α in $[p, q]$, so the condition $\sum_{(c,d) \in [p,q] \setminus P^\alpha} \omega(F, [c, d]) < \infty$ holds.

Let ε be given. Because $g \restriction P^\alpha \cap [p, q]$ is Lebesgue integrable, $G(x) = \int_{P^\alpha \cap [p,x]} g$ is AC. Let δ^* witness the absolute continuity of G for $\frac{\varepsilon}{2}$. Let $\delta < \delta^*$ be small enough that $\sum_{\substack{(c,d) \in [p,q] \setminus P^\alpha \\ d-c<\delta}} \omega(F, [c, d]) < \frac{\varepsilon}{2}$ Now suppose we are given a sequence $(a_i, b_i)_{i<k}$ of disjoint

intervals with $a_i, b_i \in P^\alpha \cap [p,q]$ and $b_i - a_i < \delta$. Then

$$\begin{aligned}\sum_i \omega(F,[a_i,b_i]) &= \sum_o \sup_{x,y\in[a_i,b_i]} |F(y)-F(x)| \\ &\le \sum_i \sup_{x,y\in[a_i,b_i]} \left(|\int_{[x,y]\cap P^\alpha} g| + \sum_{(c,d)\in[x,y]\setminus P^\alpha} |F(d)-F(c)|\right) \\ &\le \sum_i \sup_{x,y\in[a_i,b_i]} |G(y)-G(x)| + \sum_{(c,d)\in[a_i,b_i]\setminus P^\alpha} \omega(F,[c,d]) \\ &\le \frac{\varepsilon}{2} + \sum_{\substack{(c,d)\in[p,q]\setminus P^\alpha \\ \text{s.t. } d-c<\delta}} \omega(F,[c,d]) < \varepsilon.\end{aligned}$$

□

Though we will not prove the third part of the equivalence here, we remark that because of the work required to develop the constructive definition of Denjoy integration, a variant of the third part of the equivalence is often taken as the definition of Denjoy integration: a measurable function $f : [a,b] \to \mathbb{R}$ is narrow Denjoy integrable if there is a function $F \in ACG_*$ which is a.e. differentiable and whose derivative is f.

It was clear from the definition of Denjoy integration for a function g that the resulting indefinite integral F is the same for any f in the same a.e. equivalence class of g. Conversely, by the proof of Theorem 10 one sees that if F is ACG_* then the g which integrates to it is a.e. unique. Therefore, a.e. equivalence classes of Denjoy integrable functions are in one-to-one correspondence with their indefinite integrals in ACG_*.

We mentioned earlier that although every function of ACG_* is a.e. differentiable, there are a.e. differentiable functions which are not ACG_*. Perhaps the most prominent example of this sort of function is Cantor's function, also known as the Devil's Staircase.

Definition 1.4.8. Cantor's function $f_C : [0,1] \to [0,1]$ *is defined as follows. Let* C *be the standard Cantor set in* $[0,1]$. *For each* n, *there are* 2^{n-1}*-many intervals* I *of length* 3^{-n} *contiguous to* C. *Let* f_C *be constant on each such interval, taking the values* $\frac{i}{2^n}$ *for* i *odd with* $0 < i < 2^n$ *in the same order in which the intervals appear. For example,* $f_C \restriction (\frac{1}{3},\frac{2}{3}) \equiv \frac{1}{2}$, $f_C \restriction (\frac{1}{9},\frac{2}{9}) \equiv \frac{1}{4}$, *and* $f_C \restriction (\frac{7}{9},\frac{8}{9}) \equiv \frac{3}{4}$. *Define* f_C *on* C *so that* f *is continuous on* $[0,1]$.

This function is increasing and a.e. differentiable and its derivative is zero almost everywhere. To see that f_C is not ACG_*, note that if g were a function which could be integrated to obtain f_C, then g would have to be zero almost everywhere since f_C is constant almost everywhere, but then $\int_0^x g = 0$, not $f_C(x)$, a contradiction.

Cantor's function has the interesting property that $f_C(C)$ is all of $[0,1]$ except the dyadic rationals. Therefore, f_C maps a measure zero set to a set of measure 1. This is, in general, a way to tell when a function is not ACG_*. Consider the following well-known characterization.

Proposition 1.4.6. *A function $F : [a, b] \to \mathbb{R}$ is absolutely continuous if and only if it is continuous, of bounded variation, and $\mu(F(A)) = 0$ whenever $\mu(A) = 0$.*

The last property in that list is called the Lusin (N) property. And we have

Proposition 1.4.7 ([6], Theorem 6.12)**.** *If $F \in ACG_*$ then F satisfies Lusin's (N) property.*

We will have need for a slight variation of this, however. Under the assumption that F is continuous, the last property in the proposition below is equivalent to Banach's (S) property, that for every ε there is a δ such that whenever $\lambda(A) < \delta$, $\lambda(F(A)) < \varepsilon$. Since we deal only with continuous functions, we will refer to the property below as Banach's (S) property as well.

Proposition 1.4.8. *A function F is absolutely continuous if and only if it is continuous, it is of bounded variation, and for every ε there exists a δ such that whenever $(a_i, b_i)_{i<k}$ is a finite sequence of disjoint intervals for which the $(F(a_i), F(b_i))$ (indices reversed if necessary) are also disjoint, then $\sum_i |F(a_i) - F(b_i)| < \varepsilon$.*

Proof. It is clear from the definition of absolute continuity that it implies the property (S). On the other hand, our formulation of (S) implies continuity, and a continuous function satisfying Banach's (S) also satisfies Lusin's (N), by [30, Theorem 7.4]. Therefore, (S) and bounded variation imply absolute continuity. □

Also, [30, Theorem 8.8, page 233], [30, Theorem 6.3, page 279], and [30, Theorem 7.3, page 284], combined with the preceding, give us

Proposition 1.4.9. *If F is ACG_*, then F fulfills condition (S).*

In section 1.4 we show that the set $\{e : F_e \text{ is } AC\}$ is Π_3-complete. In Section 1.4 we show that for $\alpha > 1$, $\{e : F_e(x) = D_\alpha \int_0^x f \text{ for some } f\}$ is $\Sigma_{2\alpha}$. This also implies that ACG_* is Π^1_1. In Section 1.4, we provide a reduction from well-founded trees to ACG_* with the property that if $T \mapsto F$ then $|T|_{ls} = |F|_D$, where $|F|_D$ is the least α such that $F(x) = D_\alpha \int_0^x f$ for some f. This shows that for $\alpha > 1$ $\{e : F_e(x) = D_\alpha \int_0^x f \text{ for some } f\}$ is $\Sigma_{2\alpha}$-complete, and that ACG_* is Π^1_1-complete.

To complete the analogy with previous sections, we may define

Definition 1.4.9. *Let $F : [a, b] \to \mathbb{R}$. If $F \in ACG_*$, let the Denjoy rank of F, denoted $|F|_D$, be the least α such that $F(x) = D_\alpha \int_a^x f$ for some f.*

In this notation, the results of the next three sections may be summarized as

Theorem 12. *Let F_e denote the eth computable function in $C[0, 1]$. Then*

1. *The set $\{e : |F_e|_D = 1\}$ is Π_3-complete*

2. *For each constructive $\alpha > 1$, the set $\{e : |F_e|_D \leq \alpha\}$ is $\Sigma_{2\alpha}$-complete.*

Being a Lebesgue integral is Π_3-complete

In this section we prove that the image of Lebesgue integration is Π_3-complete. By F_e we mean the eth code for a continuous function in the sense of Section 1.1.

Proposition 1.4.10. *The set* $\{e : F_e$ *is absolutely continuous*$\}$ *is* Π_3*-complete.*

Proof. Under the Π_2 assumption of the continuity of F, the remainder of the definition of absolute continuity is Π_3, because in the continuous setting, rational approximations suffice for everything. That is, one may check that

$$\begin{aligned} F \in AC \iff & F \text{ is continuous and} \\ & \forall \varepsilon \exists \delta \forall (a_i, b_i)_{i<k} \\ & \text{if the } (a_i, b_i) \text{ are disjoint and } \sum_i b_i - a_i < \delta, \text{ then } \sum_i |F(b_i) - F(a_i)| \leq \varepsilon. \end{aligned}$$

where all quantifications are over the rationals or intervals with rational endpoints; the characterization on the right is Π_3.

Considering now the alternate characterization of absolute continuity given in Proposition 1.4.7, one may check that the statements "F_e is continuous" is Π_2 and "F_e is of bounded variation" is Σ_2 in the set of continuous indices. So any proof of the Π_3-completeness of being absolutely continuous must crucially use the Lusin (N) property. Our strategy is to approximate a version of the Cantor function which will converge to a Cantor-like function only if the Π_3 statement fails, and be absolutely continuous otherwise.

Any canonical representation of a Π_3 statement can be effectively re-written as $\forall n[W_{g(n)}$ is finite] for some computable total g. We now define a function $F : [0, 1] \to [0, 1]$, uniformly in g, such that F is absolutely continuous if and only if $\forall n[W_{g(n)}$ is finite] holds.

Effective in g, we define a computable sequence of functions F_s which converge effectively and uniformly to the desired computable function F. Let $F_0(x) = x$. Each F_s will be piecewise linear, containing some pieces of slope zero separated by pieces of positive slope. Wherever F_s is piecewise constant, it is equal to the limiting function F.

For each n let $I_n = [\frac{1}{n+2}, \frac{1}{n+1}]$. This is the interval in which $W_{g(n)}$'s finiteness or lack thereof will be expressed.. At stage $s + 1$, see for which $n < s$ is there a new element enumerated into $W_{g(n)}$. Let $F_{s+1} \upharpoonright [\frac{1}{n+2}, \frac{1}{n+1}] = F_s \upharpoonright [\frac{1}{n+2}, \frac{1}{n+1}]$ for all $n \geq s$ and for all $n < s$ such that no new element of $W_{g(n)}$ has been enumerated at stage s. For those n for which a new element is enumerated into $W_{g(n)}$, define $F_{s+1} \upharpoonright I_n$ as follows. For each maximal interval $I \subseteq I_n$ on which F_s is constant, let $F_{s+1} \equiv F_s$ on I. For each maximal interval I on which $F_s$is linear with positive slope, define F_{s+1} on I to satisfy:

1. $F_{s+1} = F_s$ at the endpoints
2. F_{s+1} is piecewise linear, continuous, and increasing
3. F_{s+1} has slope zero on $\frac{1}{3}$ of I, and the slope everywhere else is increased by an appropriate factor, which happens to be $\frac{3}{2}$.

4. F_s and F_{s+1} differ by no more than 2^{-s} at any point.

This can be accomplished by letting $F_{s+1} \upharpoonright I$ resemble a sufficiently fine staircase. The effect is that $\frac{1}{3}$ of the measure of I is given to points at which $F'(x) = 0$. Thus if F'_s was nonzero on a measure r subset of I_n, then F'_{s+1} is nonzero on a measure $\frac{2}{3}r$ subset of I_n.

This completes the construction. One may check that F is continuous and of bounded variation.

Now suppose that it holds that $\forall n[W_{g(n)}$ is finite]. Then for each n, there will come a stage s for which $F_s \upharpoonright I_n = F \upharpoonright I_n$, and so the final F is piecewise linear on I_n for all n. And F satisfies the Lusin N property because each I_n satisfies it, and there are only countably many I_n. So if $\mu(A) = 0$ then $\mu(F(A)) = \mu(\cup_n F(A \cap I_n)) \leq \sum_n \mu(F(A \cap I_n)) = 0$. Thus F is absolutely continuous.

On the other had, suppose that $W_{g(n)}$ is infinite for some fixed n. Then letting $Z = \cup_s\{x \in I_n : F'_s(x) = 0\}$, we have $\mu(Z) = \mu(I_n)$, but $F(Z)$ is countable, since for each s, $\{F_s(x) : F'_s(x) = 0\}$ is finite. But F is continuous, so $F(I_n) = I_n$, so $F(I_n \setminus Z)$ has measure $\mu(I_n)$, and F is not absolutely continuous. □

Being a Denjoy integral of rank α is $\Sigma_{2\alpha}$

In this section we prove that for all constructive $\alpha > 1$, $\{e : F_e(x) = D_\alpha \int_0^x f\}$ is $\Sigma_{2\alpha}$. For the purposes of showing the completeness result of the last section, it was necessary to make everything hinge on the Lusin (N) property. However, by Proposition 1.4.7, every $F \in ACG_*$ has the Lusin N property.

Even more, if $F \in ACG_*$ then F satisfies Banach's (S) property, by Proposition 1.4.9.

So this property cannot differentiate between indefinite Denjoy integrals of different ranks. So it happens that the property of bounded variation piled on top of bounded variation is what drives the hierarchy once we are past the three jumps where the Banach (S) property can make a difference. In particular, we claim that

Proposition 1.4.11. *For every recursive $\alpha > 1$, $\{e : F_e$ is a Denjoy integral of rank at most $\alpha\}$ is $\Sigma_{2\alpha}$.*

Proof. It is straightforward to verify that the Banach (S) property is Π_3. Because ACG_* is a proper subset of the continuous functions satisfying Banach (S), we may work inside this class.

Under the assumption that F satisfies the Banach (S) property, we claim that uniformly in e and in all constructive α, the set $\{(p,q) : p,q \in \mathbb{Q}$ and $[p,q] \cap P^\alpha = \emptyset\}$ is $\Sigma_{2\alpha}$, where P^α are defined as follows: $P^0 = [a,b]$, $P^{\alpha+1} = \{x \in P^\alpha : F$ is not AC_* on P^α in any neighborhood of $x\}$, $P^\lambda = \cap_{\alpha<\lambda} P^\alpha$ for λ a limit.

Note that this definition of the P^α hierarchy is the same as that given in the proof of Theorem 10, and so $|F|_D \leq \alpha$ if and only if $P^\alpha = \emptyset$. The latter is true if and only if $\{(p,q) : [p,q] \cap P^\alpha = \emptyset\}$ covers $[a,b]$ (as usual, we allow $[a,q),(p,b]$ in the cover), and if this happens then by compactness there is a finite cover. So for F satisfying the Banach (S) property,

$$\begin{aligned}|F|_D \leq \alpha &\iff P^\alpha = 0\\ &\iff \exists (p_i, q_i)_{i<k} \text{ s.t } [p_i, q_i] \cap P^\alpha = \emptyset \text{ and } \cup_i (p_i, q_i) = [a, b]\end{aligned}$$

Under the assumption that $[p_i, q_i] \cap P^\alpha = \emptyset$ is $\Sigma_{2\alpha}$ uniformly in e and in α, this is the result, because $|F|_D \leq \alpha \iff$ (F satisfies (S)) and ($P^\alpha = \emptyset$ assuming F satisfies (S)).

So now let us see why $[p_i, q_i] \cap P^\alpha = \emptyset$ is a $\Sigma_{2\alpha}$ statement for $\alpha > 0$ under the assumption that F satisfies the Banach (S) property.

When $\alpha = 1$, because F is already assumed to be continuous and have the Banach (S) property, we need only check for bounded variation. Thus

$$\begin{aligned}[p, q] \cap P^1 = \emptyset &\iff \exists (p', q') \supset [p, q] \text{ s.t. } F \text{ has bounded variation on } (p', q')\\ &\iff \exists (p', q') \supset [p, q] \exists N \text{ such that for all } (a_i, b_i)_{i<k} \text{ disjoint in } (p', q'),\\ &\quad \sum_i |F(b_i) - F(a_i)| \leq N.\end{aligned}$$

and the last expression is Σ_2.

Supposing that $\{(p, q) : [p, q] \cap P^\alpha = \emptyset\}$ is $\Sigma_{2\alpha}$, let us show $\{(p, q) : [p, q] \cap P^{\alpha+1} = \emptyset\}$ is $\Sigma_{2\alpha+2}$ under the assumption that F has the Banach (S) property.

We have by Proposition 1.4.3

$$\begin{aligned}[p, q] \cap P^{\alpha+1} = \emptyset &\iff \exists (p', q') \supset [p, q] \text{ s.t. } F \text{ is } AC_* \text{ on } [p', q'] \cap P^\alpha\\ &\iff \exists [p', q'] \supset [p, q] \text{ s.t. } F_\alpha \text{ is } AC \text{ on } [p', q']\end{aligned}$$

where $F_\alpha = F_{P^\alpha}$ Because F has the Banach (S) property, it has the Lusin (N) property. So if $\lambda(A) = 0$ then $\lambda(F(A \cap P^\alpha))$. And $\lambda(F_\alpha(A \cap (c, d))) = 0$ for any (c, d) contiguous to P^α because F_α is linear there. So $\mu(F_\alpha(A))) = 0$. So F_α has the Lusin (N) property, and we may continue with

$$[p, q] \cap P^{\alpha+1} = \emptyset \iff \exists [p', q'] \supset [p, q] \text{ s.t. } F_\alpha \text{ has bounded variation on } [p', q']$$

The last part is $\Sigma_2(F_\alpha)$, so if F_α is uniformly $\emptyset_{(2\alpha)}$-computable, we are done.

Because F was computable, it has an effective modulus of uniform continuity. Thus, given ε, we may effectively find N such that for any interval I of length at most $1/N$, $\omega(F, I) < \varepsilon$. Increase N if necessary so $\omega(F, [a, b])/N < \varepsilon$ as well. We use this division to compute an approximation G to F_α as follows.

For each natural number $i < N$ such that $[\frac{i}{N}, \frac{i+1}{N}] \cap P^alpha = \emptyset$, let $G \restriction [\frac{i}{N}, \frac{i+1}{N}] = F \restriction [\frac{i}{N}, \frac{i+1}{N}]$. For each remaining maximal interval of the form $(\frac{i}{N}, \frac{j}{N})$,

Let

$$w = \max_{i', j' \in [i, j]} \left| F(\frac{i'}{N}) - F(\frac{j'}{N}) \right|$$

where i' and j' range over integers, and let $G(\frac{i}{N}) = F(\frac{i}{N})$, $G(\frac{j}{N}) = F(\frac{j}{N})$, $G(\frac{i+j}{2N}) = w + \min(F(\frac{i}{N}), F(\frac{j}{N}))$, and let G be defined on the rest of $[\frac{i}{N}, \frac{j}{N}]$ as a linear interpolation of these points.

One may verify that G is computable from $\emptyset_{2\alpha}$ which answers questions of the form $[\frac{i}{N}, \frac{i+1}{N}] \cap P^\alpha = \emptyset$? Let us check that $||G - F_\alpha|| < 11\varepsilon$.

First we claim that $||F_\alpha - G|| < 6\varepsilon$ on the intervals $[\frac{i}{N}, \frac{i+1}{N}]$ whose intersection with P^α is nonempty. Since $F_\alpha(x) = G(x)$ for $x \in P^\alpha$, we consider the rest of the points of the interval according to the following three cases:

- Points of any interval (c, d) contiguous to P^α which is fully contained within $[\frac{i}{N}, \frac{i+1}{N}]$
- Points in the interval $(c, \frac{i+1}{N}]$, where $c \in P^\alpha$, but $(c, \frac{i+1}{N}]$ is disjoint from P^α.
- Points in the interval $[\frac{i}{N}, d)$, symmetric to the above.

In the first case, for $z \in (c, d)$, $|F_\alpha(x) - G(x)| = |F_\alpha(x) - F(x)|$. If x is closer to c, then

$$|F_\alpha(x) - F(x)| = (x - c)|F'_\alpha(x)|$$

where $|F'_\alpha(x)| \leq \frac{2\omega(F,[c,d])}{d-c} < \frac{2}{d-c}\varepsilon$. Therefore,

$$|F_\alpha(x) - F(x)| < (x - c)\frac{2}{d - c}\varepsilon < 2\varepsilon$$

since $x - c < d - c$. If x is closer to d, a symmetric argument gives the same conclusion.

In the second case (and the third, by symmetry), for $x \in (c, \frac{i+1}{N}]$, let (c, d) be the full connected component contiguous to P^α which extends the interval under consideration. Let j be an integer for which $d \in [\frac{j}{N}, \frac{j+1}{N}]$. Then by the choice of N, $\omega(F, [c, d]) < \varepsilon(j - i + 1)$, so $F'_\alpha(x) < \frac{2(j-i+1)\varepsilon}{d-c}$. Then regardless of whether x is closer to c or d, we have

$$|F_\alpha(x) - F(x)| < (x - c)\frac{2(j - i + 1)}{d - c}\varepsilon$$

. If $j - i = 1$, then since $x - c < d - c$, we have $|F_\alpha(x) - F(x)| < 4\varepsilon$. Otherwise, $d - c \geq \frac{j-i-1}{N}$, so $|F_\alpha(x) - F(x)| < (x - c)\frac{2(j-i+1)N}{j-i-1}\varepsilon < \frac{1}{N}6N\varepsilon = 6\varepsilon$, and the bound is even tighter in the wide case.

Our second claim is that $||F_\alpha - G|| < 11\varepsilon$ on $[\frac{i+j-1}{2N}, \frac{i+j+1}{2N}]$, where $(\frac{i}{N}, \frac{j}{N})$ is a maximal interval disjoint from P^α. Note that the interval on which we are evaluating the approximation need not have its boundaries on integer multiples of $\frac{1}{N}$, but it does contain both $\frac{i+j}{2N}$ and $\frac{c+d}{2}$, where (c, d) is the connected component of $[0, 1] \setminus P^\alpha$ in which $[\frac{i}{N}, \frac{j}{N}]$ lies. Then

$$|F_\alpha(x) - G(x)| = |F_\alpha(x) - F_\alpha(\frac{c + d}{2})| + |F_\alpha(\frac{c + d}{2}) - G(\frac{i + j}{2N})| + |G(\frac{i + j}{2N}) - G(x)$$

Unless the coincidence of x with $\frac{c+d}{2}$ makes it zero, the first summand is

$$|F_\alpha(x) - F_\alpha(\frac{c+d}{2})| = |\frac{c+d}{2} - x| \cdot |F'_\alpha(x)| < \frac{1}{N}\frac{2(j-i+2)N}{j-i}\varepsilon < 4\varepsilon,$$

because $|F'_\alpha(x)| \leq \frac{2\omega(F,[c,d])}{d-c}$ and as before knowing the approximate location of c and d allows us to conclude that $\omega(F,[c,d]) < (j-i+2)\varepsilon$ and $\frac{j-i}{N} < d-c$. (The numbers do not line up exactly with the previous ones because there is a off-by-one difference in the location of c.) For the second summand,

$$\begin{aligned} F_\alpha(\frac{c+d}{2}) - G(\frac{i+j}{2N})| &= |\omega(F,[c,d]) + \min(F(c),F(d)) - w - \min(F(\frac{i}{N}),F(\frac{j}{N})| \\ &\leq |\omega(F,[c,d]) - w| + |\min(F(c),F(d)) - \min(F(\frac{i}{N}),F(\frac{j}{N}))| \\ &\leq 2\varepsilon + \varepsilon \end{aligned}$$

And the last summand is also bounded by 4ε, since $w < \omega(F,[c,d])$ and the same bounds apply to G. Therefore, for $x \in [\frac{i+j-1}{2N}, \frac{i+j+1}{2N}]$, $|F_\alpha(x) - G(x)| < 11\varepsilon$.

Finally, we claim that $||F_\alpha - G|| < 11\varepsilon$ on the remaining intervals not yet discussed. That is because F_α and G are both linear on the remaining intervals, so their difference is maximized on one or both of the endpoints of those intervals, and bounds on those endpoints have already been found. Therefore, $||F_\alpha - G|| < 11\varepsilon$ everywhere. Thus F_α may be approximated as closely as desired by an effective sequence of functions computable in $\emptyset_{(2\alpha)}$.

Thus we see that $P^{\alpha+1}$ is $\Sigma_{2(\alpha+1)}$, completing the successor case of the induction. For the limit case, if λ is given as an effective sequence α_n with $\lim_{n\to\infty}\alpha_n = \lambda$, then $[p,q] \cap P^\lambda = \emptyset \iff \exists n[p,q] \cap P^{\alpha_n} = \emptyset$. Since the matrix is uniformly $\emptyset^\lambda$-computable, $\{[p,q] : [p,q] \cap P^\lambda = \emptyset\}$ is $\Sigma_\lambda = \Sigma_{2\lambda}$.

□

Being a Denjoy integral of rank α is $\Sigma_{2\alpha}$-complete

In this section we give a reduction $(WF, \neg WF) \to (ACG_*, \neg ACG)$ which maps trees of limsup rank α to functions of Denjoy rank α. The idea is that each node of the tree should contribute a finite length to the variation of the function. In most cases the total variation will be infinite as a result, but the way in which that infinite length is distributed will determine the rank of the function.

Proposition 1.4.12. *There is a computable reduction $T \mapsto F_T$ from trees to continuous functions on $[0,1]$ satisfying*

1. *If T is not well-founded, $F_T \notin ACG_*$.*

2. *If T is well-founded with $|T|_{ls} = \alpha$, then $F_T \in ACG_*$, and $|F_T|_D = \alpha$.*

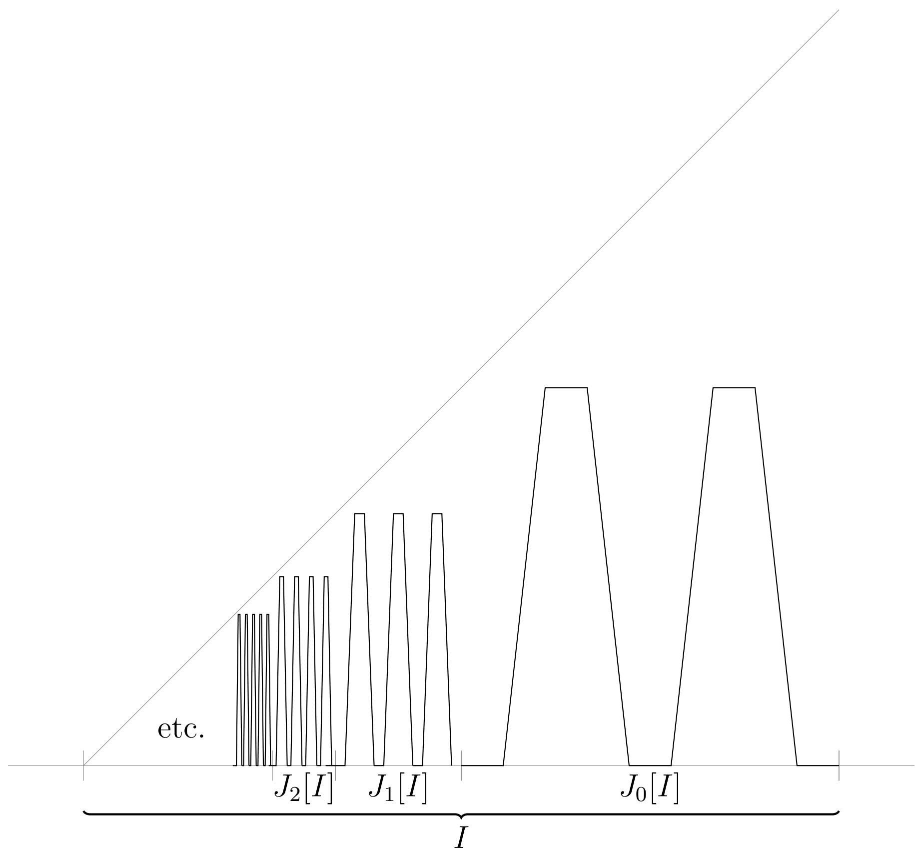


Figure 1.5: The function $G(I)$, when $\lambda(I) = 1$.

Proof. To each subset $X \subseteq \mathbb{N}$ and interval $I \subseteq \mathbb{R}$, one may associate a function $G(X, I) : \mathbb{R} \to \mathbb{R}$, uniformly in X and I, as follows. $G(X, I)$ will be supported on I. First associate to each interval I a function $G(I)$ defined as follows. For $[a, b] \subseteq [0, 1]$, let $[a, b][I]$ denote the corresponding interval in I, namely $[\min I + a\lambda(I), \min I + b\lambda(I)]$, where $\lambda(I)$ is the Lebesgue measure (here, just the length of I). For each interval $J_n = [\frac{1}{n+2}, \frac{1}{n+1}]$ define $G(I) \restriction J_n[I]$ to be a computable piecewise linear function as follows: Let $J_n[I]$ be divided into $4M + 1$ equally sized regions, where M is least such that $M \geq \frac{n+2}{\lambda(I)}$. Let $G \equiv 0$ on every fourth region including the first and last, $G \equiv \frac{\lambda(I)}{n+2}$ on the middle interval of each remaining cluster of three intervals, and let G be a linear interpolation on the rest of $J_n[I]$. See Figure 1.5. Note that M has been chosen so that the variation of $G(I)$ on $J_n[I]$ is at least 2.

Note that $G(I)(\min I + x) \leq x$ for all x. One might say that J_n supports M "hills" each of height $\frac{\lambda(I)}{n+2}$. $G(I)$ is computable uniformly in I.

Define $G(X, I)$ by defining $G(X, I) \restriction J_n[I] = G(I) \restriction J_n[I]$ for every $n \in X$, and $G(X, I) \equiv 0$ elsewhere.

Now for well founded T, we may define $F(T, I) : \mathbb{R} \to \mathbb{R}$ recursively as follows, proceeding by recursion on the usual rank of T. If $T = \emptyset$, let $F(T, I) \equiv 0$. Otherwise, assuming $F(T_n, I)$ is defined for all I and all $T_n = \{\sigma : n\sigma \in T\}$, let

$$F(T, I) = G(\{n : T_n \neq \emptyset\}, I) + \sum_{n:T_n\neq\emptyset} \sum_{\substack{H \text{ maximal} \\ \text{in } J_n[I] \text{ s.t.} \\ G(I)'(x)=0 \\ \text{for } x\in H}} F(T_n, H)$$

Let us check that $F(T, I)$ is computable uniformly in T and I, where I is represented by the numbers $\min I$, $\max I$.

Proceeding by induction on the usual rank of T, let us assume that each $F(T_n, H)$ is computable uniformly in T_n and H, and let us also assume that each $F(T_n, H)$ satisfies $F(T_n, H)(\min H + x) \leq x$. Bounded away from $\min I$, $F(T, I)$ is a sum of finitely many functions, uniformly computable by the inductive hypothesis.

Let us check that $F(T, I)(\min I + x) \leq x$ on each $J_n[I]$. Because of the shape of $G(I) \upharpoonright J_n[I]$, and because the $F(T_n, H)$ all satisfy the same bound, if $F(T, I)$ would fail to satisfy the bound anywhere, the failure would occur at the top of the first hill. Because $G(I)$ is felt there, the bound on $F(T_n, H)$ implies the bound on $F(T, I) \upharpoonright J_n[I]$. Since n was arbitrary, the bound holds for $F(T, I)$ as a whole. Together with the uniform computability of $F(T, I)$ away from zero, this bound also implies that $F(T, I)$ is computable.

So far $F(T, I)$ has been defined only for well-founded trees, but it has a natural extension to all trees. For any tree $T \in \mathbb{N}^{<\mathbb{N}}$, let $T \upharpoonright m$ be $\{\sigma \in T : |\sigma| \leq m\}$. If T is not well-founded, define $F(T, I) = \lim_{m\to\infty} F(T \upharpoonright m, I)$. To see that this is well-defined, note that the definition gives the same result for all well-founded trees, and that $||F(T \upharpoonright m+1, I) - F(T \upharpoonright m, I)|| < 2^{-n}\lambda(I)$. The bound can be seen by noting that $F(T \upharpoonright m+1, I) - F(T \upharpoonright m, I)$ consists of a sum of a large number of functions of the form $G(X, J)$ for J disjoint and $\lambda(J) < 2^{-n}\lambda(I)$ because the locally constant intervals of $G(X, H)$ satisfy $\lambda(J) < 2\lambda(H)$, and part of the induction hypothesis is that $||F(T, I)|| < \lambda(I)$.

Let us check that if $|T|_{ls}$ exists, then $F(T, I)$ is an indefinite Denjoy integral of rank $|T|_{ls}$.

We proceed by induction on the usual rank of the tree, starting with the singleton tree $\{\emptyset\}$. This tree has limsup rank 1, and $F(T, I) \equiv 0$, so it is Lebesgue integrable, and $|F(T, I)|_D = 1$ in both the narrow and wide senses. Let T be given, with limsup rank $\alpha + 1$. By induction, assume that for each n and J, $|T_n|_{ls} = |F(T_n, J)|_D$. By the definition of the limsup rank of T, there must be a number N such that for all $n \geq N$, $|T_n|_{ls} \leq \alpha$. Consider $F(T, I) \upharpoonright [0, \frac{1}{N+2}]$. For $x > 0$, $F(T, I) \upharpoonright [x, \frac{1}{N+2}]$ is a sum of finitely many functions of Denjoy rank at most α, by the induction hypothesis. Therefore, by Proposition 1.4.1, $F(T, I) \upharpoonright [x, \frac{1}{N+2}]$ is an indefinite Denjoy$_\alpha$-integral. By Proposition 1.4.2, $P^\alpha_{F(T,I)} \cap [x, \frac{1}{N+2}] = \emptyset$, where the P^α are defined as in the proof of Theorem 10, because that proof also showed that these P^α are precisely the P^α of Proposition 1.4.2 for the a.e. unique g whose D_α-integral is $F(T, I) \upharpoonright [x, \frac{1}{N+2}]$. Therefore, $P^\alpha_{F(T,I)} \cap [0, \frac{1}{N+2}] \subseteq \{0\}$. Since 0 is topologically isolated in P^α, $0 \notin P^{\alpha+1}$. So $P^{\alpha+1} \cap [0, \frac{1}{N+2}] = \emptyset$. Next consider $F(T, I) \upharpoonright [\frac{1}{N+2}, 1]$. It is the sum of finitely many functions of rank at most $\alpha + 1$, so $P^{\alpha+1} \cap [\frac{1}{N+2}, 1] = \emptyset$. Therefore $P^{\alpha+1} = \emptyset$ and the Denjoy rank of $F(T, I)$ is at most $\alpha + 1$.

Now we show that it is at least $\alpha+1$. Suppose that $|T_n|_{ls} = \alpha+1$ for some n. Then since for all J, $P^{\alpha}_{F(T_n,J)} \neq \emptyset$ by the inductive hypothesis, and the fact that $F(T,I) \restriction J = F(T_n, J)$ for some interval $J \subset J_n[I]$, we have $P^{\alpha}_{F(T,I)} \cap J_n \neq \emptyset$, so $|F(T,I)|_D \geq \alpha+1$. On the other hand, suppose that $\lim_{n\to\infty} |T_n|_{ls} = \alpha$. Then for every $\beta < \alpha$, there are infinitely many n for which $P^{\beta}_{F(T,I)} \cap J_n \neq \emptyset$. Therefore, $0 \in P^{\beta}_{F(T,I)}$. In fact, for any such n, for any maximal $J \subseteq J_n[I]$ on which $G(I)$ is constant, $P^{\beta}_{F(T,I)} \cap J \neq \emptyset$, because $P^{\beta}_{F(T_n,J)} \neq \emptyset$. Therefore, for each such n, $\sum_{(c,d)\in J_n\setminus P^{\beta}} \omega(F,(c,d)) \geq \sum_{(c,d)\in J_n\setminus P^{\beta}} |F(d)| - |F(c)|$ and the latter sum catches at least half the variance in $G(I) \restriction J_n[I]$, because if $G(I)(c) = \frac{\lambda(I)}{n+1}$ then $F(T,I)(c) \geq \frac{\lambda(I)}{n+2}$,a nd if $G(I)(c) = 0$ then $F(T,I)(c) \leq ||F(T_n,J)|| \leq \lambda(J)$. Then $\lambda(J) = \frac{\lambda(J_n[I])}{4M+1} = \frac{\lambda(I)}{(n+2)(n+1)(4M+1)} < \frac{\lambda(I)}{2(n+2)}$. So $\sum_{(c,d)\in J_n\setminus P^{\beta}} |F(d) - F(c)| \geq 1$ for each n for which $|T_n|_{ls} > \beta$, and because there are infinitely many such n, for each $x > a$, we have $\sum_{(c,d)\in[a,x]\setminus P^{\beta}} |F(d) - F(c)| = \infty$. So $F(T,I)$ is not absolutely continuous on P^{β} in any neighborhood of 0. Therefore, for every $\beta < \alpha$, $0 \in P^{\beta+1}$. If α is a limit, this implies $0 \in P^{\alpha}$. If $\alpha = \beta+1$, then directly $0 \in P^{\alpha}$. Therefore, $|F(T,I)|_D \geq \alpha+1$. This completes the proof that if $|T|_{ls}$ exists then $|F(T,I)|_D = |T|_{ls}$.

Now suppose that T is not well-founded. We will show that $F(T,I) \notin ACG_*$. To do so, we find a perfect set $E \subseteq I$ so that $F(T,I)$ is not AC_* on any $J \cap E$, which shows $F(T,I) \notin ACG_*$ by Theorem 11.

Let $\{m_i\}_{i<\omega}$ be an infinite path through T. Let $E_0 = \{I\}$. For each k, let

$$E_{k+1} = \{J \subseteq J_{m_k}[K] : K \in E_k \text{ and } G(K) \restriction J \text{ constant and } J \text{ maximal for this property}\}.$$

Let $E = \cap_{k=0}^{\infty}(\cup E_k)$.

Then for any $x \in E$ and any open interval J containing x, let ε be small enough that $(x-\varepsilon, x+\varepsilon) \subseteq J$. Let k_0 be large enough that the elements of E_{k_0} have length less than ε. Then there is an $H \in E_{k_0}$ such that $x \in H$. Let us see that $F(T,I)$ is not AC_* on $H \cap E$.

First note from the definition of the E_k that for each $K \in E_k$, there is $K' \subseteq E_{k+1}$ with $K' \subset K$. So by compactness, for each k and each $K \in E_k$, there is a $y \in K$ such that $y \in E$.

Now for each E_k with $k > k_0$, define a finite set of disjoint intervals (a_j, b_j) as follows. For each $K \in E_k$ for which $K \subseteq H$, let $c_1 < c_2 < \cdots < c_r$ be an ordered list of elements of E, consisting of one element of E for each interval in E_{k+1} that is contained in K. Again, the location of the c_i, alternating in imagine between $F(T,I)(c_i) \geq \frac{\lambda(K)}{m_k+2}$ and $F(T,I)(c_i) < \frac{\lambda(K)}{2(m_k+2)}$, guarantee that the intervals $(c_1,c_2),(c_2,c_3),\ldots,(c_{r-1},c_r)$ capture most of the variation of $G(K) \restriction J_{m_k}[K]$. That is,

$$\sum_{i=1}^{r-1} |F(T,I)(c_{i+1}) - F(T,I)(c_i)| \geq 1.$$

Taking the set of all the intervals (c_i, c_{i+1}) gleaned from all $K \in E_k$ with $K \subseteq H$, we obtain the finite sequence (a_i, b_j), which is disjoint because the different K are. Because

$\cup_j (a_j, b_j) \subseteq E_k$, we have

$$\sum_j b_j - a_j \leq \mu(E_k) < 2^{-k}\lambda(I),$$

where the last inequality is justified because each $K \in E_i$ is replaced in E_{i+1} by a set of subintervals of $J_{m_i}[K]$, and $\lambda(J_{m_i}[K]) \leq \frac{\lambda(K)}{2}$. Therefore, as k goes to infinity, the selected intervals have a total measure approaching zero. However, $E_k \cap H$ has at least 2^{k-k_0}-many intervals, because each $K \in E_i$ is replaced by at least two intervals in E_{i+1}. Therefore, $\sum_j |F(T,I)(b_j) - F(T,I)(a_j)| > 2^{k-k_0}$, and $F(T,I)$ is not AC_* on $E \cap H$. □

Chapter 2

Computability in Subshifts

In this chapter, we discuss the connections between computability theory and symbolic dynamics by examining the subshift invariants of entropy, Medevedev degree, and effective dimension spectrum. The latter is defined in Section 2.6. In Section 2.1 we establish the basic notions in computability theory and in symbolic dynamics. In Section 2.2 we review what is known about the values the entropy can take for subshifts subject to various restrictions. In Section 2.3 we review what is known about the Medvedev degrees that subshifts may inhabit. In Section 2.4 we review what is known about the effective dimensions that trajectories of a subshift may have. In Section 2.5 we show that entropy and Medvedev degree of subshifts are independent, and that every right-r.e. entropy in $[0, 1)$ may combine with every Medvedev degree in a one-dimensional Π^0_1 subshift or a two-dimensional shift of finite type using an alphabet with only to symbols. In Sectoin 2.6 we introduce the effective dimension spectrum, give conditions under which it is a simple interval, and calculate it for a certain minimal subshift.

2.1 Preliminaries

Notation

We use standard notation. In general, capital letters are either large integers or subsets of A^G, where A is a finite set, considered an alphabet, and G is $\mathbb{N}, \mathbb{Z}, \mathbb{N}^2$ or $\mathbb{Z}^2$. Letters a, b, c, d are usually numbers, f, g, h functions, i, j, k, l, m, n indices or lengths, p, q rationals, r, s, t reals, u, v, w, σ, τ strings in $2^{<\omega}$, and $x, y, z \in 2^G$ where G is as before. The length of a string σ is denoted $|\sigma|$. Concatenation of strings in 2^ω is denoted by simply writing the string names one after the other, for example $\sigma 1$. We use σ^N to denote the concatenation of σ N-many times. Finite strings are also called finite sequences. The notation σ^ω refers to the concatenation of σ infinitely many times to form an infinite sequence. For a finite or an infinite sequence x, the notation $x[a, b]$ refers to the finite sequence $x(a)x(a+1)\dots x(b-1)$. The length of $x[a, b]$ is $b - a$. If $b \geq |x|$, then $x[a, b]$ is not defined. In the special case when

$a = 0$, we write $x \upharpoonright b$ for $x[a, b]$. If $\sigma = x[a, b]$ for some a, b, we say that σ is a subword of x or σ appears in x.

If $G = \mathbb{Z}^d$ and $x \in A^G$ then $x \upharpoonright n$ refers to the d-dimensional array $u \in A^{(2n-1)^d}$ such that u is equal to the central block of x. We may also say u is a subword of x. Similarly, if $G = \mathbb{N}^d$ and $x \in A^G$, then $x \upharpoonright n \in A^{n^d}$.

If $\sigma \in A^{n^d}$ for some n, then $[\sigma]$ denotes $\{x \in A^{\mathbb{N}^d} : x \upharpoonright n = \sigma\}$. Similarly, if $\sigma \in A^{(2n-1)^d}$ for some n, then $[\sigma]$ denotes $\{x \in A^{\mathbb{Z}^d} : x \upharpoonright n = \sigma\}$. It will be clear form context whether $\mathbb{N}$ or $\mathbb{Z}$ should be used. The sets A^G have a natural topology whose basic open sets are the $[\sigma]$.

Computability-Theoretic Notions

A real number s is right-recursively-enumerable, or right-r.e., if there is an algorithm producing a decreasing sequence of rationals $q_0, q_1, \ldots$ such that $\lim_{n\to\infty} q_n = s$.

If σ is a finite string, $K(\sigma)$ refers to its prefix-free Kolmogorov complexity. For an introduction to Kolmogorov complexity, see [18]. Informally, Kolmogorov complexity measures, up to an additive constant factor, the number of bits of information contained in the string. It satisfies the following properties, which can also be found in [18].

Proposition 2.1.1. *There is a C such that for all n, $K(n) \leq 2\log n + C$.*

Proposition 2.1.2. *There is a C such that for all $\sigma, \tau \in 2^{<\mathbb{N}}$,*

$$K(\sigma\tau) \leq K(\sigma) + K(\tau) + C.$$

Proposition 2.1.3. *There is a C such that for all $\sigma, \tau \in 2^{<\mathbb{N}}$,*

$$K(\sigma, \tau) = K(\sigma) + K(\tau|\sigma^*) \pm C.$$

Proposition 2.1.4. *There is a C such that for all $\sigma, \tau \in 2^{<\mathbb{N}}$,*

$$K(\sigma, \tau) \leq K(\sigma\tau) + K(|\tau|) + C.$$

Putting the last two propositions together gives,

Proposition 2.1.5. *Letting C be large enough to satisfy the previous propositions, for all σ, τ,*

$$K(\sigma\tau) \geq K(\sigma) + K(\tau|\sigma^*) - 2\log|\tau| - 3C.$$

Proof. Combining the last two propositions,

$$\begin{aligned} K(\sigma\tau) + K(|\tau|) + C &\geq K(\sigma) + K(\tau|\sigma^*) - C \\ K(\sigma\tau) &\geq K(\sigma) + K(\tau|\sigma^*) - K(|\tau|) - 2C \\ &\geq K(\sigma) + K(\tau|\sigma^*) - 2\log|\tau| - 3C \end{aligned}$$

□

Though the constants C above may all be different, we choose one C larger than all of them, and use it frequently later.

We will use the notion of *effective dimension* defined in [20]. An equivalent definition [21] is:

Definition 2.1.1. *The* effective dimension *of a sequence $x \in 2^\omega$ is*

$$\dim(x) = \liminf_{n\to\infty} \frac{K(x \upharpoonright n)}{n}.$$

The *packing dimension* $\mathrm{Dim}(x)$ is defined similarly, using lim sup instead of lim inf.

Medvedev reducibility is a way of comparing two sets $A, B \subseteq 2^\omega$ in terms of how complicated it is to isolate an individual sequence from those sets.

Definition 2.1.2. *Given $C, B \subseteq A^G$, we say that C is* Medvedev reducible *to B, and write $C \leq_w B$, if there is a Turing functional Γ whose domain includes all of B such that for all $x \in B$, $\Gamma(x) \in C$. If $C \leq_w B$ and $B \leq_w C$ then we say C and B are* Medvedev equivalent *and write $C \equiv_w B$.*

Note that Medvedev degree can apply to all finite A and previously mentioned G, not just $A = 2$ and $G = \mathbb{N}$. For the non-$2^\mathbb{N}$ cases, assume a standard way of encoding the elements of A^G.

Subshifts, Conjugacy and Invariants

Let A be a finite set of symbols, considered as an alphabet. A subshift is a set $X \subseteq A^G$ that is topologically closed and closed under the shift operation, where $G = \mathbb{N}^d$ or $\mathbb{Z}^d$, where $d = 1, 2, \ldots$. In general there are d shift operations to be closed under, on for each direction. It can be shown that a subshift may be characterized by its set of forbidden strings, that is, $F = \{\sigma : \text{ for all } x \in X, \sigma \text{ does not appear in } x\}$ may be used to define X in the sense that $X = \{x \in A^G : \text{ no string of } F \text{ appears in } x\}$. If F is empty, $X = A^G$ and this shift is called the full shift. If X is not the full shift, then some σ is forbidden, and as a result infinitely many other strings are forbidden (all those which contain σ as a subword.) However, a finite set F can still be used to define a shift

$$X_F = \{x \in A^G : \text{ no string of } F \text{ appears in } x\},$$

and if X can be defined in this way, it is called a *shift of finite type*

Definition 2.1.3. *A subshift $X \subseteq 2^G$ is called a* shift of finite type *if it may be written as $X = \{x \in 2^G :$ no string of F appears in $x\}$ for some finite set of strings F.*

Two subshifts $X, Y \subseteq 2^G$ are *topologically conjugate* if there is a shift-invariant homeomorphism $h : X \to Y$. The conjugation function h is in fact a finite object because the compactness of the spaces brings a finiteness to the action of h on X near zero, and the shift-invariance then defines h everywhere else. Therefore, the function h is also called a *sliding block code* because the ith bit of $y = h(x) \in Y$ may be determined from $x[i-N, i+N]$ for a fixed sufficiently large N, and so one could imagine sliding a $2N$-sized window over x and using the block visible in the window to read off the appropriate symbol of y. For details and the proof of the following, we refer the reader to [19].

Proposition 2.1.6. *The property of being a shift of finite type is invariant under conjugacy.*

Despite the simplicity of the conjugation functions, equivalence of subshifts under conjugacy is a universal Borel equivalence relation [4].

Conjugacy invariants are studied both for the satisfaction of curiosity about the properties of various subshifts, and as a way to tell when two subshifts are not conjugate. By [4] there will be no simple invariant that exactly captures conjugacy, however.

A very important invariant is the entropy.

Definition 2.1.4. *The* entropy *of a subshift $X \subseteq A^G$ is defined as*

$$\operatorname{ent}(X) = \lim_{n\to\infty} \frac{\log |\{x \upharpoonright n : x \in X\}|}{|F_n|},$$

where $|F_n|$ depends on G, and is the number of symbols in $x \upharpoonright n$. For $G = \mathbb{N}^d, F_n = n^d$; for $G = \mathbb{Z}^d, F_n = (2n-1)^d$.

The entropy measures the rate of growth of the number of permitted strings of length n as n goes to infinity. For example, the full shift on $2^{\mathbb{N}}$ has entropy 1. It is well-known that the limit which defines the entropy is decreasing.

Another subshift invariant is the Medvedev degree. If two subshifts are topologically conjugate, then they have the same Medvedev degree because the homeomorphism that witnesses their conjugacy is both computable and computably invertible. Therefore, this function also witnesses the Medvedev equivalence.

A useful operation one can do with subshifts is to take their product. Given two subshifts $X \subseteq A^G$ and $Y \subseteq B^G$, where G is the same for both subshifts and A and B are any two finite alphabets, the product shift $X \times Y$ is $\{x \times y : x \in X \text{ and } y \in Y\}$, where $x \times y$ is the trajectory of $(A \times B)^G$ satisfying $(x \times y)(n) = (x(n), y(n))$. One may verify that this set is closed and shift invariant. We now consider the entropy and Medvedev degree of product shifts.

It is well-known that $\operatorname{ent}(X, Y) = \operatorname{ent}(X) + \operatorname{ent}(Y)$, and it follows from the complete independence of the x and y parts of any element of $X \times Y$. One may calculate:

$$\begin{aligned}\operatorname{ent}(X \times Y) &= \lim_{n\to\infty} \frac{\log |\{x \times y \upharpoonright n : x \times y \in X \times Y\}|}{n}\\ &= \lim_{n\to\infty} \frac{\log |\{x \upharpoonright n : x \in X\}| \cdot |\{y \upharpoonright n : y \in Y\}|}{n}\\ &= \lim_{n\to\infty} \frac{\log |\{x \upharpoonright n : x \in X\}| + \log |\{y \upharpoonright n : y \in Y\}|}{n} = \operatorname{ent}(X) + \operatorname{ent}(Y).\end{aligned}$$

As for the Medvedev degree, we can say that if X has a computable element x_0, (that is, X has Medvedev degree 0), then $X \times Y$ is Medvedev equivalent to Y. For given any $x \times y$ in $X \times Y$, one may trivially extract its y part; and given any $y \in Y$, one may computably produce $x_0 \times y$.

2.2 Entropy of subshifts

Any number in $[0, \infty)$ can be the entropy of a subshift, but certain restrictions, such as having the set of forbidden words be computably enumerable, or being a shift of finite type, also put restrictions on the possible values for the entropy. This section summarizes what is known about the possible values the entropy of a subshift can take subject to various restrictions. References are provided for everything but the development of the density-r subshifts; the author expects that they would be known to the community, but is not aware of a source.

In one dimension

Let us consider first the case of one dimension, so $G = \mathbb{N}$ or $\mathbb{Z}$. Without any other restrictions, the set of possible entropies is $[0, \infty)$. Because the full shift 2^G has entropy 1, if X_r has entropy $r \in [0, 1)$, then $X \times (2^G)^n$ has entropy $n + r$, where $n = 0, 1, 2, \ldots$. Therefore, it suffices to show that for any $r \in [0, 1)$, there is a subshift X_r with entropy r.

We construct a subshift on 2^ω with the strategy of only using a fraction r of the total bits to encode information. This strategy was already used in [11] to construct shifts in Cantor space of every entropy $r \in [0, 1]$, but we organize the coding bits differently here.

Let $\{\rho_i\}_{0<i<\omega}$ be any sequence of 0s and 1s. Intuitively, an example of a sequence X of effective dimension $s = \sum_{i=1}^{\infty} \rho_i 2^{-i}$ is one for which $X(n) = 0$ whenever $n \equiv 2^{i-1} (\mod 2^i)$ for an i such that $\rho_i = 0$, and where the remaining bits of X are taken from a sequence of effective dimension one. We would like a subshift with includes all the shifts of sequences of this form for $\{\rho_i\}_{0<i<\infty}$ with $\sum_{i=1}^{\infty} \rho_i 2^{-i} \leq r$. Thus

Definition 2.2.1. *The* density-r subshift $X_r \subseteq 2^G$ *is defined as follows. Given σ and k (the amount to shift), let $\rho_i = 1$ if $\sigma(j) = 1$ for any j in range such that $j - k \equiv 2^{i-1} \mod 2^i$. If $\sigma(j) = 0$ for all such j, let $\rho_i = 0$. The ρ_i depend on k and σ, but context will always resolve the ambiguity.*

Let X_r be the subshift obtained by forbidding σ if for all k, $\sum_{i=1}^{\infty} \rho_i 2^{-i} > r$.

Note that almost $2|\sigma|$ values of k need to be checked. Therefore, if r is right-r.e., then the set of forbidden strings is computably enumerable, so X_r is a Π^0_1 subshift.

Proposition 2.2.1. *The entropy of X_r is r.*

Proof. Let $r = \sum_{i=1}^{\infty} \tau_i 2^{-i}$. If r is a dyadic rational, chose the expression which ends in zeros. For a lower bound on the entropy, we will count those strings σ which satisfy $\sigma(j) = 0$ whenever $j \equiv 2^{i-1} (\mod 2^i)$ and $\tau_i = 0$. Such σ are permitted words regardless of what is encoded in their "free" bits. Let $f(n)$ denote the number of free bits in a string of length n under this scheme. Then $\lim_{n\to\infty} \frac{f(n)}{n} = r$, so

$$\mathrm{ent}(X_r) = \lim_{n\to\infty} \frac{\log |X_r \upharpoonright n|}{n} \geq \lim_{n\to\infty} \frac{\log 2^{f(n)}}{n} = \lim_{n\to\infty} \frac{f(n)}{n} = r.$$

Now we find an upper bound on the entropy. For any σ, writing $k = a_0 + 2^1 a_1 + \cdots + 2^l a_l$, where each $a_i \in \{0, 1\}$, one may observe that the choice of a_0 determines which half of the bits of σ to use to evaluate ρ_1, and in general, the choice of a_i determines which half of the so-far-unused bits of σ to use to evaluate ρ_{i+1}. Therefore, whenever it is possible to choose a_{i_0} so that $\rho_{i_0+1} = 0$, it is also possible to choose the other value for a_{i_0}, guaranteeing $\rho_i = 0$ for all $i > i_0 + 1$.

Therefore, if σ and k are such that $\sum_{i=1}^{\infty} \rho_i 2^{-i} \leq r$, it is possible to choose k' so that (associating the ρ_i to k and the ρ'_i to k') $\sum_{i=1}^{\infty} \rho'_i 2^{-i} \leq \sum_{i=1}^{\infty} \tau_i 2^{-i} = r$ with the additional restriction that $\rho'_i \leq \tau_i$ for each i. If $\sum_{i=1}^{\infty} \rho_i 2^{-i} < r$, but $\rho_i > \tau_i$ for some i, let i_0 be least such that $\rho_{i_0} = 0$ and $\tau_{i_0} = 1$. We know i_0 exists because $\sum \rho_i 2^{-i} < \sum \tau_i 2^{-i}$, and furthermore is the first index at which ρ_i and τ_i differ. Let $k' = k + 2^{i_0 - 1}$. Then $\rho'_{i_0} = 1 = \tau_{i_0}$, and $\rho'_i = 0 \leq \tau_i$ for all $i > i_0$. If $\sum_{i=1}^{\infty} \rho_i 2^{-i} = r$, then both are dyadic rationals and both are finite sums, so $\rho_i = \tau_i$ for all i already.

Based on the above, when counting the number of permissible strings σ of length n, it suffices to consider all choices of k, and all choices of σ which satisfy $\rho_i \leq \tau_i$ for the given k.

For any n, there are at most $2n$ choices of k which have distinct residues mod $2, 4, \ldots, 2^{\lceil \log n \rceil}$. Let $g(n)$ be the maximum, taken over all values of k, of the number of free bits a string σ can have while satisfying $j - k \equiv 2^{i-1} (\mod 2^i)$ and $\tau_i = 0 \implies \sigma(j) = 0$. Note that $\lim_{n\to\infty} \frac{g(n)}{n} = r$. Then for each value of k there are at most $2^{g(n)}$ strings σ permitted by that k, so

$$\operatorname{ent}(X_r) = \lim_{n\to\infty} \frac{\log |X_r \upharpoonright n|}{n} \leq \lim_{n\to\infty} \frac{\log 2n 2^{g(n)}}{n} = \lim_{n\to\infty} \frac{\log 2n}{n} + \frac{g(n)}{n} = r.$$

□

The construction of X_r provides answers to two more questions about how the entropy characterization changes when restrictions are made on what type of subshift can be used. Because any subshift on 2^G has entropy at most 1, and each $X_r \in 2^G$, we see that the entropies of the subshifts in 2^G are exactly the numbers in $[0, 1]$. Furthermore, by the definition of the entropy together with the that the limit in that definition is decreasing, it is clear that if a subshift in one dimension is Π^0_1, then its entropy is right-r.e. Going in the other direction, if r is right-r.e., then the X_r defined above is Π^0_1, as one may verify by checking its definition.

One-dimensional shifts of finite type

When we restrict X to be a shift of finite type, the situation becomes significantly more restricted. The entropies of the shifts of finite type are exactly the rational multiples of logarithms of Perron numbers (see e.g. [19] for the definition and discussion). The entropy of an SFT can be effectively computed, because it is the logarithm of the spectral radius of a certain matrix associated with the SFT. Because the graph underlying this matrix will interest us later, we present some standard ways of analyzing one-dimensional SFTs here. For more details, see [19].

The *nth higher block shift* of a subshift $X \subseteq 2^G$, where G is $\mathbb{N}$ or $\mathbb{Z}$, is

$$X^{[n]} = \{\{x_i^j\}_{i<n, j\in G} \in (A^n)^G : \text{ for all } j, x_0^j \dots x_{n-1}^j \text{ is a permitted word of } X \\ \text{and } x_{i+1}^j = x_i^{j+1} \text{whenever the value of the indices are in range}\}.$$

One may verify that $X^{[n]}$ is a subshift and that it is naturally conjugate to X. Thus it has the same entropy.

If X is an SFT, consider $X^{[n]}$ when n is longer than the longest forbidden word of X. Let $B \subseteq A^n$ be the collection of permitted words from A^n; these are the symbols of $X^{[n]}$. So $X^{[n]} \subseteq B^G$, and $X^{[n]}$ is not only a shift of finite type, but it is what is called a *one-step* shift of finite type, meaning that it can be characterized by its length-2 forbidden words. Therefore, one could describe $X^{[n]}$ by a directed graph whose nodes are in one-to-one correspondence with the elements of B, and for $u, v \in B$, an arrow begins at u and terminates in v if and only if uv is a permitted word of $X^{[n]}$. The set of all infinite or bi-infinite (according to whether $G = \mathbb{N}$ or $\mathbb{Z}$) paths through this graph is exactly $X^{[n]}$. The adjacency matrix of this graph is the matrix from whose spectral radius one may compute the entropy of the SFT. For details, see [19].

In two dimensions

In two dimensions, it is still true that the product of any shift with the full shift results in an increase by one in the entropy, so it suffices to consider only entropies in $[0, 1]$. Within that restriction, it is possible to make density-r subshifts just as in the one-dimensional case.

Definition 2.2.2. *Let $G = \mathbb{Z}^2$ or $\mathbb{N}^2$. The* density-r subshift *$X_r \subseteq 2^G$ is defined as follows. Given $\sigma \in 2^{n^2}$ and k (the amount to shift in one dimension), let $\rho_i = 1$ if $\sigma(j_0, j_1) = 1$ for any j_0, j_1 in range such that $j_0 - k \equiv 2^{i-1} (\mod 2^i)$. If $\sigma(j_0, j_1) = 0$ for all such j, let $\rho_i = 0$.*

Let X_r be the subshift obtained by forbidding σ if for all k, $\sum_{i=1}^{\infty} \rho_i 2^{-i} > r$.

The only difference between this and the one-dimensional denity-r subhift is that instead of free bits and constrained-to-zero bits, we have free columns and constrained-to-zero columns. With proofs almost identical to the one-dimensional case, we may see that in two dimensions also X_r has entropy r, that every entropy in $[0, 1]$ is obtainable with only two symbols, and that the right-r.e. r correspond to the Π_1^0 X_r.

Two-dimensional shifts of finite type

The problem of characterizing the entropies of multidimensional SFTs was open until [13] in 2010. In sharp contrast to the algebraic characterization of the entropies of the one-dimensional SFTs, the characterization in the multidimensional case is recursion-theoretic.

Theorem 13. *(Hochman-Meyerovitch) The entropies of the d-dimensional shifts of finite type for $d \geq 2$ are exactly the right-r.e. numbers.*

The method of the proof was to encode the workings of a certain Turing machine into the rules of the SFT, a Turing machine which in turn would enforce a structure on the SFT to make it act like one of the two-dimensional density-r subshifts above. It was already known how to force elements of an two-dimensional SFT to encode computations (see e.g. [28, 24, 25, 10]); the novelty was to use those computations to control the entropy of the subshift that resulted.

One constant source of technical difficulty in the encoding of Turing machines in SFTs is that one must balance the need for assuring that there every element of the SFT contains some computation (meaning new computations must be begun at regular spatial intervals) and the need for ensuring that computations of arbitrary size can exist. The usual solution, first described in [28], is to pepper the tilings sparsely with computations of various sizes, spread apart so that larger computations can use the space left behind by smaller computations. One technical device for making such constructions easy, which we will also use, is a substitution construction. The following is paraphrased from [13]:

Let $G = \mathbb{Z}^2$. A *substitution rule* is a map $s : A \to A^{k^2}$ which sends each symbol of A to a $k \times k$ block of symbols of A. One may iterate the rule by applying it to each symbol of the resulting block. so that $s^n : A \to A^{(k^n)^2}$. Define a subshift $W \subseteq A^G$ by saying that σ is forbidden in W if there is no $s^n(a)$ in which it appears as a subword. Define $s_\infty : W \to W$ so that $s_\infty(x)$ is the result of replacing each symbol in x with its image under s. (So the image of each sub-block increases in size by a factor of k^2.) Say that x is derived from y if $s_\infty(y) = x$, where it is permitted to first shift x less than k vertically and horizontally in order to get it to "line up" with $s_\infty(y)$. Each x is derived from some y, and if this y is unique then we say s has *unique derivation*. Then using crucially a theorem of Mozes [24], Hochman and Meyerovitch showed the following:

Proposition 2.2.2. *Let $s : A \to A^{k^2}$ be a substitution rule with unique derivation and let W be the associated subshift. Then there exists an alphabet Δ, a SFT $\widetilde{W} \subseteq \Delta^G$, and a map $\varphi : \Delta \to A$ such that $\varphi(\widetilde{W}) = W$. Furthermore,* $\operatorname{ent}(\widetilde{W}) = 0$.

In other words, one may define a subshift W as above using a substitution rule and be able to treat it as a SFT with entropy zero, the only difference being that there are finitely many versions of each symbol of W and the SFT may have different rules for each. We will use substitution to create a subshift that includes computations in the next section.

2.3 Medvedev degree of subshifts

In this section we review what is known about the Medvedev degrees of one-dimensional Π^0_1 subshifts and two-dimensional SFTs. In short, it was known that every Medvedev degree not immediately prohibited is possible ([32, 23]). Our contribution is to notice (in the one-dimensional case) and reconstruct (in the two-dimensional case) such subshifts with zero entropy.

Medvedev subshifts

Simpson [31] showed that there is a two-dimensional SFT in every Π^0_1 Medvedev degree, and asked whether there was a one-dimensional Π^0_1 subshift in every Π^0_1 Medvedev degree. Miller answered this in [23] with the following family of subshifts, one for each Π^0_1 class. We present the construction here because later we will modify it to prove that the Medvedev degree of a subshift is independent from its entropy.

The idea behind the construction is this. For a given Π^0_1 class P, each element $y \in P$ will be encoded into some $x \in M_P$ by ensuring that every bit of y is coded into the tail of x, and that these bits can be recovered from any tail of x. To encode the first bit of y into the tail of x, we demand that x be an infinite concatenation of certain words $abb, abbb$ if $y(0) = 0$, and certain other words $baa, baaa$ if $y(1) = 1$, where a and b are chosen so that the two possibilities can be distinguished (for example, $a = 0, b = 1$.) Subsequent bits are encoded by placing restrictions on the order in which $abb, abbb$ (respectively $baa, baaa$) can be concatenated.

More formally, let λ denote the empty string. Fix $a_\lambda, b_\lambda \in 2^{<\omega}$ with a self-aligning property: for every sufficiently long σ, if σ is a subword of $c_0c_1 \dots, c_k$, $c_i \in \{a_\lambda, b_\lambda\}$, then the word boundaries of the c_i contained in σ may be uniquely determined from σ. For example, Miller used $a_\lambda = 0, b_\lambda = 1$. Later we will use $a_\lambda = 0, b_\lambda = 1^N$ for some large N.

Proceeding by induction, define

$$a_{\sigma 0} = a_\sigma b_\sigma b_\sigma \quad b_{\sigma 0} = a_\sigma b_\sigma b_\sigma b_\sigma$$
$$a_{\sigma 1} = b_\sigma a_\sigma a_\sigma \quad b_{\sigma 1} = b_\sigma a_\sigma a_\sigma a_\sigma$$

Definition 2.3.1 (Miller). *Given a closed set $P \subseteq 2^\omega$, let its* Medvedev subshift M_P *be the subshift for which the following words are forbidden:*

1. *Any word which cannot be parsed as a subword of a concatenation of a_λ and b_λ.*

2. *For each σ, forbid $a_\sigma a_\sigma a_\sigma a_\sigma, a_\sigma a_\sigma b_\sigma b_\sigma, a_\sigma b_\sigma a_\sigma b_\sigma$*

3. *For each σ, forbid the same strings as above, but with a and b reversed.*

4. *For each $\sigma \notin P$, forbid a_σ, b_σ.*

Lemma 2.3.1. *If $x \in M_P$, then for every n, there is a unique σ_n of length n so that x eventually consists of a concatenation of the words a_σ, b_σ. Furthermore, all such σ are comparable.*

Proof. Proceeding by induction on n, we start with $n = 0$. The prohibition on words that cannot be parsed as a concatenation of a_λ and b_λ, together with the self-aligning condition on these words, guarantees that each $x \in M_P$ is eventually a concatenation of a_λ and b_λ. Now suppose that each $x \in M_P$ is eventually a concatenation of a_σ and b_σ. By the second and third prohibitions, two occurrences of a_σ must be followed by either b_σ or $a_\sigma b_\sigma$; this b_σ must be followed by at least two, but not more than three, a_σ. Another way of stating these restrictions is that whenever $a_\sigma a_\sigma$ occurs in x, the rest of x from then on must be a concatenation of $b_\sigma a_\sigma a_\sigma$ and $b_\sigma a_\sigma a_\sigma a_\sigma$, and the parallel fact is true if ever $b_\sigma b_\sigma$ occurs in x. Exactly one of these two must occur. So x is either a concatenation of $a_{\sigma 0}, b_{\sigma 0}$, or of $a_{\sigma 1}, b_{\sigma 1}$. To see that x is not additionally eventually parseable as a concatenation of a_τ, b_τ for $\sigma \not\prec \tau$, observe that a_σ is not a subword of any concatenation of a_τ, b_τ for such τ. □

Proposition 2.3.2. *The subshift M_P is Medvedev equivalent to P.*

Proof. To each x in M_P, one may effectively associate a $y \in 2^\omega$ by

$$y = \cup\{\sigma : x \text{ is eventually a concatenation of } a_\sigma \text{ and } b_\sigma\}.$$

The effectiveness follows from the above observation that the first occurrence of $a_\sigma a_\sigma$ (respectively $b_\sigma b_\sigma$) guarantees the eventual concatenation of $a_{\sigma 1}, b_{\sigma 1}$ (respectively $a_{\sigma 0}, b_{\sigma 0}$). To see that y is in P, observe that since a_σ, b_σ occur in y, the fourth prohibition implies that $\sigma \in P$.

Conversely, if $y \in P$, then $\cup_{n>0} a_{y\upharpoonright n} \in M_P$. For more details, see [23]. □

Note that we have not yet made any restriction on the complexity of P, so in fact there are subshifts in every Medvedev degree which contains a closed set. If P is a Π^0_1 class, then one may observe that the set of forbidden sequences is computably enumerable, and so M_P is a Π^0_1 subshift, which was Miller's original goal. He had no need to analyze the entropy of the subshifts he produced, but we will use the following fact.

Proposition 2.3.3. *The subshifts M_P all have entropy 0.*

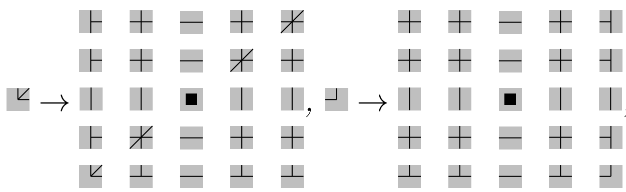
Proof. For each $x \in M_P$, and for each k, there is a σ of length k such that except for a finite initial segment, x is made of concatenations of a_σ and b_σ, where $|a_\sigma|, |b_\sigma| > k$. Therefore, the effective dimension of x is less than $\frac{1}{k}$. Since this is true for all k, the effective dimension of x is 0. Since x was arbitrary, $\sup_{x \in M_P} \dim(x) = 0$. Therefore, the constructive dimension of M_P is zero, so $\mathrm{ent}(M_P) = 0$. □

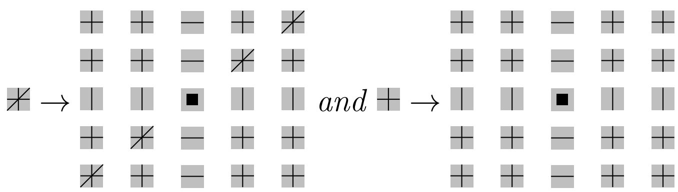
The existence of a zero entropy two-dimensional SFT with arbitrary Medvedev degree is almost guaranteed by [31, 25, 28], but in those papers the authors did not have any reason to consider the entropy, and their construction happens to have positive entropy. Below we will, for completeness, give a version of that proof that results in zero entropy.

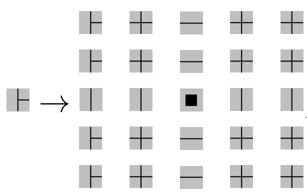
We closely follow the Hochman and Meyerovitch construction because in that construction the entropy is very carefully controlled, adding in the ideas of [31, 25] to get the result.

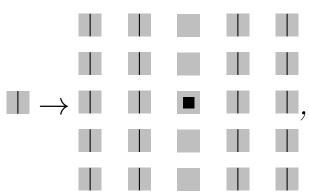
Definition 2.3.2. *The* subshift of boards B *is defined as follows.*

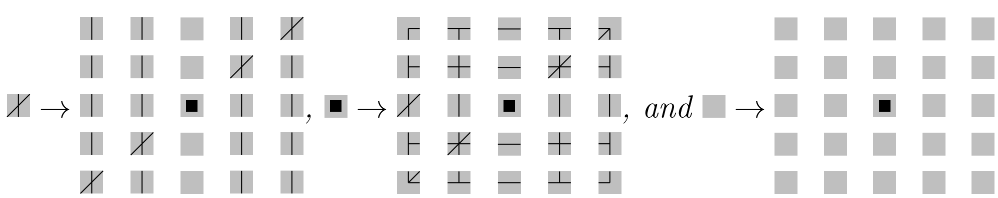
Construct an SFT $\widetilde{W}$ by substitution using the following rules:

1.  *, and their 180° rotations.*

2. 

3.  *, and its other three rotations.*

4.  *, and its 180° rotation.*

5. 

Let O_v be the shift with two symbols ↕ and ⇕, and the constraint that ↕ and ⇕ cannot appear horizontally next to each other. Let $B = O_v \times \widetilde{W}$.

By Proposition 2.2.2, $\widetilde{W}$ has zero entropy. The restriction on O_v implies that all its elements are constant on columns. Because an $n \times n$ square has only 2^n ways to satisfy these restrictions, O_v has entropy zero. Therefore, $\operatorname{ent}(B) = O$. In Section 7 of [13], it is described how to superimpose the action of a Turing machine on a certain SFT named $X \times \widetilde{R}$ in their paper. While $X \times \widetilde{R} \neq B$, the two shifts are very similar, and many of the same conclusions hold for the same reasons. In particular,

Proposition 2.3.4 (Proposition 7.1 of [13])**.** *given a Turing machine T, there exists an SFT Y_T superimposed over B such that the following are equivalent:*

1. *$(x, r) \in B$ is represented in Y_T.*

2. *For each finite or infinite board induced by r and containing the symbol corner-slash, when T is run on the sequence of 0's and 1's induced by the $\updownarrow$'s and $\Updownarrow$'s of x on the board, the number of steps it runs without halting is at least equal to the number of rows in the board.*

Furthermore, $\mathrm{ent}(Y_T) = O$.

Proof. See the proof of Proposition 7.1 in [13]. □

It is clear that the same is true of $B \times O_v$ if the Turing machine T should read from the oracles.

For the purposes of exhibiting two-dimensional, zero-entropy SFTs in each Π^0_1 Medvedev degree, it will be necessary to synchronize the oracles as in [25].

Analogous to Proposition 2.3.4 we have

Proposition 2.3.5. *Given a Turing machine T, there exists an SFT Y_T superimposed over B such that the following are equivalent:*

1. *$(x, r) \in B$ is represented in Y_T.*

2. *For each finite board induced by r and containing the symbol corner-slash, the sequences of 0's and 1's induced by the $\updownarrow$'s and $\Updownarrow$'s of x on the boards are all compatible, and their union is an oracle O on which T does not halt; and if there is an infinite board induced by r containing corner-slash, T does not halt on the sequence of 0's and 1's induced by O_v on it.*

Furthermore, $\mathrm{ent}(Y_T) = 0$ *and for any oracle O on which T runs forever, there is $(x, r) \in B$ represented in Y_T such that an initial segment of O is induced by x on each finite board induced by r.*

Proof. First note that in any element $z \in B$, the columns and rows are naturally partitioned into "0-board columns", "1-board column", etc., and similarly for rows. Formally, a column of z is an n-board column if n is the least number such that $s^n(\blacksquare)$ appears as a subword of z that intersects the given column in z. A parallel criterion defines an n-board row. See Figure 2.1.

The n-board columns of $z \in B$ are exactly the columns which induce the oracle of the n-boards.

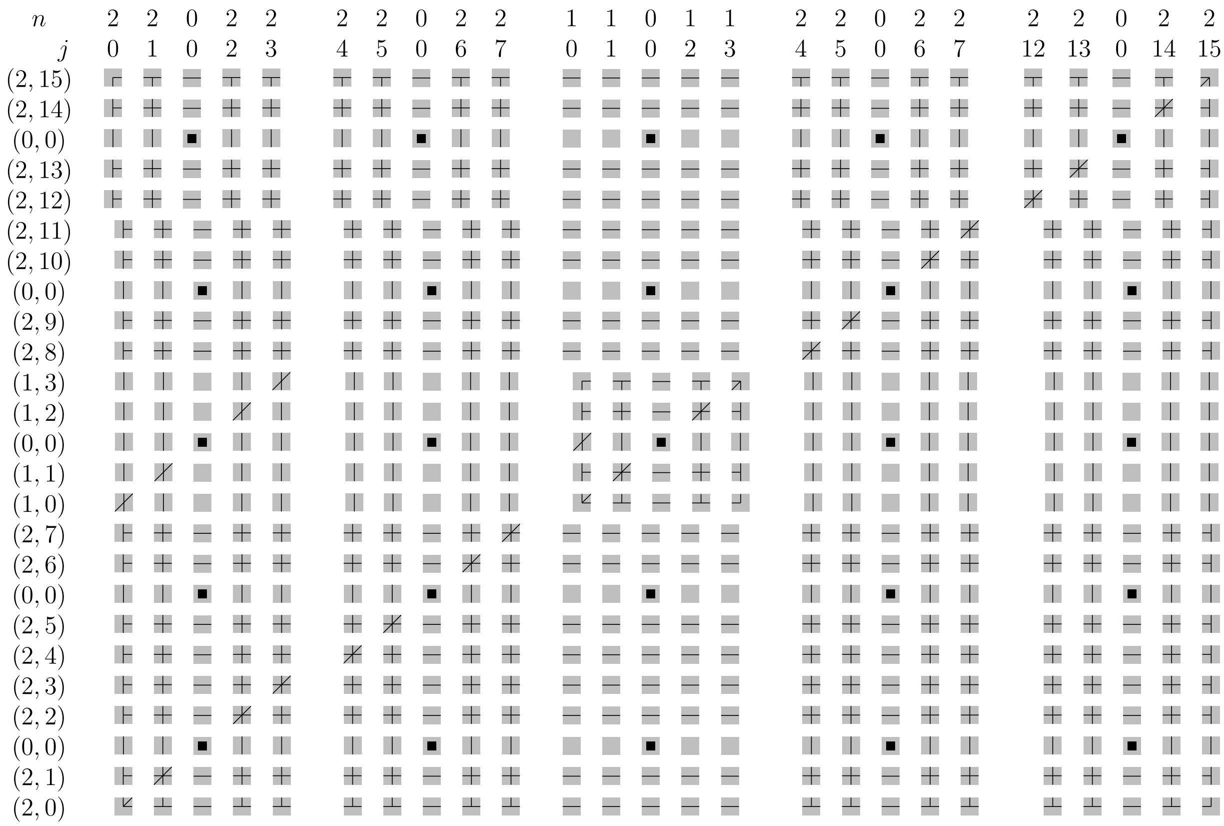
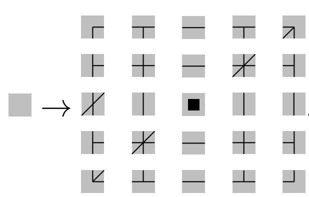
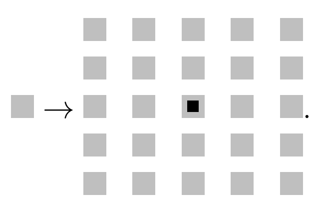


Figure 2.1: The tiles of $s^2(■)$ with the (n, j)-rows and -columns labeled.

Notice, the subshift $\widetilde{R}$ of Hochman and Meyerovitch does not have the property that the columns used by n-boards for their oracle are distinct from the columns used by m-boards for their oracle when $m \neq n$, because they used



Instead, we use



This property can be proved by induction on n, using the fact that $s^k(■)$ appears with the period s^k in both dimensions in $s^n(■)$ for $n > k$ and $s^k(■)$ appears nowhere else in $s^n(■)$

(the latter is not true in [13] because of the difference pictured above).

The columns and rows can be further subdivided into the (n, j) columns and (n, j) rows, where n indicates the board size to which the column/row belongs, and j indicates how many n-columns or n-rows appear to the left of or below the given column or row in $s^n(\blacksquare)$. So for example, when $n = 1$, j ranges from 0 to 3. One may verify by induction on n that the choice of j is well-defined. See Figure 2.1.

Let O_h be the shift on $\leftrightarrow$, $\Leftrightarrow$ which is just O_v rotated 90 degrees. Consider the shift $\hat{B}$ which is a subset of the product shift $O_h \times B$, adding the additional constraint that on ↙, ⋇, ⟋ and $\blacksquare$, the column and row markings from O_h and O_v must agree, either $\updownarrow$ and $\leftrightarrow$, or $\Updownarrow$ and $\Leftrightarrow$. It is immediate that $\hat{B}$ has zero entropy.

The constraints imposed by ↙, ⋇, $\blacksquare$ and ⟋ fall into two categories. First we consider the effect of the ↙, ⋇, and $\blacksquare$. By induction on the size of $s^n(\blacksquare)$ needed to identify the (n, j) label of the column or row, if one of these tiles occurs on an (n, j) column, then it must also occur in an (n, j) row and vice versa. Furthermore, every (n, j) column and every (n, j) row intersect in such a tile. Therefore, the ↙, ⋇, and $\blacksquare$ tiles link exactly the (n, j) rows and columns together, so that there is one (n, j)-label (either 0 or 1, as represented by $\updownarrow$ or $\Updownarrow$) which is shared by all (n, j) rows and columns. And the ↙'s, ⋇'s and $\blacksquare$'s place no other restrictions on rows and columns of finite boards.

The second kind of constraint is a ⟋. Again by induction one may show that when a ⟋ occurs in a (n, j) row, its column is an $(n+1, j)$-column. And for every n and $j < 4^n$, there is at least one place where this happens. Therefore, the effect of the ⟋'s is to link the values carried by the (n, j) and (m, j) rows and columns for any finite n and m. The ⟋ may also exist in an infinite-board row (respectively column) but in that case it must also exist in an infinite-board column (respectively row). Thus it imposes no restraint on the finite rows and columns.

Now for any $O \in 2^{\mathbb{N}}$, there is $\hat{B}$ such that $O(j)$ and the markings on each (n, j) column and row agree. And conversely, in any element of $\hat{B}$, the sequences of 0's and 1's induced on each finite board are all compatible. Then by applying Proposition 2.3.4 to $\hat{B}$, we obtain Y_T satisfying the equivalences and $\mathrm{ent}(Y_T) = 0$.

Finally, if O is an oracle on which T runs forever, let $z \in \hat{B}$ be such that $O(j)$ agrees with the common marking on all the (n, j) columns and rows, and such that there are no infinite rows. Then by the equivalence just proved, z is represented in Y_T. □

Now we are ready to re-prove Simpson's result using an SFT with zero entropy.

Proposition 2.3.6. *Given any Π^0_1 class P, there is a two-dimensional SFT with zero entropy that is Medvedev equivalent to P.*

Proof. Let a Π^0_1 class P be given. Let T be the Turing machine that halts if its oracle leaves P_0.

Let $M_P = Y_T$ from Proposition 2.3.5. We claim that M_P is Medvedev equivalent to P. Given an arbitrary $x \in Y_T$, the common oracle O for the finite boards may be uniformly read off of it, and by Proposition 2.3.5 T does not halt on O, so $O \in P$. On the other hand, given any element $O \in P$, one may compute a tiling which has O as the common oracle of the finite boards and has no infinite boards.

Thus we have assured the existence of a two-dimensional shift of finite type with zero entropy in any Π^0_1 Medvedev degree. $\square$

2.4 Entropy and effective dimension

In this section we introduce a theorem of Simpson relating effective dimension and entropy of subshifts, and we give an elementary proof of a special case. We also discuss shift-complex sequences and what they reveal about the entropy of related subshifts.

An element of maximal dimension

The entropy of a subshift is related to effective dimension through the following theorem of Simpson [32]. Recall that $\dim(x)$ is the effective dimension of x. If $x \in A^{\mathbb{Z}^d}$ then $\dim(x) = \lim_{n\to\infty} \frac{K(x\upharpoonright n)}{(2n-1)^d}$ and similarly for $G = \mathbb{N}$.

Theorem 14 (Simpson). *Let $X \subseteq A^G$ be a subshift, where $G = \mathbb{N}^d$ or $G = \mathbb{Z}^d$. Then $\{\dim x : x \in X\}$ has a maximum element, and $\max\{\dim x : x \in X\} = \mathrm{ent}(X)$.*

The proof required the use of measure-theoretic entropy. Simpson asked whether there as a more elementary proof. Here we note that there is an elementary proof in the case $G = \mathbb{N}$, using an argument very similar to the ones Furstenburg had used to prove that the entropy of a subshift on $A^{\mathbb{N}}$ was equal to its Hausdorff dimension.

Simpson had considered a general measure of complexity, speaking simultaneously both of prefix free and plain Kolmogorov complexity, and some other variants as well. In the below we assume prefix-free Kolmogorov complexity, but because all the universal measures of complexity differ from each other by at most a logarithmic factor, the effective dimension calculated by all of them is the same. We will use the fact that $\mathrm{ent}(X)$ is an upper bound on the effective dimensions (in fact, on the packing dimensions) of $x \in X$.

Lemma 2.4.1 (Simpson). *Let $X \subseteq A^G$ be a subshift. Then for any $x \in X$,*

$$\limsup_{n\to\infty} \frac{K(x \upharpoonright n)}{|x \upharpoonright n|} \le \mathrm{ent}(X).$$

Now here is the alternate proof of Simpson's result for the case $G = \mathbb{N}$.

Proposition 2.4.2 (Simpson). *If $X \subseteq 2^{\mathbb{N}}$ is a subshift, then there is an $x \in X$ for which $\dim(x) = \mathrm{ent}(X)$.*

Proof. The idea is similar to Furstenburg's proof that $\mathrm{ent}(X) = \dim(X)$ for $X \subseteq A^{\mathbb{N}}$. First, any $x \in X$ with the property that $K(x \upharpoonright n)/n \ge \mathrm{ent}(X)$ for all n also satisfies $\lim_{n\to\infty} K(x \upharpoonright n)/n = \mathrm{ent}(X)$, by Lemma 2.4.1. Now for contradiction, assume there is no x with this property. Then there is a finite set $I \subset A^{<\mathbb{N}}$ such that $\cup_{\sigma\in I}[\sigma]$ covers X, and for each $\sigma \in I$, $K(\sigma)/|\sigma| < \mathrm{ent}(X)$. Since I is finite, let $s < \mathrm{ent}(X)$ be such that $\frac{K(\sigma)}{|\sigma|} \le s$ for $\sigma \in I$. We will show $\mathrm{ent}(X) \le s$ to get the contradiction.

Let m be the maximum length of $\sigma \in I$. Fix any $z \in X$ and make a code for $z \upharpoonright n$ as follows. Using σ_i from I, write $z \upharpoonright n$ as $\sigma_1 + ... + \sigma_k$, where the latter is too long by at most

m. Then in the code, give first a prefix free description of n (so that the universal machine knows when to stop reading) and then, in order, the optimal prefix free codes for $\sigma_1, ..., \sigma_k$. The universal machine can recover z from this information by reading the codes for σ_i and appending them until it has read enough codes to fill n bits. Calculating the complexity of this code:

$$\begin{aligned} K(z \upharpoonright n) &< C + \log n + K(\sigma_1) + ... + K(\sigma_k) \\ &\leq C + \log n + s(|\sigma_1| + ... + |\sigma_k|) \\ &< C + \log n + s(n + m). \end{aligned}$$

This bound holds for any $z \in X$. Therefore, there are at most $2^{C+\log n+s(n+m)}$ permitted strings of length n in X. Thus

$$\lim_{n\to\infty} \frac{\log(|\{x \upharpoonright n : x \in X\}|)}{n} \leq \lim_{n\to\infty} \frac{C + \log n + s(n+m)}{n} = s.$$

□

An alternate proof for the multidimensional case would complete the picture.

Shift Complex Subshifts

A nice family of example subshifts comes from the notion of shift-complex sequences, which were introduced in [5]. A good survey is [15].

Definition 2.4.1. *An $x \in 2^{\mathbb{N}}$ is* (d, b)-shift-complex *if for any σ appearing as a subword of x, $K(\sigma) \geq d|\sigma| - b$. An $x \in 2^{\mathbb{N}}$ is* d-shift-complex *if there is a b for which x is (d, b)-shift-complex.*

The existence of a d-shift-complex sequence for any $d < 1$ was established by [5]. Miller [23] gave an alternative proof in which he showed that for each d there is a b such that the set $S_{d,b}$ of (d, b)-shift-complex sequences is a nonempty subshift. The reader may directly verify that each such set is a subshift, so the work is to show that one is nonempty. Because the effective dimension of every d-shift-complex sequence is at least d, by Theorem 14 we have d as a lower bound of the entropy of $S_{d,b}$ whenever this subshift is non-empty. In general, the entropy of this subshift depends on the universal machine used.

We will build on the method of proof of the following result:

Theorem 15 (Hirschfeldt and Kach [12])**.** *For every $d \in (0, 1)$, there is a d-shift-complex sequence with packing dimension d.*

By controlling the packing dimension, they guarantee that the effective dimension of the resulting sequence is also d.

2.5 Orthogonality of entropy and Medvedev degree

In this section we show that entropy and Medvedev degree of subshifts are quite independent by constructing subshifts of every possible combination of the two, while varying the dimension of the subshift, the number of symbols, and the presence of computability restrictions. The main difficulty is in passing from multiple symbols to two symbols.

In one dimension

Proposition 2.5.1. *For any $r \in [0, \infty)$, and any closed set P, there is a one-dimensional subshift Medvedev equivalent to P with entropy r. Furthermore, if r is right-r.e., and P is a Π^0_1 class, the subshift may be made Π^0_1. If $r \in [0, 1)$, the subshift may be found in 2^G, where $G = \mathbb{N}$ or $\mathbb{Z}$.*

If one does not include the restriction that the subshift must exist on two symbols if $r \in [0, 1)$, the proof is much simpler, so we give that first.

Proposition 2.5.2. *For any $r \in [0, \infty)$ and any closed set $P \subseteq 2^\omega$, there is a one-dimensional subshift Medvedev equivalent to P withe entropy r.*

Proof. By taking the product with as many full shifts as needed (each contributes entropy 1), we may assume that $r \in [0, 1)$. By Definition 2.3.1 and Proposition 2.3.2, for every closed set P there is a zero-entropy subshift M_P in $2^{\mathbb{N}}$ Medvedev equivalent to P. On the other hand, for every $r \in [0, 1)$, the density-r subshift has entropy r. By taking the product $M_P \times X_r$ and noting that the density-r subshift does contain a computable element, we obtain a subshift of entropy r which is Medvedev equivalent to P. □

The difficulty in repeating the above using only two symbols is that a product of subshifts results in a minimum of four symbols. So we cannot make the Medvedev portion and the entropy portion of our subshift completely independent anymore. Instead, we reserve some small-density fraction of the bits in the subshift for Medvedev-related encoding, and use the rest for entropy. We interleave a subshift M_P with a subshift $X_{r'}$, where M_P has the appropriate Medvedev degree and $r' > r$ is precisely the right size to compensate for the loss of entropy when we reserve some density of bits for encoding M_P.

Note that when one is restricted to two symbols, the only subshift with entropy 1 is the full shift, so the Medvedev degree can only be zero for that entropy. This is why the restriction to $r \in [0, 1)$ is necessary.

Proof of Proposition 2.5.1. Let $r \in [0, 1)$ be given. Fix N odd and large enough that $\frac{N}{N-1} r < 1$. Let $r' = \frac{N}{N-1} r$. We build a subshift X with the idea that its trajectories x should be of the form

$$x(j) = \begin{cases} m(\frac{j-l}{N}) & \text{if } j \equiv l \mod N \\ y(j) & \text{otherwise} \end{cases}$$

where $l < N$ is fixed but arbitrary, y is a trajectory from the density-r' subshift $X_{r'}$, and m is a trajectory from a subshift M_P defined as in 2.3.1. Because of the choice of the subshift from which m is drawn, it will be possible to compute, uniformly in x, the value l. Because m has no information content, its presence overwriting bits of y will decrease the information content of x, and we will show that this decrease is exactly by a factor of $\frac{N-1}{N}$. Now we give a formal definition of the subshift X.

Let M_P be a subshift defined as in 2.3.1, using $a_\lambda = 0$ and $b_\lambda = 1^{2^{i_0}}$ where i_0 is least such that $\tau_{i_0} = 0$ in the decomposition $r' = \sum_{i=1}^{\infty} \tau_i 2^{-i}$ (choose the decomposition ending in zeros has been chosen if r' is a dyadic rational). The purpose of making long strings of 1s is to be able to later identify which bits came from M_P.

Let X be the subshift defined by the following set of forbidden strings. Given a string σ, consider for each $l < N$ the subsampled string $\sigma_m = \sigma(l)\sigma(l+N)\sigma(l+2N)\ldots\sigma(l+pN)$ and its "background" $\sigma_y = \sigma[0,l]0\sigma[l+1,l+N]0\sigma[l+N+1,l+2N]0\ldots0\sigma[l+pN+1,|\sigma|_1]$. We say the alignment l fails if σ_m is forbidden in M_P or σ_y is forbidden in $X_{r'}$. If all of the alignments fail, we forbid σ.

Note that if r is right-r.e., then r' is as well, and if P is a Π^0_1 class, then X is also Π^0_1.

First we check the entropy of X. The key observation is that because the chosen N is odd, it is relatively prime to 2^i for all i. Proceeding as in the proof of the entropy of X_r, let us first compute a lower bound on the entropy by considering only those σ which are permitted with

1. Alignment $l = 0$
2. σ_y belongs to $X_{r'}$ with shift $k = 0$ and $\tau_i = 0$ implies $\sigma_y(j) = 0$ for all $j \equiv 2^{i-1} (\mod 2^i)$.
3. Anything can happen in σ_y on bits not constrained by the above, or by the $\sigma_y(bN) = 0$ restriction.

If it were not for the fact that $\sigma_y(bN) = 0$ for this alignment, the number of free bits in σ_y would take up approximately r' of the length of σ_y. To take that additional restriction into account, first note that the free bits $f(n)$ in a string of length n and shift 0 before that restriction may be broken down as

$$f(n) = \sum_{i<\infty:\tau_i=1} |\{j < n : j \equiv 2^{i-1} (\mod 2^i)\}|,$$

and $\lim_{n\to\infty} \frac{f(n)}{n} = r'$. Since N and 2^i are relatively prime,the set of $j \equiv 0 \mod N$ has density $\frac{1}{N}$ in $\{j : j \equiv 2^{i-1} (\mod 2^i)\}$. So the number of actual free bits $f'(n)$ on a string of length n under the additional restriction $\sigma(bN) = 0$ satisfies

$$\begin{aligned} f'(n) &= \sum_{i<\infty:\tau_i=1} |\{j < n : j \equiv 2^{i-1} (\mod 2^i) \text{ and } j \not\equiv 0 \mod N\}| \\ &\leq \sum_{i<\lceil \log n\rceil:\tau_i=1} \left(\frac{N-1}{N}|\{j < n : j \equiv 2^{i-1}(\mod 2^i)\}| + 1\right) \\ &= \frac{N-1}{N} f(n) + \lceil \log n \rceil. \end{aligned}$$

Therefore,

$$\begin{aligned} \text{ent}(X) &= \lim_{n\to\infty} \frac{\log|X \upharpoonright n|}{n} \\ &\leq \lim_{n\to\infty} \frac{\log|M_P \upharpoonright \frac{n}{N}| + \log 2^{f'(n)}}{n} \\ &= \lim_{n\to\infty} N\frac{\log|M_P \upharpoonright \frac{n}{N}|}{\frac{n}{N}} + \frac{\frac{N-1}{N}f(n) + \log n}{n} \\ &= 0 + \frac{N-1}{N} r' = r \end{aligned}$$

On the other hand, if we overcount by considering

1. All values of $l < N$
2. All values of k which result in a distinct partitioning of the bits (at most $2n$ values of k)

define $g'(n)$ as the maximum number of free bits a string of length n with any alignment l and shift k can have under $\sum_{i=1}^{\infty} \tau_i 2^{-i}$, then $|X \upharpoonright n| \leq |M_P \upharpoonright \frac{n}{N}| \cdot 2^{g'(n)} \cdot N \cdot 2n$ where $\lim_{n\to\infty} \frac{g'(n)}{n} = \frac{N-1}{N} r' = r$. So

$$\text{ent}(X) \leq \frac{1}{N}\text{ent}(M_P) + r = r.$$

Finally, we must verify the Medvedev equivalence. Since $0^\omega \in X_{r'}$, any element of M_P computes an element of X by just placing the bits of a trajectory of M_P on every Nth bit and zeros otherwise. On the other hand, given a trajectory $x \in X$ we may uniformly find an l which allows us to recover a trajectory $m \in M_P$ as follows. The idea is that due to the very long strings of 1s in the starting words for M_P, for a sufficiently long σ there will be only one l that could possibly be the alignment of the m part. Again because N and 2^i are relatively prime, any sufficiently long σ will have at least 2^{i_0}-many 1s in a row on its m part. Regardless of the choice of k, these 1s will appear at every residue mod 2^{i_0}. Because $\tau_{i_0} = 0$, these 1s would invalidate the string were they niot to be blanked out by a unique selection of l. Therefore, there is an effective algorithm to find l, and from there the embedded $m \in M_P$ is easily retrieved. □

In multiple dimensions

Multidimensional shifts of finite type, must have right-r.e. entropy and they must be in a Π^0_1 Medvedev degree. In this section we show that within those restrictions, the entropy and Medvedev degree are independent. Furthermore, any combination of entropy and Medvedev degree that is not immediately prohibited by the use of only two symbols can be obtained using only two symbols.

As in the one-dimensional case, there is a direct product-based construction using arbitrarily many symbols, while the reduction to only two symbols uses a division of the bits into a low-density computation portion and a high-density entropy portion.

Proposition 2.5.3. *Given any right-r.e. entropy and any Π^0_1 Medvedev degree, there is a 2-dimensional shift of finite type which lies in the given Medvedev degree and has the given entropy.*

Proof. The strategy is just as in the many-symbol one-dimensional case – we take the product of a two-dimensional SFT with the given entropy (guaranteed by [13]) with a two-dimensional SFT of zero entropy with the appropriate Medvedev degree guaranteed by Proposition 2.3.6. Then just as before, because there is a computable element of each subshift constructed by [13], we will have the desired result. □

Next we show how to modify the construction of Hochman and Meyerovitch in order to make 2-dimensional SFTs of every possible right-r.e. entropy using only two symbols. As in the one-dimensional case, the two symbol restriction means that the entropy must be in $[0, 1]$, and the only way to have the entropy equal to 1 is to use the full shift. Aside from these basic restrictions, everything else is possible.

Proposition 2.5.4. *Given any right-r.e. $r \in [0, 1]$, there is an SFT on $2^{\mathbb{Z}^2}$ with entropy r.*

Proof. If $r = 1$, the full shift suffices. Otherwise, the key to the proof is the observation that if the entropy is to be less than 1, then it will be possible to reserve some tiny but definite fraction of the total area for computation. This computation makes no contribution to the entropy in itself, but it controls the rest of the space, which is used for entropy.

For convenience of visualization we will speak of white tiles and black tiles instead of the symbols $\{0, 1\}$.

We impose the following constraints on the geometric structure of the elements of our SFT. There will be a grid, formed mostly of white tiles, made with horizontal and vertical bands of width a (where a is a number to be determined later). These bands will be spaced far from each other, so that the area not used by the white grid is divided into an infinite array of square regions of side length b (another number to be determined later). Within the white grid, at the place where a horizontal and vertical white band intersect, there is an $(a-2) \times (a-2)$ square at the center which may take any one of its $2^{(a-2)^2}$ configurations. This coding square is centered, surrounded by an $a \times a$ ring of white tiles. Turning now to the $b \times b$ square, it is required that the edges of the square consist of black tiles. In this

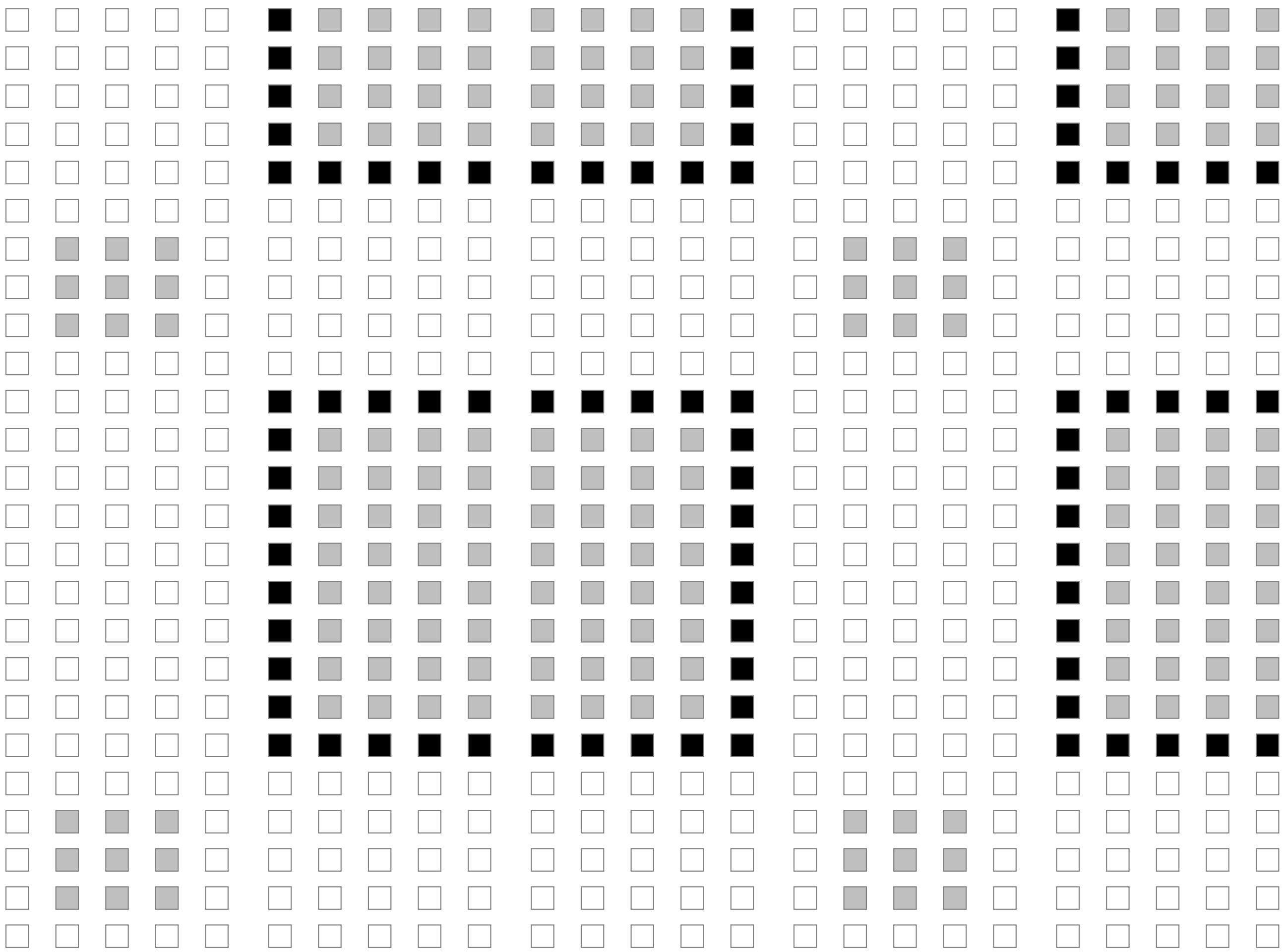


Figure 2.2: A permitted word in the white grid SFT with $a = 5$ and $b = 10$. Any combination of white and black tiles are permissible for the tiles pictured gray here.

way there can be no ambiguity about the location of the grid and the coding squares. See Figure 2.2. One may verify that these geometric restrictions may be accomplished with a finite number of local restrictions.

Now, in any square region of side length $b + a$, the area taken by tiles belonging to the interior of the $b \times b$ square is $(b - 2)^2$, while the rest of the area is taken by tiles belonging to the white grid and black outline. The interior of the $b \times b$ square will be the place where entropy is encoded, so we must verify that it make take up as great a percentage of the total area as is needed. But this is clear because its percentage of the total area is

$$\frac{(b-2)^2}{(b+a)^2}$$

and so regardless of a, this ratio can be made arbitrarily close to 1 by an appropriate increase of b.

Now returning strictly to the context of [13], consider a Turing machine T which has two input tapes and does the following.

1. It computes over time better and better approximations to a target right-r.e. number $r \in [0, 1)$.

2. It reads from one of its input tapes codes for two numbers a and b.

3. It proceeds computing a new right-r.e. number $r' = \frac{(b+a)^2}{(b-2)^2} r$.

4. Exactly as in [13], it halts if the evidence of its second oracle indicates that the density of the 1's exceeds r' in the element of X which induced the second oracle, where X is the SFT from [13] Section 6.

Now apply Propositions 2.3.4 and 2.3.5 to obtain Y_T with the following properties.

1. $\mathrm{ent}(Y_T) = 0$.

2. The first oracle, n the manner of Proposition 2.3.5, is synchronized across all finite boards.

3. The second oracle, in the manner of Proposition 2.3.4, is induced by an unsynchronized $x \in O_v$, and in general the induced second oracles on the finite boards may not be compatible.

Let X be the SFT defined in Section 6 of [13]. Let $Z \subseteq Y_T \times X$ be the SFT $Y_T \times X$ under the additional constraint that the $\updownarrow$ and $\Updownarrow$ markings which induce the second oracle on the boards of Y_T must correspond to the 0 and 1 markings in X, respectively. The entropy of Z is zero, and for every $x \in Z$, if its common first oracle encodes a, b, then the density of 1's in its X part is at most $\frac{(b+a)^2}{(b-2)^2} r$. This follows form the definition of the Turing machine T to interact appropriately with X as described in [13].

Fix a large enough that the total number of symbols in the alphabet of Z is at most $2^{(a-2)^2}$. Associate to each symbol in the alphabet of Z a unique $(a-2) \times (a-2)$ square of black and white tiles, which will serve as the code for that tile. Let b be large enough that the fraction of the total area in the interior of the $b \times b$ squares is at least r. Note that this also guarantees that $r' < 1$.

Returning to the original grid of black and white tiles, now we add the following additional restrictions, which one may verify consist of finitely many local restrictions. Call the resulting SFT W.

1. The only codes which can appear in the coding locations of the white grid are codes from the alphabet of Z.

2. The first oracle (whose values can be read in the $(a-2)\times(a-2)$ square codes) must encode the same numbers a and b which were fixed above. Because the code for a, b is finite, and the amount of space from an $(n, 0)$ column to an (n, j) column is dependent only on j, not n, this can be accomplished with finitely many local restrictions.

3. A local configuration of adjacent (in the sense of the white grid) codes is forbidden in the grid if the symbols associated to the codes were forbidden in that configuration in Z.

4. Whenever a code refers to a symbol whose X part is marked with 0, all the tiles of the $b \times b$ square directly up-right of the code tile must be black.

If the code refers to a symbol whose X part is marked with 1, then there is no additional restriction on the associated $b \times b$ square, so all $2^{(b-2)^2}$ configurations are permitted.

One may see that the elements of Z in which the first input tape contains "a, b" are in computable one-to-one correspondence with the configurations of the white grid in W. Since Z had zero entropy, the configuration of the white grid contributes nothing to the entropy of W. Because the location of the free bits is determined entirely by the configuration of the white grid, we may estimate $|W \restriction n| = |\{u \in 2^{n^2} : u \text{ may be extended to } x \in W\}|$ as follows. A free bit is a bit contained in the interior of a $b \times b$ square whose associated code has 1 in its X part. Let $f(n)$ be the greatest number of free bits in any $n \times n$ square of any $x \in W$. There are $x \in W$ or which the density of 1's in the X part of the corresponding $z \in Z$ is r', so $\lim_{n\to\infty} \frac{f(n)}{n^2} = \frac{(b-2)^2}{(b+a)^2} r' = r$.

For a lower bound on $|W \restriction n|$, consider that if some $x \in W$ has a pattern u whose white grid configuration permits $f(n)$ free bits, then u could be replaced with any one of $2^{f(n)}$ variations of itself which vary the free bits in all possible ways and the result would still be an element of W. Therefore,

$$\lim_{n\to\infty} \frac{\log |W \restriction n|}{n^2} \leq \frac{2^{f(n)}}{n^2} = r.$$

For an upper bound, it suffices to consider all the ways the white grid could be aligned in the $n \times n$ region, which is $(a+b)^2$ ways; all the possible configurations of the white grid that size, which is $|Z \restriction \frac{n}{(a+b))}|$; and for each configuration of the white grid, the maximum number of distinct choices of its free bits, which is at most $2^{f(n)}$.

Thus,

$$|W \restriction n| \leq (a+b)^2 \cdot |Z \restriction \frac{n}{a+b}| \cdot 2^{f(n)}$$
$$\frac{|W \restriction n|}{n^2} \leq \frac{\log(a+b)^2}{n^2} + \frac{\log |Z \restriction \frac{n}{a+b}|}{n^2} + \frac{f(n)}{n^2}$$

and by taking limits in the above, $\text{ent}(W) = r$.

□

Of course, these two ideas may be combined to create an SFT on two symbols with any desired combination of Medvedev degree and entropy in $[0,1)$.

Corollary 2.5.5. *For any Π^0_1 class P and any right-r.e. $r \in [0,1)$, there is a two-dimensional SFT on two symbols with entropy r that is Medvedev equivalent to P.*

Proof. Replace Z in the previous proof with $Z \times M_P$, where M_P has entropy zero and has the desired Medvedev degree. A proof of the existence of M_P was described in Proposition 2.3.6. Then proceed with the construction of W as above. The entropy is determined as before, and is unaffected by the additional superimposed SFT because its entropy is also zero. From any element of the W, we may read off the oracle for the finite boards of M_P by decoding the codes in the white grid, obtaining an element of P. And from any element $O \in P$, we may uniformly compute an element of W which combines an element of M_P which has finite board oracle O, with a computable element of the original Z. □

2.6 The effective dimension spectrum

We consider the *effective dimension spectrum*:

Definition 2.6.1. *The* effective dimension spectrum *of a subshift* X *is* $\mathcal{DS}(X) = \{\dim(x) : x \in X\}$.

If $\Phi : X \to Y$ is a shift-invariant homeomorphism of subshifts, then $\dim(\Phi(x)) = \dim(x)$ for all $x \in X$. So the effective dimension spectrum is a conjugacy invariant for subshifts. Simpson [32] proved a topological restriction on $\mathcal{DS}(X)$ – it must always have a maximum element. We would like to characterize the $A \subseteq [0, 1]$ for which there exists a subshift X such that $\mathcal{DS}(X) = A$. For simplicity, we work only with the case $X \subseteq 2^{\mathbb{N}}$.

Motivating examples

In this section we relate what is known about the dimension spectrum by finding the dimension spectrum of the important examples. We also include a more general result which lets us characterize the effective dimension spectra of shifts of finite type.

Density-r subshifts

If X_s and X_t are density-r subshifts with given entropies $s < t$, then by the definition, $X_s \subseteq X_t$. Combined with Simpson's result [32] that the effective dimension spectrum of a subshift X always has a maximal element, $\mathrm{ent}(X)$, this set of inclusions guarantees that $\mathcal{DS}(X_s) = [0, s]$ for all $s \in [0, 1]$.

Shift-complex subshifts

Because any d-shift-complex sequence has effective dimension at least d, the shift-complex subshift $S_{d,b}$, when it is nonempty, satisfies $\mathcal{DS}(S_{d,b}) \subset [d, 1)$. An effective dimension of one is excluded because forbidding any string puts an upper bound on the packing dimension. The entropy of $S_{d,b}$ depends not only on d and b, but also on the universal machine used. It is not known whether the effective dimension spectrum of $S_{d,b}$ is an interval. By Theorem 15 for $d \in (0, 1)$ and for sufficiently large b, there is $x \in S_{d,b}$ with $\dim(x) = d$.

Medvedev subshifts

Because each Medvedev subshift has entropy zero, each of its elements has effective dimension zero, so the effective dimension spectrum of a Medvedev subshift is simply $\{0\}$.

Conditions under which the spectrum is [0, ent(X)]

In this section we give some sufficient conditions for the dimension spectrum to be the commonly encountered $[0, \mathrm{ent}(X)]$.

Definition 2.6.2. *A subshift $X \subseteq 2^{\mathbb{N}}$ is* computably extendible *if uniformly in every σ such that $\sigma \prec x \in X$, one may compute such an x extending σ.*

Definition 2.6.3. *A subshift $X \subseteq 2^{\mathbb{N}}$ is* uniformly full *if*

$$\lim_{n\to\infty} \min_{\sigma\prec x\in X} \frac{\log|\{x \upharpoonright n : \sigma x \in X\}|}{n} = \mathrm{ent}(X).$$

Theorem 16. *If a subshift $X \in 2^{\mathbb{N}}$ is computably extendible and uniformly full, then $\mathcal{DS}(X) = [0, \mathrm{ent}(X)]$*

Proof. The proof is very similar to Hirschfeldt and Kach's construction [12] of a d-shift-complex sequence with packing dimension d for a given $d \in (0,1)$. The difference is that we run the construction inside the subshift. Though we only need sequences of effective dimension r for each $r \in (0, \mathrm{ent}(X))$, we actually construct r-shift-complex sequences with packing dimension r.

Because X is computably extendible, it has a computable element, so $0 \in \mathcal{DS}(X)$. By [32], $\mathrm{ent}(X) \in \mathcal{DS}(X)$. Let $r \in (0, \mathrm{ent}(X))$ be given. We will build $x \in X$ in stages, at each stage appending a string of length m, where m will be chosen later. At stage n, given $x \upharpoonright mn$, if $K(x \upharpoonright mn)/mn \leq r$, we choose τ so that $K(\tau : x \upharpoonright mn)$ is as large as possible; otherwise, we extend $x \upharpoonright mn$ computably.

Now to specify the choice of m and check that this construction works.

First we will check that there is a constant D such that for all n, $rmn - D \leq K(x \upharpoonright mn) \leq rmn + D$.

There is some inertia to Kolmogorov complexity. If $K(\sigma) \leq r|\sigma|$ and $K(\sigma^\frown\tau) > r|\sigma^\frown\tau|$, then $K(\sigma^\frown\tau) - r|\sigma^\frown\tau|$ can be bounded by a constant:

$$\begin{aligned} K(\sigma^\frown\tau) - r|\sigma^\frown\tau| &< K(\sigma) + K(\tau) + C - r|\sigma^\frown\tau| \\ &\leq r|\sigma| + m + 2\log m + 2C - r|\sigma^\frown\tau| \\ &= m + 2\log m + 2C - rm \\ &= (1-r)m + 2\log m + 2C \end{aligned}$$

Similarly, if $K(\sigma) > r|\sigma|$ and $K(\sigma^\frown\tau) \leq r|\sigma^\frown\tau|$,

$$\begin{aligned} r|\sigma^\frown\tau| - K(\sigma^\frown\tau) &< r|\sigma| + rm - K(\sigma) + K(|\tau|) + C \\ &< rm + 2\log m + 2C \end{aligned}$$

where we used the fact that $K(\sigma) < K(\sigma^\frown\tau) + K(|\tau|) + C$. So, if there is any trouble, it comes from the situation where our attempts to change the compressive ratio actually move it in the wrong direction.

When we are trying to get the ratio to increase, in fact we will always be able to find τ such that $K(\sigma\tau) - K(\sigma) \geq rm$. If one takes this fact as given (it will be justified later) then one sees that when trying to go up, one can never go down by accident, and thus in fact $rmn - (rm + 2\log m + C) \leq K(x \upharpoonright mn)$ for all n.

However, when we are trying to get the ratio to decrease by adding τ according to an algorithm, it may be possible to get unlucky and actually increase the ratio for some period of stages before succeeding in bringing the ratio down. We can quantify this, assuming that initially $r|\sigma| < K(\sigma) < r|\sigma| + ((1-r)m + 2\log m + 2C)$, since this is the situation we will be in when first trying to decrease the ratio. Then, letting k be the number of iterations of continuing to extend the string algorithmically, that is $|\tau| = km$, we have

$$\begin{aligned} K(\sigma\tau) - r|\sigma\tau| &\leq K(\sigma) + K(\tau : \sigma^*) + C - r|\sigma| - rkm \\ &\leq K(\tau : \sigma^*) + C - rkm + (1-r)m + 2\log m + 2C \\ &\leq 2\log km + 2C - rkm + (1-r)m + 2\log m + 2C \\ &= -rkm + 2\log k + (1-r)m + 4(\log m + C) \end{aligned}$$

Since only the first two terms vary with k, and the negative term dominates, there is some k^* (which does not depend on $|\sigma|$) such that within k^* iterations of choosing τ algorithmically, $K(\sigma\tau) < r|\sigma\tau|$ will be achieved, and there is some D which upper bounds the values the right hand side can take as k increases from 0. Thus we have shown that always

$$rmn - D \leq K(x \upharpoonright mn) \leq rmn + D$$

pending justification for the assertion that when we wish the ratio to go up, we may in fact choose a τ such that $K(\sigma\tau) - K(\sigma) \geq rm$.

Now we make use of uniform fullness. Since $r < \text{ent}(X)$, fix t with $r < t < \text{ent}(X)$, and let m be long enough that for all σ,

$$\frac{\log |x \upharpoonright m : \sigma x \in X|}{m} > t.$$

By counting, there must be some τ such that $K(\tau : \sigma^*) > tm$. Let m also be long enough that $tm - 2\log m - 3C > rm$. Then, by proposition 2.1.5, this τ also satisfies $K(\sigma\tau) - K(\tau) \geq rm$.

Thus the bound $K(x \upharpoonright mn) = rmn \pm D$ holds, and $\dim(x) = \text{Dim}(x) = r$, as desired.

To see that x is also r-shift-complex, consider first σ of the form $x[am, bm]$, that is, a σ whose boundaries are multiples of m. Then it follows from $K(x \upharpoonright bm) \leq K(x \upharpoonright am) + K(\sigma) + C$ that

$$\begin{aligned} K(\sigma) &\geq K(x \upharpoonright bm) - K(x \upharpoonright am) - C \\ &\geq rbm - ram - C - 2D \\ &= r|\sigma| - (C + 2D) \end{aligned}$$

Now consider an arbitrary substring σ of x, whose boundaries may lie anywhere. So $\sigma = \sigma_b\sigma_m\sigma_e$, where σ_m is aligned and $|\sigma_b|, |\sigma_e| < m$. Then we have

$$\begin{aligned} K(\sigma) + K(|\sigma_b|) + K(|\sigma_e|) + C &\geq K(\sigma_b, \sigma_m, \sigma_e) \\ &\geq K(\sigma_m) - C \\ K(\sigma) &\geq K(\sigma_m) - K(|\sigma_b|) - K(|\sigma_e|) - 2C \\ &\geq r|\sigma_m| - (C + 2D) - 4(\log m + C) \\ &\geq r|\sigma| - (2rm + 2D + 4\log m + 5C) \end{aligned}$$

Thus x is (r, b)-shift-complex, where $b = 2rm + 2D + 4\log m + 5C$. □

Shifts of finite type

The theorem of the previous section may be applied to characterize the effective dimension spectra of shifts of finite type. All SFTs are computably extendible, but not all are uniformly full. For example, the SFT which forbids $\{10100, 01011\}$ has positive entropy, but the string $\sigma = 0101$ has only one infinite extension. However, the following can be proved using techniques standard to the analysis of SFTs.

Proposition 2.6.1. *Let $X \leq 2^{\mathbb{N}}$ be a SFT with entropy r. Then X contains an SFT with entropy r which is uniformly full.*

Proof. Consider $X^{[n]}$, the nth higher block shift of X, for some n longer than the longest forbidden word of X. Let $B \subseteq 2^n$ (here 2^n refers to the set of binary sequences of length n) be the collection of permitted words from 2^n; these are the symbols of $X^{[n]}$. For each pair of symbols $a, b \in B$, say that $a \equiv b$ if and only if there are words u and v made of the symbols of B such that aub and bva appear as subwords of some elements $x, y \in X^{[n]}$. This relation is transitive and symmetric, but it may not be reflexive. Restrict the relation only to those $a \in B$ such that $a \equiv a$ and the result is an equivalence relation. It is nonempty because $X^{[n]}$ is non-empty and there are only finitely many symbols in B, so at least one of them must occur infinitely often in some $x \in X^{[n]}$.

In fact, given any $x \in X^{[n]}$, the set of $\{a \in B : a \text{ appears infinitely often in } x\}$ is contained in some equivalence class. Fix $x \in X^{[n]}$ to be an element such that $\dim(x) = \text{ent}(X)$, such an element being guaranteed by [32]. Let $U \subseteq B$ be the set of symbols which appear infinitely often in x. Let $Y^{[n]} \subseteq X^{[n]}$ be the subshift of $X^{[n]}$ all of whose elements use only symbols of U. Then $Y^{[n]}$ is non-empty because some tail of x is in $Y^{[n]}$.

The entropy of $Y^{[n]}$ is equal to $\text{ent}(X)$ because the tail of x has the same effective dimension as x, and this is the largest possible entropy. Finally, $Y^{[n]}$ is naturally conjugate to a shift $Y \subseteq X$ via the map $\{x_i^j\}_{i<n} \mapsto x_0^j$. Because being an SFT is invariant under conjugation, Y is an SFT (Proposition 2.1.6). Note that $Y^{[n]}$ actually is the nth higher block shift of Y, justifying the notation.

It remains to show that Y is uniformly full. We work in $Y^{[n]}$ because it is easier. Fix $a \in U$. Because $Y^{[n]}$ is a one-step SFT, starting from any σ there is a uniform bound on the number of symbols needed in a word u so that σua is valid. From that point on, anything that may follow a may follow σua. A tail of x may follow a, so $Y^{[n]}$ is uniformly full. □

Thus we have

Corollary 2.6.2. *If a subshift $X \subseteq 2^{\mathbb{N}}$ is an SFT, then $\mathcal{DS}(X) = [0, \text{ent}(X)]$.*

Shifts with gaps in the dimension spectrum

All the examples of subshifts we have seen so far have dimension spectrum that either is a closed interval, or, in the case of the shift-complex subshifts, is consistent with being a closed interval. There are shifts whose effective dimension spectra have gaps in them, however. For $s < d$ and appropriately large b, the subshift $X = X_s \cup S_{d,b}$ satisfies

$$[0, s] \cup \{d\} \subseteq \mathcal{DS}(X) \subseteq [0, s] \cup [d, 1).$$

However, this example is unsatisfactory, being just the union of two subshifts with non-overlapping spectrum. We address this issue further in the next section by considering minimal subshifts, those which have no subshift as a proper subset.

Minimal subshifts

Definition 2.6.4. *A subshift is* minimal *if it has no proper sub-subshifts.*

Solving the characterization problem for the dimension spectra of minimal subshifts will not solve it for general subshifts because not every $x \in 2^{\mathbb{N}}$ is contained in a minimal subshift. However, considering this simplification allows us to avoid the issue of simply taking the union of two subshifts in order to form an interesting dimension spectrum.

Minimality allows us to prove

Proposition 2.6.3. *If X is minimal, $\mathcal{DS}(X)$ has a least element.*

Proof. Going in stages, construct x by finite extension, $\sigma_0 \prec \sigma_1 \prec \cdots$. We will maintain always that σ_n is a subword of some $y \in X$, and thus by minimality, that σ_n is a subword of every $y \in X$. At stage $n + 1$, find $y \in X$ so that $\dim(y) - \inf \mathcal{DS}(X) < 1/n$. Let $\sigma_{n+1} = y[k, k + m]$, where k is the start of σ_n in y, and m is long enough that

$$\frac{K(\sigma_{n+1})}{|\sigma_{n+1}|} - \dim(y) < 1/n.$$

Then

$$\frac{K(\sigma_{n+1})}{|\sigma_{n+1}|} - \inf \mathcal{DS}(X) < 2/n.$$

Therefore $\dim(x) = \inf \mathcal{DS}(X)$. □

We note that this was possible because in a minimal subshift, every element contains every word, and because it is impossible to end up with $\dim(x) < \inf \mathcal{DS}(X)$.

An example

The existence of minimal subshifts is guaranteed by Zorn's lemma, but the first explicit construction of a minimal subshift was by Grillenberger [8]. His construction was simplified to the following by Bruin [2]:

Definition 2.6.5. *Let $\mathcal{E}_1$ consist of n_1 strings of length l_1, where n_1 and l_1 will be chosen later. For $i > 1$, define $\mathcal{E}_i$ to be the set of all concatenations of permutations of the elements of $\mathcal{E}_{i-1}$. Let X be the subshift which forbids exactly the words which never appear as a subword of any σ in any E_i.*

Note that for $i > 1$, $|\mathcal{E}_i| = n_i = n_{i-1}!$ and the length of an element of $\mathcal{E}_i$ is $l_i = n_{i-1}l_{i-1}$.

Proposition 2.6.4 (Grillenberger [8], Bruin [2])**.** *X is a minimal subshift.*

Proof. If X had a proper sub-subshift Y, there would be some σ permitted in X that was not permitted in Y. Let τ and i be such that σ is a subword of $\tau \in \mathcal{E}_i$. Then σ is a subword of every word in $\mathcal{E}_k$ for $k > i$. In fact, σ occurs repeatedly in each word of such $\mathcal{E}_k$, with occurrences of σ separated by at most $2l_{i+1}$, since σ is a subword of each word of $\mathcal{E}_{i+1}$, each word of $\mathcal{E}_{i+1}$ has length l_{i+1}, and each word of $\mathcal{E}_k$ is a concatenation of words of $\mathcal{E}_{i+1}$. Therefore, Y has no permitted words longer than $2l_{i+1}$, so Y is empty. □

As noted in the original constructions, this subshift has positive entropy if n_1 and l_1 are chosen sufficiently.

Proposition 2.6.5 (Grillenberger[8], Bruin[2])**.** *The parameters n_1 and l_1 may be chosen so that X has positive entropy.*

Proof. By subsampling the convergent sequence which defines the entropy of X, we have

$$\text{ent}(X) = \lim_{i\to\infty} \frac{\log N_{l_i}}{l_i}$$

where N_k is the number of permitted strings in X of length k. For a lower bound, because $n_i < N_{l_i}$,

$$\lim_{i\to\infty} \frac{\log N_{l_i}}{l_i} \geq \lim_{i\to\infty} \frac{\log n_i}{l_i} = \lim_{i\to\infty} \frac{\log n_{i-1}!}{n_{i-1}l_{i-1}}$$

Then using the approximation $\sum_{k=1}^{n} \log k \geq \int_1^n \log t dt = n \log n - n + 1$ we have

$$\frac{\log n_{i-1}!}{n_{i-1}l_{i-1}} \geq \frac{n_{i-1}\log n_{i-1} - n_{i-1} + 1}{n_{i-1}l_{i-1}} > \frac{\log n_{i-1}}{l_{i-1}} - \frac{1}{l_{i-1}}.$$

Therefore, by repeating this,

$$\frac{\log n_{i-1}!}{n_{i-1}l_{i-1}} \geq \frac{\log n_1}{l_1} - \sum_{k=1}^{i-1} \frac{1}{l_k},$$

and

$$\mathrm{ent}(X) \geq \frac{\log n_1}{l_1} - \sum_{k=1}^{\infty} \frac{1}{l_k}.$$

Because l_i increases as the factorial, the sum on the right can be made arbitrarily small by choose l_1 large enough, and $\frac{\log n_1}{l_1}$ can be made arbitrarily close to 1 by choosing n_1 arbitrarily close to 2^{l_1}. Therefore, not only can the entropy be made positive, it can be made arbitrarily close to 1. □

Proposition 2.6.6. *The entropy of X can be computed as*

$$\mathrm{ent}(X) = \lim_{i\to\infty} \frac{\log n_i}{l_i}.$$

Proof. In the previous proposition, we saw that $\mathrm{ent}(X) \geq \lim_{i\to\infty} \frac{\log n_i}{l_i}$. For an upper bound, note that if σ is permitted in X with $|\sigma| = l_i$, then σ is a subword of some permitted τ which is a concatenation of elements of $\mathcal{E}_{i-1}$. Because $|\sigma| = n_{i-1}l_{i-1}$, σ intersects at most $n_{i-1}+1$ of the elements of $\mathcal{E}_{i-1}$ making up the concatenation, and σ starts at most l_i of the way into the first of them. Therefore, letting N_k be the number of permitted words of X of length k, we have the upper bound

$$N_{l_i} \leq n_{i-1}^{n_{i-1}+1} l_{i-1}$$

and so

$$\lim_{i\to\infty} \frac{\log N_{l_i}}{l_i} \leq \lim_{i\to\infty} \frac{n_{i-1}\log n_{i-1} + \log n_{i-1}l_{i-1}}{n_{i-1}l_{i-1}} \leq \lim_{i\to\infty} \frac{\log n_{i-1}}{l_{i-1}} + \frac{\log l_i}{l_i} = \lim_{i\to\infty} \frac{\log n_i}{l_i}.$$

□

The main result of this section is the calculation of the dimension spectrum of this subshift, which we show to be $[0, \mathrm{ent}(X)]$. The strategy is similar to the one used in the proof of Theorem 16, but it cannot be used as-is because X is not uniformly full. We explain why below.

Definition 2.6.6. *A word σ is* aligned *in X if σ is permitted in X and σ is an initial segment (as opposed to merely a subword) of some τ in some $\mathcal{E}_i$.*

For an aligned word of length N and any $l_i < N$, one may write a unique decomposition $N = al_i + b$ where $b < l_i$. In that case, we would say that σ has a-many l_i-blocks, which are the $\sigma[jl_i, (j+1)l_i]$ for $j < a$.

If a word τ is permitted but not aligned, we may still refer to its l_i-blocks, which will generally be of the form $\tau[jl_i + k, (j+1)l_i + k]$ and correspond to the l_i-blocks of an aligned block of which τ is a subword. Each l_i-block of a permitted word is an element of $\mathcal{E}_i$.

Now consider a string $\sigma = w_1 w_1 w_2 w_3 \ldots w_{n_{i-1}-1}$ where each $w_k \in \mathcal{E}_i$ and $j \neq k$ implies $w_j \neq w_k$. Then σ is permitted even though its first two l_i-blocks are the same, because it is possible that the l_{i+1}-block boundary lies between the first two l_i-blocks.

In fact, this is the only scenario where σ is permitted, so the only way it can be extended is with $w_{n_{i-1}}$, the last unused block of length l_i. But i was arbitrary, so for any N there is a σ which is permitted but for which $|\{y \upharpoonright N : \sigma y \in X\}| = 1$. Therefore, X is not uniformly full.

However, note that for this example, the number of bits beyond σ that are fully determined make up a very small fraction $\frac{1}{n_{i-1}}$ of the original length of σ. This phenomenon turns out to be true more generally, inspiring the following strategy for constructing elements of effective dimension t for every $t \in (0, \mathrm{ent}(X))$. We still build up a trajectory out of finite initial segments, but instead of adding a finite number of bits each time, we add bits in such a way as to increase the length of the initial segment by a fixed proportion, so that $|\sigma_{s+1}| \approx |\sigma_s|\frac{n+1}{n}$ where σ_s is the initial segment produced at stage s.

Most of the work goes into showing that when a sufficiently long string is extended by a fixed proportion of its length, it will always be possible to extend it in a way that causes $\frac{K(\sigma)}{|\sigma|}$ to increase by a fixed amount. This is the content of Lemma 2.6.7 below. Then by letting n go to infinity, the effect of any individual extension on the value of $\frac{K(\sigma)}{|\sigma|}$ is reduced and the effective and packing dimensions converge to the target t.

Lemma 2.6.7. *For every $n > 0$ and $\varepsilon > 0$, there is an N such that for all aligned σ permitted in X with $|\sigma| > N$, there is a τ for which $\sigma\tau$ is aligned and permitted in X, and $|\tau| = \lfloor \frac{\sigma}{n} \rfloor$ and*

$$\mathrm{ent}(X) - \frac{K(\tau|\sigma^*)}{|\tau|} < \varepsilon.$$

Proof. Let N be long enough to satisfy the following list of requirements, whose necessity will be seen in the course of the proof.

1. Because $\lim_{i\to\infty} \frac{\log n_i}{l_i} = \mathrm{ent}(X)$, let N be large enough that $|\frac{\log n_{i-1}}{l_{i-1}} - \mathrm{ent}(X)| < \frac{\varepsilon}{4}$ for all $l_i > N$.
2. Let N be large enough that $\left(\frac{16}{\varepsilon} + 2\right) n < (n_{i-1} - \frac{1}{n}) \left(\frac{n+1}{n}\right)$ for all $l_i > N$.
3. Let N be large enough that $\frac{8}{\varepsilon} \frac{n \log 3n}{l_{i-1}} < \frac{\varepsilon}{4}$ for all $l_i > N$.
4. Let N be large enough that $\frac{1+4n}{n_{i-1}} < \frac{\varepsilon}{4}$ for all $l_i > N$.

Let σ be aligned in X with $|\sigma| > N$. Then $|\sigma|$ can be written in a canonical way as

$$|\sigma| = a_i l_i + a_{i-1} l_{i-1} + \cdots + a_1 l_1$$

where $a_i > 0$ and each $a_k < n_k$. Because of the sufficiently large size of N, the numbers $l_i, n_i, l_{i-1}, n_{i-1}$ satisfy the relationships to n and ε enumerated above.

The analysis splits into two cases, depending on whether we will find an appropriate τ by appending blocks of size l_i or blocks of size l_{i-1}. Which we choose depends on the size of a_i. If $a_i > \left(\frac{16}{\varepsilon} + 2\right) n$, we will build τ out of blocks of length l_i. If $a_i < (n_{i-1} - \frac{1}{n}) \left(\frac{n+1}{n}\right)$, we will build τ out of blocks of length l_{i-1}. By condition (2) on the size of σ, for each value of a_i one of these cases will hold.

Case 1: Appending l_i-blocks

Assuming $a_i > \left(\frac{16}{\varepsilon} + 2\right) n$, we will now give a lower bound on the number of strings τ of length $\lfloor \frac{|\sigma|}{n} \rfloor$ for which $\sigma\tau$ is aligned in X. Because we wish to add l_i-blocks, we must start on an l_i-boundary, but σ may not end on one. Therefore, under-counting, we fix the first bits of τ to be some arbitrary way of extending σ to the nearest l_i-boundary. For example, there is a computable way to do this. Then we will count the number of different ways to choose the order of the subsequent l_i-blocks of τ, until its length is exhausted but for some number of bits again less than l_i. Because we will only count ways of laying l_i-blocks which stay inside the subshift, there will be a way to continue τ to its full length by choosing the last bits arbitrarily, again perhaps computably. Therefore, letting b be the number of l_i-blocks within the domain of τ, we have the following bounds:

$$\frac{a_i}{n} - 3 < b < \frac{a_i + 1}{n}.$$

The bound on the left takes into account losing up to l_i on the left and right sides of τ, minus one more to allow us to forget the values of a_k for $k < i$. The bound on the right is larger than $\frac{|\tau|}{l_i}$.

Now if $a_i + 1 + b < n_i \left(1 - \frac{1}{2n}\right)$, then the blocks we are adding are "far" from the end of the first l_{i+1}-block. Because of the restriction on repeating l_i-blocks within an aligned l_{i+1}-block, the number of choices per block decreases by one for each block added. But since we are far from the end, for each of b-many l_i-blocks chosen, there will be at least $\frac{n_i}{2n}$-many choices for each block, because there are at least that many choices for the last block. Therefore, a lower bound on the number of τ satisfying the conditions of the lemma is $\left(\frac{n_i}{2n}\right)^b$. Therefore, by counting there is a τ for which

$$K(\tau|\sigma^*) \geq b \log \frac{n_i}{2n}$$

and so

$$\begin{aligned} \frac{K(\tau : \sigma^*)}{|\tau|} &\geq \frac{b \log n_i - b \log 2n}{\frac{a_i+1}{n} l_i} \\ &= \left(\frac{bn}{a_i + 1}\right) \left[\frac{\log n_i}{l_i} - \frac{\log 2n}{l_i}\right] \\ &= \left(\frac{bn}{a_i + 1}\right) \frac{\log n_i}{l_i} - \frac{\varepsilon}{4} \end{aligned}$$

where the last replacement follows by constraint (3) on N and the fact that $\frac{bn}{a_i+1} < 1$.

Now on the other hand, if $a_i + 1 + b \geq n_i \left(1 - \frac{1}{2n}\right)$, then the blocks we are adding are near to the end of the first l_{i+1}-block. In fact, they might spill over in to the next l_{i+1}-block. If they do not spill over, then there are $\frac{(n_i - a_i - 1)!}{(n_i - a_i - b - 1)!}$ possibilities. If they do spill over, there are

$(n_i - a_i - 1)!\frac{n_i!}{(2n_i - a_i - 1 - b)!}$ possibilities. A lower bound on both these numbers is $b!$. Therefore, by counting, there is a τ for which

$$K(\tau|\sigma^*) \geq \log b! \geq b \log b - b$$

where the last inequality is obtained using an integral approximation to the logarithm. Now, because $a_i + 1 + b \geq n_i(1 - \frac{1}{2n})$ and $b < \frac{a_i+1}{n}$, we have

$$\begin{aligned}(a_i + 1)\frac{n+1}{n} &\geq n_i(1 - \frac{1}{2n}) \\ a_i &\geq \frac{n}{n+1} n_i(1 - \frac{1}{2n}) - 1 \geq \frac{n}{2(n+1)} n_i\end{aligned}$$

And then because $b > \frac{a_i}{n} - 3$, we have

$$b \geq \frac{a_i}{n} - 3 \geq \frac{n_i}{2(n+1)} - 3 \geq \frac{n_i}{3n}$$

Therefore,

$$\begin{aligned}\frac{K(\tau|\sigma^*)}{|\tau|} &\geq \frac{b \log b - b}{\frac{a_i+1}{n} l_i} \\ &\geq \frac{b \log \frac{n_i}{3n} - b}{\frac{a_i+1}{n} l_i} \\ &= \left(\frac{bn}{a_i+1}\right)\left(\frac{\log n_i}{l_i} - \frac{1 + \log 3n}{l_i}\right) \\ &\geq \left(\frac{bn}{a_i+1}\right)\frac{\log n_i}{l_i} - \frac{\varepsilon}{4}\end{aligned}$$

where the last replacement is justified by the condition (3) on N and the fact that $\frac{bn}{a_i+1} < 1$.

Therefore, in either case there is a τ for which

$$\begin{aligned}\operatorname{ent}(X) - \frac{K(\tau|\sigma^*)}{|\tau|} &< \left|\operatorname{ent}(X) - \frac{\log n_i}{l_i}\right| + \left(\frac{\log n_i}{l_i} - \frac{K(\tau|\sigma^*)}{|\tau|}\right) \\ &< \frac{\varepsilon}{4} + \left(\frac{\log n_i}{l_i} - \left(\frac{bn}{a_i+1}\right)\frac{\log n_i}{l_i} + \frac{\varepsilon}{4}\right) \\ &= \frac{\varepsilon}{2} + \left(\frac{\log n_i}{l_i}\right)\left(1 - \frac{bn}{a_i+1}\right) \\ &< \frac{\varepsilon}{2} + \frac{a_i + 1 - bn}{a_i + 1} \\ &< \frac{\varepsilon}{2} + \frac{a_i + 1 - a_i + 3n}{\left(\frac{16}{\varepsilon} + 2\right)n + 1} \\ &< \frac{\varepsilon}{2} + \varepsilon\frac{1 + 3n}{(16 + 2\varepsilon)n + \varepsilon} \\ &< \varepsilon.\end{aligned}$$

Case 2: Appending l_{i-1}-blocks

Assuming $a_i < (n_{i-1} - \frac{1}{n})\left(\frac{n+1}{n}\right)$, we will again bound the number of strings τ of the appropriate length. That restriction translates into $a_i + \frac{a_i+1}{n} < n_{i-1}$. In this case, we will add approximately $\frac{a_i n_{i-1}+a_{i-1}}{n}$-many length-$l_i$ blocks. As before, we will fill in the ends computably and just count full blocks. We will also waste one l_{i-1} block at the beginning, so letting $a = a_i n_{i-1} + a_{i-1}$, the total number b of blocks to be added will satisfy

$$\frac{a}{n} - 4 < b < \frac{a+1}{n}.$$

Since a_i-many l_i-length blocks have already been started, and we may start as many as $\frac{a_i+1}{n}$ more of them, we need some strategy to ensure that we do not repeat any. The simple strategy we use is to start each new l_i-block with an l_{i-1}-block that has never been used at the start of a previous l_i-block. For the very first l_{i-1}-block, which may come in the middle of an l_i-block which may be threatening to repeat a previous l_i-block, we will also choose it deterministically, so that it does not repeat an l_{i-1}-block that has stood in its position in previous l_i-blocks. This strategy works as long as $a_i + \frac{a_i+1}{n} < n_{i-1}$, so that we cannot run out of l_{i-1}-blocks.

Now let us count. Expressing b as $b = b_i n_{i-1} + b_{i-1}$, with $b_{i-1} < n_{i-1}$, notice that whenever we lay n_{i-1}-many l_{i-1}-blocks in a row (not necessarily on an l_i-boundary, though a total length of l_i is laid down), there are at least $(n_{i-1}-1)!$ ways to do this. (There would be $n_{i-1}!$ if we let the first block be chosen freely, but instead it is chosen according to the strategy described above.) And for laying b_{i-1}-many l_{i-1}-blocks in a row, in the worst case there are still at least $(b_{i-1}-1)!$ ways to do this (the very worst case being when the last block starts a new l_i-block). Therefore, the total number of options for the b-many l_{i-1}-blocks of τ is bounded by $(n_{i-1}-1)!^{b_i}(b_{i-1}-1)!$. By counting, there is a τ of the appropriate length such that

$$\begin{aligned} K(\tau|\sigma^*) &\geq b_i \log(n_{i-1}-1)! + \log(b_{i-1}-1)! \\ &\geq b_i(n_{i-1}-1)(\log(n_{i-1}-1)-1) + (b_{i-1}-1)(\log(b_{i-1}-1)-1) \\ &\geq b_i n_{i-1}\log\frac{n_{i-1}}{2} + b_{i-1}\log\frac{b_{i-1}}{2} - b_i n_{i-1} - b_i \log n_{i-1} - b_{i-1} - \log b_{i-1} \\ &\geq b_i n_{i-1}\log n_{i-1} + b_{i-1}\log b_{i-1} - 3b \end{aligned}$$

therefore

$$\frac{K(\tau|\sigma^*)}{|\tau|} \geq \frac{b_i n_{i-1}\log n_{i-1} + b_{i-1}\log b_{i-1}}{\frac{a+1}{n} l_{i-1}} - \frac{3b}{\frac{a+1}{n} l_{i-1}}$$

Now $\frac{3bn}{(a+1)l_{i-1}} < \frac{3}{l_{i-1}} < \frac{\varepsilon}{4}$. Furthermore, if $b < n_{i-1}$, then $b_i = 0$ and $b_{i-1} = b$, so in that case

$$\begin{aligned}\frac{K(\tau|\sigma^*)}{|\tau|} &\geq \frac{b\log b}{\frac{a+1}{n}l_{i-1}} - \frac{\varepsilon}{4} \\ &\geq \frac{bn}{a+1}\frac{\log\frac{a}{2n}}{l_{i-1}} - \frac{\varepsilon}{4} \\ &\geq \frac{bn}{a+1}\left(\frac{\log n_{i-1}}{l_{i-1}} - \frac{\log 2n}{l_{i-1}}\right) - \frac{\varepsilon}{4} \\ &\geq \frac{bn}{a+1}\frac{\log n_{i-1}}{l_{i-1}} - \frac{\varepsilon}{2}\end{aligned}$$

On the other hand, if $b \geq n_{i-1}$, then (using the convexity of $x\log x$)

$$\begin{aligned}\frac{K(\tau|\sigma^*)}{|\tau|} &\geq \frac{(b_i-1)n_{i-1}\log(n_{i-1}) + (n_{i-1}+b_{i-1})\log(\frac{n_{i-1}+b_{i-1}}{2})}{\frac{a+1}{n}l_{i-1}} \\ \frac{K(\tau|\sigma^*)}{|\tau|} &\geq \frac{(b_i-1)n_{i-1}\log(n_{i-1}) + (n_{i-1}+b_{i-1})\log(n_{i-1}) - (n_{i-1}+b_{i-1})}{\frac{a+1}{n}l_{i-1}} \\ &\geq \frac{bn}{a+1}\left(\frac{\log n_{i-1}}{l_{i-1}} - \frac{1}{l_{i-1}}\right) \\ &\geq \frac{bn}{a+1}\frac{\log n_{i-1}}{l_{i-1}} - \frac{\varepsilon}{2}\end{aligned}$$

Therefore, in either case there is a τ for which

$$\begin{aligned}\mathrm{ent}(X) - \frac{K(\tau|\sigma^*)}{|\tau|} &< \left|\mathrm{ent}(X) - \frac{\log n_i}{l_i}\right| + \left(\frac{\log n_i}{l_i} - \frac{K(\tau|\sigma^*)}{|\tau|}\right) \\ &< \frac{\varepsilon}{4} + \left(\frac{\log n_{i-1}}{l_{i-1}} - \left(\frac{bn}{a+1}\right)\frac{\log n_{i-1}}{l_{i-1}} + \frac{\varepsilon}{2}\right) \\ &= \frac{3\varepsilon}{4} + \left(\frac{\log n_{i-1}}{l_{i-1}}\right)\left(1 - \frac{bn}{a+1}\right) \\ &< \frac{3\varepsilon}{4} + \frac{a+1-bn}{a+1} \\ &< \frac{3\varepsilon}{4} + \frac{a+1-a+4n}{a_i n_{i-1} + a_{i-1}1} \\ &< \frac{3\varepsilon}{4} + \frac{1+4n}{n_{i-1}} \\ &< \varepsilon,\end{aligned}$$

where the last step is justified by condition (4) on N.

□

Now we can prove the main theorem of the section.

Theorem 17. *The minimal subshift X from Definition 2.6.5 has effective dimension spectrum* $[0, \mathrm{ent}(X)]$.

Proof. It is clear that X contains a computable element and by [32] it contains an element of effective dimension $\mathrm{ent}(X)$. So, fixing $t \in (0, \mathrm{ent}(X))$, we construct an element of X with effective dimension t by extending initial segments. All our initial segments will be aligned. We also build this element with packing dimension t. Fix $\varepsilon = \frac{\mathrm{ent}(X)-t}{4}$.

Let $n_0 = 1$. By Lemma 2.6.7, let N be such that whenever $|\sigma| > N$, for an aligned σ, then there is a τ with $|\tau| = |\sigma|$ such that $\sigma\tau$ is aligned and $\mathrm{ent}(X) - \frac{K(\tau|\sigma^*)}{|\tau|} < \varepsilon$.. Let N also be large enough that $\frac{2\log N+2C}{N} < \frac{t}{2}$. Let σ_0 be any aligned string of length at least N.

At each stage, we attempt to extend σ_s in such a way as to bring $\frac{K(\sigma_{s+1})}{|\sigma_{s+1}|}$ closer to t. If $\frac{K(\sigma_s)}{|\sigma_s|} \geq t$, we let σ_{s+1} be the result of computably extending σ for $\frac{|\sigma_s|}{n_s}$ bits. On the other hand, if $\frac{K(\sigma_s)}{|\sigma_s|} < t$, we choose τ by the guarantee of Lemma 2.6.7 with $n = n_s$ and $\varepsilon = \frac{\mathrm{ent}(X)-t}{4}$. This is possible because $|\sigma_s|$ will always great enough to satisfy the lemma for $n = n_s$, a fact which is normally guaranteed by setting $n_{s+1} = n_s$, and using the fact that $|\sigma|$ was large enough at the previous step. But we do set $n_{s+1} = n_s + 1$ whenever the following conditions are satisfied:

1. $|\sigma_{s+1}|$ is large enough to guarantee the existence of a τ with $|\tau| = \frac{|\sigma_{s+1}|}{n_s+1}$ for which $\mathrm{ent}(X) - \frac{K(\tau|\sigma^*)}{|\tau|} < \varepsilon$.

2. $|\sigma_{s+1}| > N$ for which $\dfrac{2\log\frac{N}{n_s+1} + 2C}{\frac{N}{n_s+1}} < \min(\dfrac{t}{2}, \dfrac{\mathrm{ent}(X)-t}{4})$

Since we add a large chunk at a time, it is necessary to verify that not only is $K(\sigma_{s+1})/|\sigma_{s+1}|$ approaching t, but any intermediate τ such that $\sigma_s \preccurlyeq \tau \preccurlyeq \sigma_{s+1}$ also has $\frac{K(\tau)}{|\tau|}$ bounded near t. So, first let us elaborate on why no individual step changes the compressive ratio by too much. Suppose we have a string σ and will add to it a string τ which is $1/n$ the length of σ. On the one hand, by Proposition 2.1.2,

$$\begin{aligned}
\frac{K(\sigma\tau)}{|\sigma\tau|} &\leq \frac{K(\sigma)+K(\tau)+C}{|\sigma\tau|} \\
&= \frac{K(\sigma)}{|\sigma|} + \frac{-\frac{1}{n}K(\sigma)+K(\tau)+C}{|\sigma|(\frac{n+1}{n})} \\
&\leq \frac{K(\sigma)}{|\sigma|} + \frac{1}{n+1}\left(\frac{K(\tau)}{|\tau|} - \frac{K(\sigma)}{|\sigma|}\right) + \frac{C}{|\sigma|} \\
&\leq \frac{K(\sigma)}{|\sigma|} + \frac{1}{n+1} + \frac{C}{|\sigma|}
\end{aligned}$$

And since this holds for all n, considering a smaller length τ is equivalent to increasing n, so this upper bound bounds the entire possible amount of deviation upwards in the compressive ratio that could be experienced when performing a single stage with proportion determined by n and starting from length $|\sigma|$.

For the other direction, by Proposition 2.1.4,

$$\begin{aligned}\frac{K(\sigma\tau)}{|\sigma\tau|} &\geq \frac{K(\sigma) - 2\log|\tau| - C}{|\sigma\tau|} \\ &= \frac{K(\sigma)}{|\sigma|} - \frac{\frac{1}{n}K(\sigma) + 2\log(|\tau|) + C}{|\sigma|\frac{n+1}{n}} \\ &\geq \frac{K(\sigma)}{|\sigma|} - \frac{1}{n+1} - \frac{2\log|\sigma| + C}{|\sigma|}\end{aligned}$$

So we see that as n and $|\sigma|$ grown large, the amount of deviation introduced over a single stage decreases to zero. Therefore, if $\lim_{s\to\infty}\frac{K(\sigma_s)}{|\sigma_s|} = t$, and $x = \cup_s \sigma_s$, then $\lim_{m\to\infty}\frac{K(x\upharpoonright m)}{m} = t$.

Next let us verify that if $\frac{K(\sigma_s)}{|\sigma_s|} > t$, then

$$\frac{K(\sigma_s)}{|\sigma_s|} - \frac{K(\sigma_{s+1})}{|\sigma_{s+1}|} \geq \frac{t}{2n_s + 2}.$$

Expressing σ_{s+1} as $\sigma\tau$, where $\sigma = \sigma_s$:

$$\begin{aligned}\frac{K(\sigma\tau)}{|\sigma\tau|} &\leq \frac{K(\sigma) + K(\tau|\sigma) + C}{|\sigma|\frac{n+1}{n}} \\ &\leq \frac{K(\sigma)}{|\sigma|} + \frac{1}{n+1}\left(\frac{K(\tau|\sigma) + C}{|\tau|} - \frac{K(\sigma)}{|\sigma|}\right) \\ &\leq \frac{K(\sigma)}{|\sigma|} + \frac{1}{n+1}\left(\frac{2\log|\tau| + 2C}{|\tau|} - \frac{K(\sigma)}{|\sigma|}\right)\end{aligned}$$

Because n_s is not increased until σ is long enough to satisfy condition (2) on the increase of n_s, $|\sigma|$ must be long enough that $\frac{2\log|\sigma|/n_s + 2C}{|\sigma|/n_s} < t/2$. Thus we may continue:

$$\frac{K(\sigma\tau)}{|\sigma\tau|} \leq \frac{K(\sigma)}{|\sigma|} + \frac{1}{n+1}\left(\frac{t}{2} - t\right)$$

So whenever $\frac{K(\sigma_s)}{|\sigma_s|} > t$, it will be decreased by at least $t/(2n_s + 2)$ until it becomes less than t again. Because this decrease may take many steps, it is possible that n_s will vary over the course of it, so that the decrease is not guaranteed by the same constant each stage. However, because n_s increases by at most one per stage, and in the worst case it is increased at every stage, in which case the resulting series would be harmonic, so it will eventually make it below t again.

Lemma 2.6.7 guarantees the increase in compressive ratio when $\frac{K(\sigma_s)}{|\sigma_s|} < t$. Therefore, by Proposition 2.1.5,, and letting $\sigma_{s+1} = \sigma_s\tau$,

$$\begin{aligned}\frac{K(\sigma_s\tau)}{|\sigma_s\tau|} &\geq \frac{K(\sigma_s) + K(\tau|\sigma_s^*) - 2\log|\tau| - 3C}{|\sigma_s|\frac{n+1}{n}} \\ &\geq \frac{K(\sigma_s)}{|\sigma_s|} + \frac{1}{n+1}\left(\frac{K(\tau|\sigma_s^*) - 2\log|\tau| - 3C}{|\tau|} - \frac{K(\sigma_s)}{|\sigma_s|}\right)\end{aligned}$$

Now we claim that

$$\mathrm{ent}(X) - \frac{K(\tau|\sigma_s^*) - 2\log|\tau| - 3C}{|\tau|} < \frac{\mathrm{ent}(X) - t}{2}.$$

This follows because

$$\begin{aligned}\mathrm{ent}(X) - \frac{K(\tau|\sigma_s^*) - 2\log|\tau| - 3C}{|\tau|} &\leq \left(\mathrm{ent}(X) - \frac{K(\tau|\sigma_s^*)}{|\tau|}\right) + \frac{2\log|\tau| + 3C}{|\tau|} \\ &\leq \frac{\mathrm{ent}(X) - t}{4} + \frac{\mathrm{ent}(X) - t}{4}.\end{aligned}$$

Therefore, we may continue as before to conclude that if $\frac{K(\sigma_s)}{|\sigma_s|} < t$, then at each stage s the ratio will increase by at least

$$\begin{aligned}\frac{1}{n_s+1}\left(\frac{K(\tau|\sigma_s^*) - 2\log|\tau| - 3C}{|\tau|} - \frac{K(\sigma_s)}{|\sigma_s|}\right) &\geq \frac{1}{n_s+1}\left(\mathrm{ent}(X) - \frac{K(\sigma_s)}{|\sigma_s|} - \frac{\mathrm{ent}(X) - t}{2}\right) \\ &\geq \frac{1}{n_s+1}\left((\mathrm{ent}(X) - t) - \frac{\mathrm{ent}(X) - t}{2}\right) \\ &= \frac{\mathrm{ent}(X) - t}{2n_s+2}\end{aligned}$$

Again, even though n_s may increase over time, it does so slowly enough that the ratio will eventually surpass t again. Therefore, $\lim_{s\to\infty} \frac{K(\sigma_s)}{|\sigma_s|} = t$. This completes the proof. □

Bibliography

[1] C. J. Ash and J. Knight. *Computable structures and the hyperarithmetical hierarchy.* Vol. 144. Studies in Logic and the Foundations of Mathematics. Amsterdam: North-Holland Publishing Co., 2000, pp. xvi+346. ISBN: 0-444-50072-3.

[2] Henk Bruin. “Dimensions of recurrence times and minimal subshifts”. In: *Dynamical systems (Luminy-Marseille, 1998).* World Sci. Publ., River Edge, NJ, 2000, pp. 117–124.

[3] Douglas Cenzer and Jeffrey B. Remmel. “Index sets for computable differential equations”. In: *MLQ Math. Log. Q.* 50.4-5 (2004), pp. 329–344. ISSN: 0942-5616. DOI: 10.1002/malq.200310102. URL: http://dx.doi.org/10.1002/malq.200310102.

[4] John D. Clemens. “Isomorphism of subshifts is a universal countable Borel equivalence relation”. In: *Israel J. Math.* 170 (2009), pp. 113–123. ISSN: 0021-2172. DOI: 10.1007/s11856-009-0022-0. URL: http://dx.doi.org/10.1007/s11856-009-0022-0.

[5] Bruno Durand, Leonid A. Levin, and Alexander Shen. “Complex tilings”. In: *J. Symbolic Logic* 73.2 (2008), pp. 593–613. ISSN: 0022-4812. DOI: 10.2178/jsl/1208359062. URL: http://dx.doi.org/10.2178/jsl/1208359062.

[6] Russell A. Gordon. *The integrals of Lebesgue, Denjoy, Perron, and Henstock.* Vol. 4. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 1994, pp. xii+395. ISBN: 0-8218-3805-9.

[7] Noam Greenberg, Antonio Montalbán, and Theodore A. Slaman. “The Slaman-Wehner theorem in higher recursion theory”. In: *Proc. Amer. Math. Soc.* 139.5 (2011), pp. 1865–1869. ISSN: 0002-9939. DOI: 10.1090/S0002-9939-2010-10693-7. URL: http://dx.doi.org/10.1090/S0002-9939-2010-10693-7.

[8] Christian Grillenberger. “Constructions of strictly ergodic systems. I. Given entropy”. In: *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 25 (1972/73), pp. 323–334.

[9] A. Grzegorczyk. “On the definitions of computable real continuous functions”. In: *Fund. Math.* 44 (1957), pp. 61–71. ISSN: 0016-2736.

[10] William Hanf. “Nonrecursive tilings of the plane. I”. In: *J. Symbolic Logic* 39 (1974), pp. 283–285. ISSN: 0022-4812.

[11] Peter Hertling and Christoph Spandl. “Shifts with decidable language and non-computable entropy”. In: *Discrete Math. Theor. Comput. Sci.* 10.3 (2008), pp. 75–93. ISSN: 1365-8050.

[12] Denis Hirschfeldt and Asher Kach. “Shift complex sequences”. Talk presented at the AMS Eastern Sectional Meeting, 2012.

[13] Michael Hochman and Tom Meyerovitch. “A characterization of the entropies of multidimensional shifts of finite type”. In: *Ann. of Math. (2)* 171.3 (2010), pp. 2011–2038. ISSN: 0003-486X. DOI: `10.4007/annals.2010.171.2011`. URL: `http://dx.doi.org/10.4007/annals.2010.171.2011`.

[14] Alexander S. Kechris and W. Hugh Woodin. “Ranks of differentiable functions”. In: *Mathematika* 33.2 (1986), 252–278 (1987). ISSN: 0025-5793. DOI: `10.1112/S0025579300011244`. URL: `http://dx.doi.org/10.1112/S0025579300011244`.

[15] Mushfeq Khan. “Shift-complex sequences”. In: *Bull. Symbolic Logic* 19.2 (2013), pp. 199–215. ISSN: 1079-8986. DOI: `10.2178/bsl.1902020`. URL: `http://dx.doi.org/10.2178/bsl.1902020`.

[16] Daniel Lacombe. “Extension de la notion de fonction récursive aux fonctions d’une ou plusieurs variables réelles. II, III”. In: *C. R. Acad. Sci. Paris* 241 (1955), pp. 13–14, 151–153.

[17] Steffen Lempp. “Hyperarithmetical index sets in recursion theory”. In: *Trans. Amer. Math. Soc.* 303.2 (1987), pp. 559–583. ISSN: 0002-9947. DOI: `10.2307/2000684`. URL: `http://dx.doi.org/10.2307/2000684`.

[18] Ming Li and Paul Vitányi. *An introduction to Kolmogorov complexity and its applications*. Third. Texts in Computer Science. Springer, New York, 2008, pp. xxiv+790. ISBN: 978-0-387-33998-6. DOI: `10.1007/978-0-387-49820-1`. URL: `http://dx.doi.org/10.1007/978-0-387-49820-1`.

[19] Douglas Lind and Brian Marcus. *An introduction to symbolic dynamics and coding*. Cambridge University Press, Cambridge, 1995, pp. xvi+495. ISBN: 0-521-55124-2; 0-521-55900-6. DOI: `10.1017/CBO9780511626302`. URL: `http://dx.doi.org/10.1017/CBO9780511626302`.

[20] Jack H. Lutz. “Gales and the constructive dimension of individual sequences”. In: *Automata, languages and programming (Geneva, 2000)*. Vol. 1853. Lecture Notes in Comput. Sci. Springer, Berlin, 2000, pp. 902–913. DOI: `10.1007/3-540-45022-X_76`. URL: `http://dx.doi.org/10.1007/3-540-45022-X_76`.

[21] Elvira Mayordomo. “A Kolmogorov complexity characterization of constructive Hausdorff dimension”. In: *Inform. Process. Lett.* 84.1 (2002), pp. 1–3. ISSN: 0020-0190. DOI: `10.1016/S0020-0190(02)00343-5`. URL: `http://dx.doi.org/10.1016/S0020-0190(02)00343-5`.

[22] S. Mazurkiewicz. “Über die Menge der differenzierbaren Funktionen.” In: *Fund. Math.* 27 (1936), pp. 244–249.

[23] Joseph S. Miller. "Two notes on subshifts". In: *Proc. Amer. Math. Soc.* 140.5 (2012), pp. 1617–1622. ISSN: 0002-9939. DOI: 10.1090/S0002-9939-2011-11000-1. URL: http://dx.doi.org/10.1090/S0002-9939-2011-11000-1.

[24] Shahar Mozes. "Tilings, substitution systems and dynamical systems generated by them". In: *J. Analyse Math.* 53 (1989), pp. 139–186. ISSN: 0021-7670. DOI: 10.1007/BF02793412. URL: http://dx.doi.org/10.1007/BF02793412.

[25] Dale Myers. "Nonrecursive tilings of the plane. II". In: *J. Symbolic Logic* 39 (1974), pp. 286–294. ISSN: 0022-4812.

[26] J. Myhill. "A recursive function, defined on a compact interval and having a continuous derivative that is not recursive". In: *Michigan Math. J.* 18 (1971), pp. 97–98. ISSN: 0026-2285.

[27] Marian B. Pour-El and J. Ian Richards. *Computability in analysis and physics*. Perspectives in Mathematical Logic. Berlin: Springer-Verlag, 1989, pp. xii+206. ISBN: 3-540-50035-9.

[28] Raphael M. Robinson. "Undecidability and nonperiodicity for tilings of the plane". In: *Invent. Math.* 12 (1971), pp. 177–209. ISSN: 0020-9910.

[29] Gerald E. Sacks. *Higher recursion theory*. Perspectives in Mathematical Logic. Berlin: Springer-Verlag, 1990, pp. xvi+344. ISBN: 3-540-19305-7.

[30] Stanisław Saks. *Theory of the integral*. Second revised edition. English translation by L. C. Young. With two additional notes by Stefan Banach. Dover Publications, Inc., New York, 1964, pp. xv+343.

[31] Stephen Simpson. "Medvedev degrees of 2-dimensional subshifts of finite type". In: *Ergodic Theory Dynam. Systems* (), to appear.

[32] Stephen Simpson. "Symbolic dynamics: entropy = dimension = complexity". In: *Theory Comput. Syst.* (), to appear.

[33] Stephen G. Simpson. *Subsystems of second order arithmetic*. Second. Perspectives in Logic. Cambridge: Cambridge University Press, 2009, pp. xvi+444. ISBN: 978-0-521-88439-6. DOI: 10.1017/CBO9780511581007. URL: http://dx.doi.org/10.1017/CBO9780511581007.

[34] Robert I. Soare. *Recursively enumerable sets and degrees*. Perspectives in Mathematical Logic. A study of computable functions and computably generated sets. Berlin: Springer-Verlag, 1987, pp. xviii+437. ISBN: 3-540-15299-7.

[35] Clifford Spector. "Recursive well-orderings". In: *J. Symb. Logic* 20 (1955), pp. 151–163. ISSN: 0022-4812.

[36] Sean Walsh. "Denjoy integration: descriptive set theory and model theory". In: *Fund. Math.* (), to appear.